



Universidad Nacional del Nordeste
Facultad de Ciencias Exactas y Naturales y
Agrimensura

Maestría en Tecnologías de la Información

Trabajo Final

**Implementación de Ingeniería de Explotación
de la Información para la Identificación de
Patrones que contribuyan a la toma de
decisiones en una unidad de cuidados
intensivos coronarios**

Autora: Sofia J. Vallejos

Directora: Dra. María Inés Pisarello

Co-Director: Dr. Emanuel Irrazábal

Año 2.021

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

No fue fácil el camino para llegar hasta donde estoy, pero gracias a su apoyo, a su amor incondicional y acompañamiento, lo difícil se hizo más fácil. Es por ello que dedico esta tesis a, mi hermosa familia.

Resumen

Uno de los desafíos que los médicos de la unidad de cuidados intensivos coronarios (UCIC) enfrentan diariamente, es manejar e interpretar un gran volumen de datos dispersos en las historias clínicas electrónicas (HCE). En este camino, es necesario, gestionar y aumentar el valor de los datos e impulsar las capacidades analíticas de los médicos. En el presente trabajo se analizará e implementará un Modelo de Procesos para una Ingeniería de Explotación de Información (MoProPEI) con el objetivo de identificar los patrones de comportamiento y las características distintivas de pacientes de la unidad de cuidados intensivos coronarios. Por ello, el análisis e implementación de las tecnologías de Ingeniería de Explotación de la Información se realizarán sobre una base de datos extraída de las HCE de una institución médica especializada en cardiología con el fin de colaborar en la transformación de la información en conocimiento para que contribuya en la toma de decisiones, capacidades

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

operativas e institucionales y la realización de acciones correctivas correspondientes.

Palabras claves: *Explotación de la información, Minería de Datos, Historia Clínica Electrónicas, Unidad Coronaria, Informática Médica.*

Abstract

One of the major challenges that physicians in charge of the Coronary Intensive Care Unit (CICU) face daily, is the fact that they need to handle, read and interpret a great volume of scattered data in patient's Electronic Medical Records (EMR). So, it becomes necessary to manage stored data, increase its value and boost the analytical capabilities of physicians. In this thesis, a Process Model applied to Information Mining Projects (MoProPEI - Modelo de Procesos para Proyectos de Explotación de Información) will be analyzed and implemented in order to identify behavior patterns and distinctive characteristics of patients in Coronary Intensive Care Unit. Therefore, the analysis and implementation of IEI Technology will be

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

carried out on a database extracted from the EMRs of a medical institution specialized in cardiology, in order to collaborate in the transformation of information into knowledge so that it contributes in decision making processes, as well as in operational and institutional capacities and the implementation of corresponding corrective actions.

Keywords: *Information Mining, Data Mining, Electronic Medical Records, Coronary Unit, Medical Informatics.*

Agradecimientos

A mis padres Nieves y Romy quienes con su amor, paciencia y esfuerzo me han permitido cumplir hoy un sueño más, gracias por inculcar en mí el ejemplo de esfuerzo y valentía.

A mi esposo Matías, por darme su amor, por la confianza y no dejarme caer, por darme el mayor tesoro que tenemos, ese rayito de luz y esperanza que es nuestra Hija Agustina

A mi hija Agustina por transformar mi vida, por alimentar de amor mi corazón y por deleitar mis oídos escuchando esa vocecita diciéndome mamá, porque eres la razón de ser una mejor persona todos los días, mi vida es hermosa por tu existencia.

A mi hermano Gastón y cuñada Jesica que con sus palabras me hacen sentir orgullosa de lo que soy, por compartir momentos significativos conmigo y por siempre estar dispuestos a escucharme y ayudarme en cualquier momento.

Al Dr. Pedro Cassani, a través de sus consejos y confianza, me permite creer aún más en mí, empoderando a la mujer SIEMPRE, su apoyo y presencia, un pilar clave en mi vida.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

A mi directora y codirector de tesis, Dra. María Inés Pisarello y Dr. Emanuel Irrazábal, quienes con la enseñanza de sus valiosos conocimientos hicieron que pueda crecer día a día como profesional, gracias por su generosidad, paciencia, y dedicación.

Al Ing. Fabián Bobadilla, de la Empresa Aliare, por el tiempo brindado y el material facilitado.

Al Dr. Jorge Parras, por su gran aporte desinteresado, horas de trabajo, paciencia y amor a su profesión.

A mis docentes de la Maestría en Tecnologías de la Información que, con su sabiduría, conocimiento y apoyo, me motivaron a desarrollarme como persona y profesional.

A mis amigos, siempre me han prestado un gran apoyo moral y humano, en especial al Lic. Darío Báez, a quien le debo su incondicionalidad, por facilitarme los caminos para seguir, sin pedir nada a cambio y sin dudar de mi capacidad.

A Dios por haberme dado la sabiduría y fuerza guiándome en el trayecto de mi vida.

Índice de contenidos

<i>1. Introducción</i>	2
1.1. Objetivos.....	4
1.1.1. Objetivo general.....	4
1.1.2. Objetivos específicos	5
1.2. Fundamentación.....	6
1.3. Organización del trabajo.....	7
1.4. Escenario	8
1.5. Producción Científica	10
1.5.1. Actividades de Divulgación Científica	10
1.5.2. Expositora en panel de Ciencia de Datos.....	10
<i>2. Estado de la cuestión</i>	12
2.1. Unidad de cuidados intensivos	12
2.1.1. Historias Clínicas Electrónicas	16
2.2. Ciencia de Datos	17
2.2.1. Informática Medica	17
2.2.2. Big Data y la informática Medica	19
2.3. Explotación de Información	21
2.3.1. Minería de Datos en el área de la Salud.....	22

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

2.4.	Ingeniería de Explotación de la información...	24
3.	<i>Ingeniería de Explotación de Información</i>	28
3.1.	Modelo de Proceso para Ingeniería de Explotación de Información	28
3.1.1.	Sub-Proceso: Gestión.....	31
3.1.2.	Sub-Proceso: Desarrollo	31
3.2.	Requerimientos en Proyectos de Explotación de Información	34
3.3.	Viabilidad y Estimación de Proyectos de Explotación de Información	41
3.3.1.	Propuesta del Modelo para la Evaluación de la Viabilidad	41
3.3.2.	Modelo de Estimación para Proyectos de Explotación de Información	45
3.4.	Propuesta de explotación de información.....	46
3.4.1.	Descubrimiento de Reglas de Comportamiento	47
3.4.2.	Descubrimiento de Grupos	49
3.4.3.	Descubrimiento de Reglas de Pertenencia a Grupos	50
3.5.	Descripción de las herramientas	51
3.5.1.	WEKA.....	52

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

3.5.2.	Tanagra	53
4.	<i>Modelo de Proceso Propuesto</i>	56
4.1.	Problemática	56
4.1.1.	Pregunta de investigación	58
4.1.2.	Caso de Estudio.....	58
4.1.3.	Propuesta.....	58
4.2.	<i>Subproceso de Gestión (G)</i>	71
4.2.1.	Iniciación del Proyecto (G.IP)	71
4.2.1.1.	Exploración Inicial del Proyecto (G.IP.EIP)	72
4.2.1.1.1.	Recursos Humanos Involucrados (G.IP.EIP.ReHI)	73
4.2.1.1.2.	Riesgos del Proyecto (G.IP.EIP.RiPr).....	74
4.2.1.1.3.	Plan de Contingencias (G.IP.EIP.PICo) ...	75
4.2.1.2.	Definición de la Comunicación (G.IP.DCo)	76
4.2.1.2.1.	Plan de Comunicación (G.IP.DCo.PCom)	77
4.2.1.3.	Evaluación de la Situación (G.IP.EvS)	78
4.2.1.3.1.	Reporte de Evaluación de Herramientas (G.IP.EvS.EvHe)	79
4.2.1.3.2.	Reporte de Evaluación de Viabilidad (G.IP.EvS.REVi)	83

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.2.1.4.	Definición del Ciclo de Vida (G.IP.DCV)	87
4.2.1.4.1.	Modelo de Ciclo de Vida (G.IP.DCV.MCV)	88
4.2.2.	Planificación (G.Pl)	89
4.2.2.1.	Planificación de la Mediciones (G.Pl.PIM)	90
4.2.2.1.1.	Listado de Métricas (G.Pl.PIM.LiMe)	91
4.2.2.1.2.	Estimación del Proyecto (G.Pl.PIM.EsPr)	92
4.2.2.2.	Planificación de las Actividades (G.Pl.PIA)	100
4.2.2.2.1.	Mapa de Actividades (G.Pl.PIA.MaAc)	100
4.2.2.2.2.	Plan de Acción (G.Pl.PIA.PIaC)	104
4.2.2.3.	Planificación de los Recursos (G.Pl.PIR)	109
4.2.2.3.1.	Plan de Necesidad de Recursos (G.Pl.PIR.PINR)	110
4.2.2.4.	Planificación de las Responsabilidades (G.Pl.PRe)	112
4.2.2.4.1.	Matriz de Responsabilidades (G.Pl.PRe.MaRe)	113
4.2.2.4.2.	Propuesta del Proyecto (G.Pl.PRe.PrPr)	117
4.2.3.	Soporte (G.So)	118

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.2.3.1.	Mediciones del Proyecto (G.So.MeP)....	119
4.2.3.1.1.	Registro de Mediciones (G.So.MeP.ReMe)	119
4.2.3.2.	Gestión de la Configuración (G.So.GeC)	121
4.2.3.2.1.	Reporte de Versionado (G.So.GeC.ReVe)	121
4.2.3.2.2.	Informe del Estado de la Configuración (G.So.GeC.InEC).....	123
4.2.4.	Control (G.Co)	127
4.2.4.1.	Gestión del Desarrollo (G.Co.GeD)	128
4.2.4.1.1.	Reporte de Estado (G.Co.GeD.ReEs)	128
4.2.4.2.	Control de las Actividades (G.Co.CoA).	129
4.2.4.2.1.	Registro de Riesgos Acontecidos (G.Co.CoA.ReRA)	130
4.2.4.3.	Gestión del Cambio (G.Co.GCa)	131
4.2.4.3.1.	Reporte de Evaluación del Cambio (G.Co.Gca.RECa)	132
4.2.5.	Cierre (G.Ci)	133
4.2.5.1.	Formalización Externa del Cierre del Proyecto (G.Ci.FEC)	133

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.2.5.1.1. Documento de Aceptación (G.Ci.FEC.DoAc).....	134
4.2.5.2. Formalización Interna del Cierre del Proyecto (G.Ci.FIC).	135
4.2.5.2.1. Reporte de Cierre (G.Ci.FIC.ReCi).....	136
4.3. Subproceso Desarrollo (D).....	138
4.3.1. Entendimiento del Negocio (D.EN).....	138
4.3.1.1. Análisis del negocio (D.EN.AnN).....	139
4.3.1.1.1. Fuentes de Información del Cliente (D.EN.AnN.FuIC)	140
4.3.1.1.2. Definiciones, Acrónimos y Abreviaturas (D.EN.AnN.DAAb).....	141
4.3.1.1.3. Objetivo del Proyecto (D.EN.AnN.ObPr) 144	
4.3.1.1.4. Criterios de éxito del proyecto (D.EN.AnN.CrEP).....	145
4.3.1.1.5. Expectativas del proyecto (D.EN.AnN.ExPr).....	146
4.3.1.1.6. Suposiciones del proyecto (D.EN.AnN.SuPr).....	147
4.3.1.1.7. Restricciones del proyecto (D.EN.AnN.RePr).....	149

4.3.1.2. Comprensión del Problema de Negocio (D.EN.CPN).....	150
4.3.1.2.1. Problema del Negocio (D.EN.CPN.PrNe)	151
4.3.1.2.2. Criterios de Éxito del Problema de Negocio (D.EN.CPN.CEPN)	152
4.3.2. Entendimiento de los Datos (D.ED)	153
4.3.2.1. Análisis de los Datos (D.ED.AnD)	153
4.3.2.1.1. Diccionario de Fuentes de Datos (D.ED.AnD.DiFD)	156
4.3.2.1.2. Campos Relacionados con el Problema de Negocio (D.ED.AnD.CRPN).....	158
4.3.2.2. Exploración de los Datos (D.ED.ExD) ..	160
4.3.2.2.1. Fuente Integrada de datos (D.ED.ExD.FuID).....	162
4.3.2.2.2. Reporte de Datos Explorados (D.ED.ExD.ReDE)	163
4.3.2.3. Evaluación de los Datos (D.ED.EvD)	165
4.3.2.3.1. Reporte de calidad de los datos (D.ED.EvD.ReCD)	166
4.3.3. Modelado de los Datos (D.Mo)	167
4.3.3.1. Modelado del problema (D.Mo.MoP)....	167

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.3.3.1.1. Diseño del Proceso de Explotación de Información (D.Mo.MoP.DPEI).....	168
4.3.3.2. Configuración del Modelo (D.MO.CoM)	171
4.3.3.2.1. Selección de Algoritmos de Explotación de Información (D.Mo.CoM.SAEI)	172
4.3.3.2.2. Selección de Variables del Modelo (D.Mo.CoM.SeVM)	175
4.3.3.2.3. Estrategias de Evaluación de Modelos (D.Mo.CoM.EsEM).....	176
4.3.4. Preparación de los Datos (D.PD)	177
4.3.4.1. Construcción de la Fuente Temporal de Datos (D.PD.CFT).....	178
4.3.4.1.1. Fuente Temporal de Datos (D.PD.CFT.FuTD).....	179
4.3.4.1.2. Reporte de Generación de la Fuente Temporal de datos (D.PD.CFT.RGFT)	179
4.3.4.2. Adecuación de la Fuente Temporal de Datos (D.PD.AFT).....	182
4.3.4.2.1. Reporte de Adecuación de la Fuente Temporal de Datos (D.PD.AFT.RAFT)	182
4.3.5. Implementación (D.Im).....	191
4.3.5.1. Selección del Modelo (D.Im.SeM)	191

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.3.5.1.1. Reporte de Estrategia de Parametrización del Modelo (D.Im.SeM.REPM)	192
4.3.5.2. Explotación de la Información (D.Im.ExI)	193
4.3.5.2.1. Reporte de Implementación del Modelo (D.Im.ExI.ReIM)	194
4.3.5.2.2. Patrones de Conocimiento (D.Im.ExI.PaCo)	195
4.3.6. Evaluación y Presentación (D.EP)	199
4.3.6.1. Evaluación de los Resultados (D.EP.EvR)	199
4.3.6.1.1. Reporte de Evaluación de los Resultados (D.EP.EvR.ReER)	200
4.3.6.2. Presentación de los Resultados (D.EP.PrR)	203
4.3.6.2.1. Reporte del Proyecto (D.EP.PrR.RepP) .	204
5. <i>Conclusión y Futuras Líneas de Investigación</i>	210
5.1. <i>Conclusiones</i>	210
5.2. <i>Futuras Líneas de Investigación</i>	212
6. <i>Referencias</i>	214

Índice de figuras

<i>Fig. 1: Estructura General - MoProPEI: Fuente [1]. ...</i>	<i>30</i>
<i>Fig. 2: Estructura del subproceso: Gestión – MoProPEI. Fuente [1].....</i>	<i>33</i>
<i>Fig. 3: Estructura del subproceso: Desarrollo – MoProPEI. Fuente [1].</i>	<i>34</i>
<i>Fig. 4: Representación de la Función de Pertenencia y asignación de Intervalo Difuso para los Valores Lingüísticos. [39].</i>	<i>43</i>
<i>Fig. 5: Cálculo de estimación de esfuerzo. Fuente [37].</i>	<i>46</i>
<i>Fig. 6: Esquema y subproductos resultantes de aplicar TDIDT al descubrimiento de reglas de comportamiento: Fuente [18].</i>	<i>48</i>
<i>Fig. 7: Esquema de proceso: Descubrimiento de grupos: Fuente [18].</i>	<i>49</i>
<i>Fig. 8: Esquema y subproductos resultantes de SOM y TDIDT aplicados al descubrimiento de reglas de pertenencia a grupos: Fuente[18].</i>	<i>51</i>
<i>Fig. 9: Interfaz de Usuario. WEKA.</i>	<i>53</i>
<i>Fig. 10: Tanagra GUI.</i>	<i>54</i>
<i>Fig. 11: Fuente de Diagrama de Datos.</i>	<i>163</i>
<i>Fig. 12: Fuente Temporal de datos.</i>	<i>179</i>
<i>Fig. 13: Patrones de Conocimiento (Tanagra).A.</i>	<i>196</i>
<i>Fig. 14: Patrones de Conocimiento. B.</i>	<i>196</i>

Índice de tablas

<i>Tabla 1: Plantilla de Definiciones, Acrónimos y Abreviaturas.</i>	<i>36</i>
<i>Tabla 2: Plantilla de Objetivos del Proyecto.</i>	<i>37</i>
<i>Tabla 3: Plantilla de Criterios de Éxito del Proyecto. ..</i>	<i>38</i>
<i>Tabla 4: Plantilla de Expectativas del Proyecto.</i>	<i>38</i>
<i>Tabla 5: Plantilla de Suposiciones del Proyecto.</i>	<i>39</i>
<i>Tabla 6: Plantilla de Restricciones del Proyecto.</i>	<i>40</i>
<i>Tabla 7: Características a ser evaluadas por el método propuesto [38].</i>	<i>42</i>
<i>Tabla 8: Fórmulas utilizadas por el modelo propuesto [39].</i>	<i>44</i>
<i>Tabla 9: G. Estructura del Subproceso Gestión.</i>	<i>60</i>
<i>Tabla 10: D. Estructura del Subproceso Desarrollo.</i>	<i>64</i>
<i>Tabla 11: G. Recursos Humanos Involucrados.</i>	<i>74</i>
<i>Tabla 12: G. Riesgo del Proyecto.</i>	<i>75</i>
<i>Tabla 13: G. Plan de Contingencias.</i>	<i>76</i>
<i>Tabla 14: G. Plan de Comunicación.</i>	<i>77</i>
<i>Tabla 15: G. Reporte de Evaluación de herramientas.</i>	<i>80</i>
<i>Tabla 16: Preguntas asociadas a la caracterización.</i>	<i>84</i>
<i>Tabla 17: G. Reporte de Evaluación de Viabilidad.</i>	<i>85</i>
<i>Tabla 18: Traducción y cálculo de intervalos por dimensión.</i>	<i>86</i>
<i>Tabla 19: Evaluación de la situación.</i>	<i>87</i>
<i>Tabla 20: G. Definición del ciclo de vida.</i>	<i>89</i>
<i>Tabla 21: G. Listado de Métricas.</i>	<i>91</i>

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

<i>Tabla 22: Tipo de objetivo de explotación de información.</i>	<i>93</i>
<i>Tabla 23: Grado de apoyo de los miembros de la organización.</i>	<i>94</i>
<i>Tabla 24: Cantidad y tipo de los repositorios de datos disponibles.</i>	<i>95</i>
<i>Tabla 25: Cantidad de tuplas disponibles en la tabla principal.</i>	<i>95</i>
<i>Tabla 26: Cantidad de tuplas disponibles en las tablas auxiliares.</i>	<i>96</i>
<i>Tabla 27: Nivel de conocimiento sobre los datos.</i>	<i>97</i>
<i>Tabla 28: Nivel de conocimiento y experiencia del equipo de trabajo.</i>	<i>97</i>
<i>Tabla 29: Funcionalidad de las herramientas disponibles</i>	<i>98</i>
<i>Tabla 30: G. Estimación del Proyecto</i>	<i>99</i>
<i>Tabla 31: G. Mapa de Actividades.</i>	<i>101</i>
<i>Tabla 32: G. Plan de Acción.</i>	<i>105</i>
<i>Tabla 33: G. Plan de Necesidad de Recursos.</i>	<i>111</i>
<i>Tabla 34: G. Matriz de Responsabilidades.</i>	<i>114</i>
<i>Tabla 35: G. Propuesta del Proyecto.</i>	<i>118</i>
<i>Tabla 36: G. Registro de Mediciones.</i>	<i>120</i>
<i>Tabla 37: G. Reporte de Versionado.</i>	<i>123</i>
<i>Tabla 38: G. Informe del Estado de la Configuración.</i>	<i>124</i>
<i>Tabla 39: G. Reporte de Estado.</i>	<i>129</i>
<i>Tabla 40: G. Registro de Riesgos Acontecidos.</i>	<i>131</i>
<i>Tabla 41: G. Reporte de Evaluación del Cambio</i>	<i>132</i>
<i>Tabla 42: G. Documento de Aceptación.</i>	<i>134</i>

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

<i>Tabla 43: G. Reporte de Cierre</i>	<i>137</i>
<i>Tabla 44: D. Fuentes de Información del Cliente.....</i>	<i>140</i>
<i>Tabla 45: D. Definiciones, Acrónimos y Abreviaturas.</i>	<i>141</i>
<i>Tabla 46: D. Objetivos del Proyecto.....</i>	<i>145</i>
<i>Tabla 47: D. Criterio de éxito del Proyecto.....</i>	<i>146</i>
<i>Tabla 48: D. Expectativa del Proyecto.</i>	<i>147</i>
<i>Tabla 49: D. Suposición del Proyecto.</i>	<i>148</i>
<i>Tabla 50: D. Restricciones del Proyecto.</i>	<i>149</i>
<i>Tabla 51: D. Problema del Negocio.</i>	<i>152</i>
<i>Tabla 52: D. Criterio de Éxito del Problema del Negocio.....</i>	<i>153</i>
<i>Tabla 53: D. Diccionario de Fuente de Datos.....</i>	<i>156</i>
<i>Tabla 54: D. Campos Relacionados con el Problema de Negocio.....</i>	<i>159</i>
<i>Tabla 55: D. Reporte de Datos Explorados.....</i>	<i>164</i>
<i>Tabla 56: D. Reporte de calidad de los datos.....</i>	<i>166</i>
<i>Tabla 57: D. Diseño del Proceso de Explotación de Información</i>	<i>170</i>
<i>Tabla 58: D. Selección de Algoritmos de Explotación de Información</i>	<i>175</i>
<i>Tabla 59: D. Selección de Variables del Modelo.....</i>	<i>176</i>
<i>Tabla 60: D. Estrategias de Evaluación de Modelos... </i>	<i>177</i>
<i>Tabla 61: D. Reporte de Generación de la Fuente Temporaria de datos.</i>	<i>180</i>
<i>Tabla 62: D. Reporte de Adecuación de la Fuente Temporaria de Datos.</i>	<i>185</i>
<i>Tabla 63: D. Reporte de Estrategia de Parametrización del Modelo.....</i>	<i>192</i>

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

<i>Tabla 64: D. Reporte de Implementación del Modelo.</i>	<i>194</i>
<i>Tabla 65: D. Patrones de Conocimiento.....</i>	<i>197</i>
<i>Tabla 66: D. Reporte de Evaluación de los Resultados</i> <i>.....</i>	<i>203</i>
<i>Tabla 67: D. Reporte del Proyecto.</i>	<i>205</i>

Índice de siglas

A

AA: Aprendizaje Automático, 18

AREP: Cantidad y tipo de los repositorios de datos disponibles, 96

D

DER: Diagrama Entidad Relación, 164

E

EV: Valor Global de la Viabilidad del Proyecto, 45

F

FEy: Fracción de eyección, 145

G

GUI: Interfaz Gráfica de Usuario, 55

H

HCE: historias clínicas electrónicas, 3

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

I

IAM: Infarto Agudo de Miocardio, 15

IEI: Ingeniera de explotación de información, 2

K

KEXT: Nivel de conocimiento y experiencia del equipo de trabajo, 98

KLDS: Nivel de conocimiento sobre los datos, 98

L

LECO: Grado de apoyo de los miembros de la organización, 96

M

MCV: Modelo de Ciclo de Vida, 90

MoProPEI: Modelo de Proceso para una Ingeniería de Explotación de Información, 5

O

OBTY: Tipo de objetivo de explotación de información, 95

Q

QTUA: Cantidad de tuplas disponibles en las tablas auxiliares, 98

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

QTUM: Cantidad de tuplas disponibles en la tabla principal, 97

S

SI: Sistemas de Información, 48

SOM: Self-Organizing Maps, 50

SUH: Servicio de Urgencias Hospitalario, 16

T

TDIDT: Top Down Induction Decision Trees, 49

TOOL: Funcionalidad de las herramientas disponibles, 100

U

UCI: Unidad de Cuidados Intensivos, 14

UCIC: unidad de cuidados intensivos coronarios, 2

V

Vd: Valor numérico de la Dimensión d para el Modelo de Viabilidad., 46

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Capítulo 1

Introducción

En este capítulo, se presenta la argumentación necesaria para el desarrollo del trabajo de investigación. Se plantean los objetivos generales y específicos. A continuación, se desarrollan la fundamentación, la organización del trabajo y el escenario. Por último, se exponen los aportes de conocimientos producidos por esta tesis.

1. Introducción

La calidad de las condiciones de salud, es una necesidad que debe prevalecer a medida que surgen nuevos avances y al mismo tiempo nuevas amenazas a la misma. Esta investigación se centra en implementar Ingeniería de Explotación de Información (IEI) para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios (UCIC), esto mediante técnicas y modelado de procesos.

Hoy en día los usos de la minería de datos resuelven o predicen problemáticas en el entorno digital, hacen uso de información de los individuos quienes con sus datos

alimentan las bases que a su vez son el insumo de los análisis.

Los médicos de cuidados intensivos coronarios utilizan su experiencia, conocimiento, juicio clínico y las historias clínicas electrónicas (HCE) para valorar, planificar, implementar y evaluar la atención a los pacientes de acuerdo a los procesos de las distintas internaciones. Por ello surge la necesidad de aplicar una metodología para el desarrollo de proyectos de ingeniería de explotación de información, para evitar: aprender conocimientos que no son verdaderos y aprender conocimiento poco útil. El objetivo para el cual se desarrolla un proyecto de explotación de información, es la generación de conocimiento para dar soporte a la toma de decisiones, el riesgo que conlleva basar dichas decisiones sobre conocimientos no válidos, y puede provocar grandes pérdidas a la organización [1].

1.1. Objetivos

1.1.1. Objetivo general

En el presente trabajo se analizará e implementará Ingeniería de Explotación de Información con el objetivo de identificar los patrones de comportamiento y las características distintivas de pacientes de la unidad de cuidados intensivos coronarios. El análisis e implementación de las tecnologías de IEI se realizará sobre una base de datos extraída de una institución médica especializada en cardiología con el fin de colaborar en la transformación de la información en conocimiento para que contribuya en la toma de decisiones, capacidades operativas e institucionales y la realización de acciones correctivas correspondientes.

Al implantar IEI se logrará un uso pleno y la diseminación de las bases de datos de pacientes de la unidad coronaria, con el objetivo de contar con sistemas que sean a la vez más eficientes y dar soporte a los profesionales de salud en la toma de decisiones y contribuir a mejorar la interacción entre la evidencia científica y la información del paciente.

1.1.2. Objetivos específicos

Para cumplimentar el objetivo general definido en el apartado anterior, se han propuesto la realización de determinados objetivos específicos, que lo componen y se indican a continuación.

- Investigar acerca del marco teórico sobre la Ingeniería de Explotación de la Información aplicados en las unidades de cuidados intensivos coronarios.
- Implementar el modelo MoProPEI como guía de modelado del proceso de desarrollo del proyecto de ingeniería de explotación de información en donde se identifiquen las fases, actividades, tareas, técnicas de representación y procedimientos de ejecución.
- Presentar los resultados donde los datos estén organizados y sintetizen la información de los pacientes, la que resulta necesaria en el momento crítico al momento de la toma de decisiones en la unidad de cuidados intensivos coronarios.

1.2. Fundamentación

Uno de los desafíos que los médicos tienen que enfrentar es el gran volumen o cantidad de grupos de datos dispersos en diferentes bases. El proceso de reconocimiento de patrones y la minería de datos para los volúmenes grandes de información son ejemplos de ellos, optimizar la cantidad de veces que se recorre el conjunto de datos, disminuye el tiempo de procesamiento.

Es necesario el avance en el conocimiento en las historias clínicas electrónicas de UCIC. En este camino, es necesario gestionar, aumentar el valor de los datos e impulsar las capacidades analíticas de los médicos. Por ello, se propone la implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios que permita la visualización de información relevante para apoyar la toma de decisiones. Esto tiene por objetivo proveer información inteligente para colaborar con la correcta gestión de las camas hospitalarias básicas, intermedias y críticas. A su vez, la implementación

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

permitirá predecir cuáles son las posibles variables predictorias para la obtener la eficacia en esos servicios.

1.3. Organización del trabajo

El trabajo investigativo se estructuró en etapas organizadas de manera secuenciada que permitieran cumplir los objetivos propuestos.

Dando lugar primero a la investigación acerca del marco teórico sobre la unidad de cuidados intensivos, las historias clínicas electrónicas y la Ingeniería de Explotación de la Información. Posteriormente se aplicó MoProPEI al caso de estudio propuesto, considerando tanto los aspectos de gestión y desarrollo con el fin de generar piezas de conocimiento que sirvan como soporte para la toma de decisiones mediante su estructuración jerárquica de tres niveles, cada uno de los cuales presenta un mayor grado de especificidad en los objetivos de sus elementos. Se muestran los resultados del modelo implementado, y a través de una discusión con el médico se compararon los resultados arrojados.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Finalmente, se establecen las conclusiones a partir de los hallazgos encontrado y en función de lo planteado se exponen las posibles líneas de investigación para futuras investigaciones.

1.4. Escenario

El Proyecto de investigación en el que se inserta la propuesta corresponde a dos grupos de Investigación consolidados, pertenecientes a la Facultad de Ciencias Exactas y Naturales y Agrimensura, de la UNNE, conforman el equipo de trabajo.

El Grupo De Ingeniería Biomédica, viene desarrollando desde hace más de una década sistemas computacionales especialmente diseñados para el procesamiento de biodatos. La producción es amplia e involucra trabajos publicados en revistas, congresos, capítulos de libro, todos ellos en ámbitos internacionales y nacionales. Cabe mencionar, también la co-dirección de proyectos de Investigación acreditados por la Secretaría General de Ciencia y Técnica de la Universidad Nacional del Nordeste e incluidos en el Programa Nacional de Incentivos; la

dirección y co-dirección de tesinas de grado y tesis de posgrado en el área y la formación de recursos, alumnos de grado y becarios de investigación.

El proyecto vigente dentro del cual se enmarca el plan de tesis aquí presentado es el PI 18F004 “Procesamiento Digital de Biopotenciales”, dirigido por el Dr. Jorge Emilio Monzón y co-dirigido por la Dra. María Inés Pisarello.

El grupo de investigación GICS-UNNE, se enfoca en el estudio y la aplicación de estándares, métodos y herramientas para contribuir con la calidad de software, tanto del proceso como del producto. En este marco, se ha realizado el estudio particular de procesos de explotación de información, y generado un sistema de apoyo a la decisión en el ámbito académico, como aporte a la analítica académica.

El grupo cuenta con publicaciones en revistas y congresos, y realiza una intensa actividad de formación de recursos humanos, mediante la dirección de tesis de Doctorado/Maestría/Especialidades y trabajos de fin de

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

carrera. El co-director de la tesis, Dr Emanuel Irrazabal pertenece a este grupo mencionado.

De los grupos mencionados anteriormente el plan de tesis propuesto es parte sustancial de este convenio.

1.5. Producción Científica

S. Vallejos, E. Báez, F. Bobadilla, M. Pisarello, “Sistemas de apoyo a la toma de decisiones clínicas (SADC) en una Unidad de Cuidados Intensivos Coronarios”, V Congreso Argentino de Ingeniería, XI Congreso Argentino de la Enseñanza de la Ingeniería, 2021.

1.5.1. Actividades de Divulgación Científica

S. Vallejos, E. Báez, F. Bobadilla, G. Dapozo, E. Irrazabal, J. Monzón M. Pisarello, “Sistema de Soporte a las Decisiones Clínicas en UCI”, 6° Edición, Jornadas de Articulación para la Innovación, 2020.

1.5.2. Expositora en panel de Ciencia de Datos

S. Vallejos, E. Báez, M. Pisarello, “Sistema de Soporte a las Decisiones Clínicas (SSDC)”, Datos: Nuevas oportunidades de desarrollo. Ciencia-Empresa, UNRaf, 2021.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Capítulo 2

Estado de la cuestión

2. Estado de la cuestión

En este capítulo se presenta la información relevante, tanto a nivel médico como a nivel informático, que se utilizan a través del presente documento y en los cuales se fundamenta el desarrollo del proyecto.

2.1. Unidad de cuidados intensivos

La medicina crítica se ha especializado en el cuidado y manejo de pacientes en estado crítico o en riesgo de desarrollarlo, atención que no puede proveerse en las salas regulares del hospital. Históricamente, el cuidado intensivo nace en la década de 1850 ante la necesidad de tener salas especiales para la atención médica en situaciones de guerra; luego, en la década de 1920 resurge ante la urgencia de contar con salas exclusivas para el cuidado postoperatorio. La terapia intensiva, a su vez, emerge en la década de 1940, cuando se aplican dispositivos y procedimientos específicos en pacientes con disfunción única, como ventilación mecánica, hemodiálisis y desfibriladores eléctricos. Finalmente, los cuidados críticos corresponden a la etapa contemporánea e

incluyen el monitoreo multiparámetro automatizado para el manejo de pacientes con deterioro multiorgánico, exámenes complementarios, dispositivos para el sostén básico y avanzado a la cabecera del enfermo y un equipo clínico multidisciplinario. La medicina crítica como disciplina se refiere a la ciencia del monitoreo y manejo del paciente crítico [2].

El paciente crítico es definido por la Sociedad Americana de Medicina Intensiva, como aquel que se encuentra fisiológicamente inestable, que requiere soporte vital avanzado y una evaluación clínica estrecha con ajustes continuos de terapia según evolución [3].

La unidad de cuidados intensivos (UCI) es sin duda el lugar que está dotado del personal humano y la infraestructura adecuada para cumplir con los requerimientos de cuidado anteriormente destos. El análisis de la forma en la que funciona la Red de Atención de Salud, muestra que muchos pacientes críticamente enfermos son evaluados y manejados inicialmente en

unidades fuera de la UCI, desde el mundo pre-hospitalario, servicios de urgencia (SU) y recuperaciones quirúrgicas.

Las atenciones en la UCI han mejorado con la llegada de los protocolos de estandarización en las terapias de pacientes críticos y con la generación de unidades de accidente vascular cerebral o de dolor torácico por fuera de la unidad. Cabe considerar por otra parte que, los recursos técnicos y humanos disponibles en estas unidades externas a las UCIs son insuficientes para manejar pacientes críticos más allá de su estabilización inicial. De este modo, y debido a la carencia de camas críticas, muchos pacientes deben permanecer fuera de las UCIs por tiempos prolongados [4].

La UCIC en los inicios se caracterizaba por tratar el infarto agudo de miocardio (IAM). Posteriormente incluyeron el manejo y la prevención de la disfunción miocárdica, consecuencia de la enfermedad coronaria. Además, introdujeron categorías clínicas y hemodinámicas que contribuyeron a discriminar grupos de pacientes con IAM

y precisar mejor el pronóstico y el manejo de los mismos [5].

La permanencia excesiva en espacios propios del servicio de urgencias hospitalario (SUH) de pacientes que ya han sido ingresados en servicios de hospitalización es una de las principales causas evitables de saturación del SUH.

Las dos vías de acceso a una cama hospitalaria son la admisión programada o la admisión urgente. La admisión urgente proviene de la totalidad del SUH y en un alto porcentaje se debe a una patología médica. En cambio, la mayoría de pacientes que ingresan programados son tributarios de realizar algún procedimiento quirúrgico, y es excepcional el ingreso programado por causa de una patología médica. La gestión adecuada de camas hospitalarias debería garantizar todos los días del año un número suficiente de camas para la admisión urgente sin necesidad de suspender las admisiones programadas [6].

2.1.1. Historias Clínicas Electrónicas

La historia clínica electrónica (HCE) es un registro de información de salud que reside en un sistema electrónico conformada por datos personales, familiares y del entorno que recopila de manera sistemática, automatizada, y detallada la información para el cuidado del paciente. Su implementación permitió la automatización de tareas, interoperabilidad en las pruebas de laboratorio y pruebas diagnósticas, identificación de factores de riesgo de la población y mayor seguridad en cuanto a prescripción farmacológica [7].

El valor potencial del procesamiento de los datos en medicina ha sido reconocido por años y la informática médica ha surgido como un campo de investigación interdisciplinaria importante en medicina. Avances en la tecnología de las HCE han hecho posible que reemplacen muchas funciones de la tradicional historia clínica en papel y su uso promete logros significativos en el cuidado del paciente

2.2. Ciencia de Datos

La ciencia de datos es la disciplina que implica el estudio de las fuentes de información, representaciones, procesamiento y finalmente conversión de los datos en conocimiento. Esta conversión es una parte integral del desarrollo empresarial y de otras decisiones estratégicas. La ciencia de datos se puede denominar como una disciplina híbrida que involucra matemáticas, estadística e informática. Específicamente, desde el dominio de la informática, implica técnicas de aprendizaje automático (AA), minería de datos y visualización o representación [8].

En [9] denomina a la ciencia de datos como la ciencia y el arte de usar métodos computacionales para identificar y descubrir patrones influyentes en los datos. El objetivo es obtener información a partir de los datos y, a menudo, influir en las decisiones para hacerlas más confiable.

2.2.1. Informática Médica

La informática médica (IM) es la ciencia aplicada de la gestión de datos de pacientes con el fin de mejorar la

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

comprensión de la salud y las biociencias. Es, por su naturaleza, una ciencia multidisciplinaria con interacciones en varios campos. La informática médica fue estática hasta la invención de la primera generación de computadoras digitales en la década de 1940. Desde entonces ha desempeñado un papel cada vez más importante en la atención de la salud y, como nueva disciplina académica, actúa como puente entre las ciencias médicas y de la información [10].

En [11] se recopiló información sobre el estado actual de las aplicaciones de minería de datos empleadas por medio de técnicas de aprendizaje automático en el campo de la informática médica. De modo que se evidencia el éxito de varias técnicas de aprendizaje automático para tomar decisiones significativas, apoyadas en el dominio de la informática médica. Además, aporta a la comprensión de como la extracción de conocimiento ofrecen beneficios en varios entornos de atención médica.

2.2.2. Big Data y la informática Medica

El término Big Data a menudo se refiere a la enorme cantidad de información digital que las empresas y organizaciones gubernamentales recopilan sobre personas y cosas, las interacciones y operaciones de sistemas en físico o entornos de vida virtual [12].

El uso de big data ha aumentado en los campos de la salud, la seguridad, los negocios, la educación y las redes sociales. En el área de salud los investigadores han comenzado a examinar cómo se pueden almacenar mejor los datos, analizarlos, y utilizar métodos para extraer información y conocimientos [13] , [14].

En [15] consideran que, dentro del mundo de Big Data existe un conocimiento oculto que podría cambiar la vida de un paciente o, en gran medida, cambiar el mundo en sí mismo. Además, declaran que las tecnologías de ciencia de datos definitivamente abrirán nuevas oportunidades y permitirán avances relacionados, entre otros, con el análisis de datos sanitarios abordando diferentes perspectivas: descriptiva, para responder a lo sucedido;

diagnóstico, para responder a la razón por la que sucedió; predictiva, para comprender lo que sucederá y/o, para detectar cómo podemos hacer que suceda.

Big Data en salud se pueden utilizar para mejorar la eficiencia y la eficacia de las estrategias de predicción y prevención o de las intervenciones médicas, los servicios de salud y las políticas de salud. El acceso a datos relacionados con la salud bien seleccionados y de alta calidad probablemente tendrá una serie de beneficios en una amplia gama de situaciones. En la práctica clínica, estos datos mejorarán los resultados de los pacientes individuales mediante la personalización de las predicciones, el diagnóstico más temprano, mejores tratamientos y un mejor apoyo a las decisiones de los médicos en los procesos cíclicos [15].

En [16] plantean que los desafíos relacionados con el análisis de datos, para los investigadores básicos como los clínicos es la necesidad de nuevas herramientas computacionales para mejorar el acceso a los datos y facilitar el análisis de datos para facilitar la toma de

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

decisiones en la clínica. Los médicos necesitan nuevas herramientas que rastreen y brinden retroalimentación rápida para la atención individual del paciente. Los investigadores necesitan herramientas que puedan adaptarse a diferentes conjuntos de datos y análisis.

En el dominio de la salud en su conjunto: médicos, pacientes, administración, y políticas: pueden beneficiarse significativamente de los avances actuales en ciencia de datos y, en particular, en análisis de datos.

2.3. Explotación de Información

La Explotación de Información es la subdisciplina Informática que contribuye a la Inteligencia de Negocio, aporta las herramientas necesarias para la transformación de información en conocimiento. Éste conocimiento es desconocido previamente y puede ser apropiado para la toma de decisiones en una organización [17]. Para un experto, o para el encargado de un sistema de información, regularmente no son los datos en sí lo más importante, sino el conocimiento que se encierra en sus relaciones, fluctuaciones y dependencias [18] y [19]. La explotación

de información se ha definido como la búsqueda de patrones interesantes y de regularidades importantes en grandes volúmenes de información [20]. Al hablar de explotación de información basada en sistemas inteligentes se refiere específicamente a los datos en sí lo más importante, sino el conocimiento que se encierra en sus relaciones, fluctuaciones y dependencias [21].

2.3.1. Minería de Datos en el área de la Salud

El termino minería de datos está relacionado a los conceptos de bases de datos y a la utilización de algoritmos para búsqueda de patrones de conocimiento en grandes volúmenes de datos. Teniendo en cuenta que hoy existen líneas de investigación en temas como: minería de textos, minería de imágenes, minería de patrones en flujos de información, minería en la web, entre otras. Por ello, es conviene utilizar el término “explotación de información” como referencia genérica a cualquiera de los tipos de minería mencionado anteriormente [1].

La minería de datos se ha utilizado para encontrar una gran cantidad de patrones de descubrimiento de datos entre

investigadores, ya que proporciona varios beneficios a la industria de la salud.

Diferentes autores han representado su investigación explorando diversas técnicas de Minería de Datos que incluyen clasificación, agrupamiento, árboles de decisión en diferentes campos de la salud [8], [22], [23].

En la era actual, millones de personas padecen algún tipo de enfermedad cardíaca cada año. La enfermedad cardíaca se ha convertido en una de las principales causas de tanto para hombres como para mujeres alrededor del mundo. El diagnóstico médico juega un papel muy importante para diagnosticar cualquier tipo de enfermedad cardíaca. La aplicación de técnicas de minería para clasificar diferentes tipos de enfermedades en función de los registros de los pacientes es importante para la industria médica. Diferentes algoritmos pueden ser aplicados para la detección de diferentes tipos de enfermedades [24].

2.4. Ingeniería de Explotación de la información

La Inteligencia de Negocio ofrece un abordaje interdisciplinario, el cual se centra en generar conocimiento que asista a la toma de decisiones de gestión y generación de planes estratégicos en las organizaciones, tomando todos los recursos disponibles para recabar información y usar herramientas analíticas y de síntesis con capacidad de transformar la información en conocimiento [25].

Con base en que la Ingeniería de Software ha sido definida en el SWEBOK [26] como: “la aplicación de un enfoque sistemático, disciplinado y cuantificable al desarrollo, operación y mantenimiento de software, y el estudio de estos enfoques, es decir, la aplicación de la ingeniería al software”; en [27] define a la Ingeniería de Explotación de Información “como la aplicación de un enfoque sistemático, disciplinado y cuantificable al desarrollo de proyectos de explotación de información, y el estudio de este enfoque, es decir, la aplicación de la ingeniería a la explotación de información”.

La IEI entiende en los procesos y las metodologías utilizadas para: ordenar, controlar y gestionar la tarea de encontrar patrones de conocimiento en masas de información [28].

Los proyectos de IEI tienen características diferentes a los proyectos de desarrollo de software tradicional [29], sobre todo en la parte operativa del proyecto. La discrepancia se presenta en los procesos de desarrollo y mantenimiento en los cuales las fases del ciclo de un proyecto de software tradicional: inicio, requisitos, análisis y diseño, construcción, integración y pruebas no resultan naturales en un proyecto de explotación de información [30].

La ingeniería de Explotación de Información (en ocasiones también referida como Minería de Datos o Extracción de Conocimiento) es la sub-disciplina de los Sistemas de Información que aporta a la Inteligencia de Negocio [31] las herramientas necesarias para la transformación de información en conocimiento [32]. Es decir, la búsqueda de patrones distintivos en grandes masas de información

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

[33] y [34], las cuales pueden estar almacenadas en distintos medios y formatos [35].

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Capítulo 3

Ingeniería de Explotación de la Información para la identificación de patrones

3. Ingeniería de Explotación de Información

En este capítulo se desarrolla el modelo de Procesos de Explotación de la información, la Viabilidad y Estimación de Proyectos de Explotación de Información, el Modelo de Estimación para Proyectos de Explotación de Información y Propuesta de explotación de información que se va a utilizar para dar respuesta a la problemática propuesta (ver Capítulo 4, sección 4.1, página 56). Por último, se describe la herramienta de minería de dato a utilizar.

3.1. Modelo de Proceso para Ingeniería de Explotación de Información

MoProPEI es un proceso para proyectos de explotación de información, en [35, p. 68] especifica la necesidad:

“surge a partir del análisis de las problemáticas identificadas en las propuestas existentes y las dificultades experimentadas en la implementación de las mismas, el cual introduce la visión de control y gestión de proyectos en el proceso de extracción de piezas de conocimiento para dar soporte al proceso de toma de decisiones”.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

En [1] se describe el modelo de procesos de explotación de información (MoProPEI), en donde detalla los aspectos generales asociados al modelo, introduce la estructura general del mismo y describe detalladamente los subprocesos que integran al modelo: el subproceso Gestión, y el subproceso Desarrollo. El modelo de proceso cuenta con una estructura jerárquica dividido en cuatro niveles: Subprocesos, Fases, Actividades y Tareas; cada uno de los cuales presenta un mayor nivel de especificidad.

El modelo como se mencionó previamente, se compone de dos subprocesos principales son la Gestión, enfocado la administración y control del proyecto; y el Desarrollo, encauzado a las actividades técnicas. En la Fig. 1 muestra cuales son los componentes generales del proceso y la vinculación externa entre los dos subprocesos.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

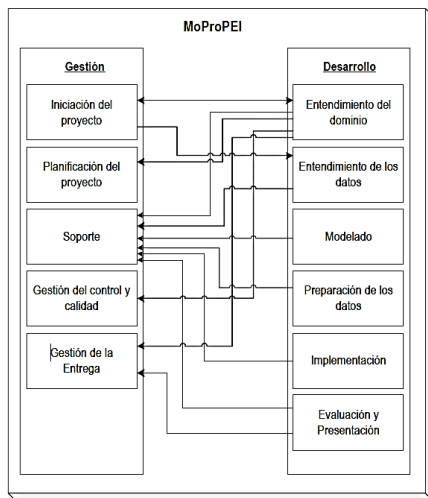


Fig. 1: Estructura General - MoProPEI: Fuente [1].

El primer subproceso (Gestión) consta de cinco fases: Iniciación del Proyecto, Planificación del Proyecto, Soporte, Gestión del Control y Calidad, y Gestión de la Entrega. Y el segundo, está compuesto por seis fases: Entendimiento del Dominio, Entendimiento de los Datos, Modelado, Preparación de los Datos, Implementación, y Evaluación y Presentación. Las ejecuciones de dichos subprocesos se realizan en paralelo, siendo el primero soporte para las actividades principales del segundo [1].

Se debe tener en cuenta, que la ejecución de los subprocesos no es secuencial, sino que la misma debe realizarse de forma paralela, debido a que el subproceso de gestión provee de soporte a las actividades pertenecientes al subproceso de Desarrollo considerando las destinadas a la construcción del producto final [35].

3.1.1. Sub-Proceso: Gestión

El subproceso de gestión abarca todas aquellas actividades relacionadas a la interacción inicial con el cliente, la estructuración del proyecto, planificación y administración de los recursos, control y calidad del desarrollo del proyecto [1]. Las cinco fases que contiene el subproceso Gestión son: Iniciación del Proyecto, Planificación del Proyecto, Soporte, Gestión del Control y Calidad, y Gestión de la Entrega.

3.1.2. Sub-Proceso: Desarrollo

El proceso de desarrollo contiene todas las actividades referidas a la identificación de patrones relevantes para la organización, de calidad y novedosas, así como su análisis y comprensión para la generación de piezas de

conocimiento validables e interesantes que aporten valor agregado a la misma [1]. Como afirma Martins en [35] que la estrategia de implementación del proyecto, define la estructuración y ejecución del subproceso, de acuerdo a los alcances, intereses, conocimiento y comprensión del dominio y del problema, entre otras cosas.

Las seis fases que contiene el subproceso Desarrollo son: Entendimiento del Dominio, Entendimiento del Negocio, Modelado, Preparación de los Datos, Implementación y Evaluación.

Como se mencionó en la Fig. 1 (ver sección 3.1 de la página 30) presenta las relaciones y dependencias desde la perspectiva de las fases, abstrayéndose del elemento específicamente vinculado y las asociaciones internas entre las actividades pertenecientes a una misma fase.

En las Fig. 2 y Fig. 3 se desarrolla las distintas actividades que componen al subproceso de gestión y al subproceso de desarrollo. La lectura de las figuras es la siguiente, en la parte izquierda se advierten las fases que componen al subproceso, y a su derecha en la misma posición horizontal

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

cada una de las actividades que las integran, junto con sus elementos de entrada y salida.

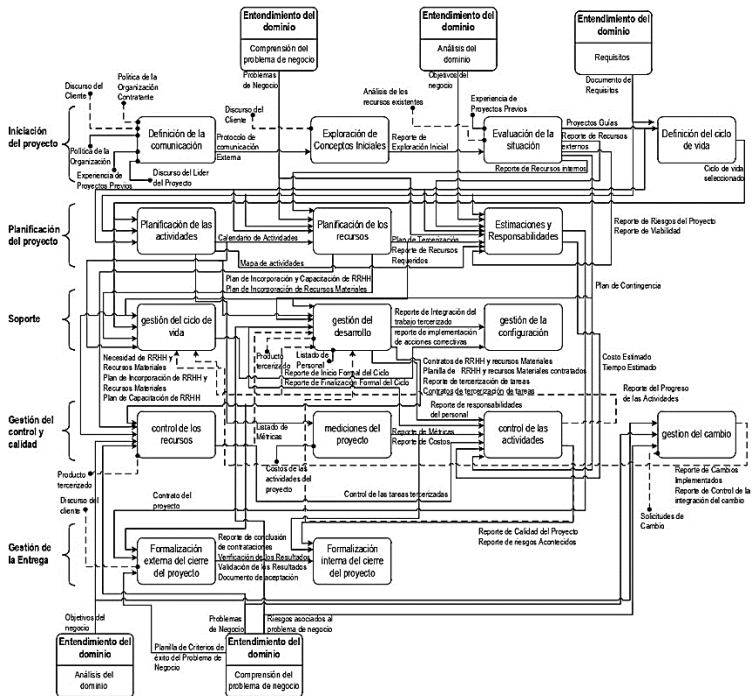


Fig. 2: Estructura del subproceso: Gestión – MoProPEI. Fuente [1].

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

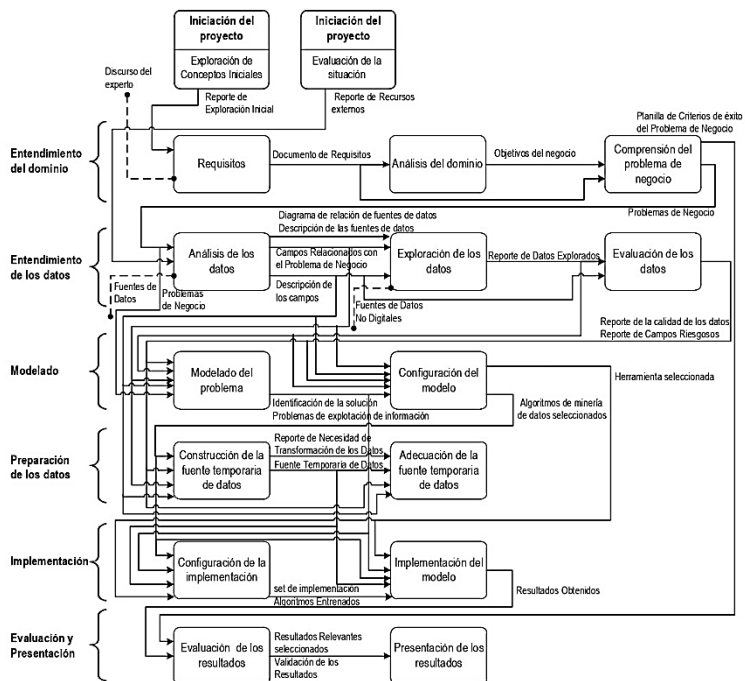


Fig. 3: Estructura del subproceso: Desarrollo – MoProPEI. Fuente [1].

3.2. Requerimientos en Proyectos de Explotación de Información

Dentro de las fases del Modelo de Procesos para Proyectos de Explotación de Información se realiza la especificación de los requerimientos. Es por esto, que se utilizará el formalismo propuesto por Britos en [36], donde se

identifica a los miembros de interés de la organización solicitante, cuyo conocimiento u opinión sea relevante para el desarrollo del proyecto, así como aquellas personas que integran al equipo que llevarán a cabo el proyecto (tanto pertenecientes a la organización desarrolladora, como externos contratados). Para cada individuo identificado se incorpora una descripción del mismo (rol, área a la que pertenece e información descriptiva del área de pericia y de interés para el proyecto) y su información de contacto [35].

A continuación, se describen brevemente cada uno de los conceptos de la actividad Comprensión del Problema de Negocio correspondiente al subproceso desarrollo, los cuales pueden ser deducidos [36]:

- Definiciones, Acrónimos y Abreviaturas: Se identifican las definiciones, acrónimos y abreviaturas, para establecer el vocabulario que será utilizado por todas las personas relacionadas con el proyecto de Explotación de Información. La plantilla que se confecciona para

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

documentar las definiciones, acrónimos y abreviaturas se exhibe en la Tabla 1.

Tabla 1: Plantilla de Definiciones, Acrónimos y Abreviaturas.

DEFINICIONES, ACRONIMOS Y ABREVIATURAS			
Responsable	<nombre del analista>	Fecha	<fecha de confección>
ID	<identificador del proyecto>		
Termino	Descripción	Tipo	Referencia
<nombre del término>	<breve explicación del significado del término>	<tipología: acrónimo/ definición / abreviatura>	<documentos y sesiones de adquisición de conocimientos que proporcionen la información acerca de ese término>

– **Objetivos del Proyecto:** Se identifican los objetivos del proyecto de Explotación de Información y su motivación para poder determinar los requisitos del cliente. La plantilla que se confecciona para documentar los objetivos del proyecto se exhibe en la Tabla 2.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 2: Plantilla de Objetivos del Proyecto.

OBJETIVOS DEL PROYECTO			
Responsable	<nombre del analista>	Fecha	<fecha de confección>
ID	<identificador del proyecto>		
ID Objetivo	Descripción	Referencia	
<identificador secuencial para cada objetivo>	<descripción breve del objetivo>	<documentos y sesiones de adquisición de conocimientos que proporcionen la información acerca del objetivo>	

- Criterios de Éxito del Proyecto: Se identifican los criterios en términos de logros esperados para convertir el proyecto de Explotación de Información en un éxito. La plantilla que se confecciona para documentar los criterios de éxito del proyecto se exhibe en la Tabla 3.
- Expectativas del proyecto: Se identifica lo que se espera conseguir por medio de la ejecución del proyecto de Explotación de Información para poder comprobar el cumplimiento de las expectativas de los clientes. Estas expectativas se deben relacionar con los objetivos y los criterios de éxito del proyecto. La plantilla que se confecciona para documentar las exceptivas del proyecto se exhibe en la Tabla 4.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 3: Plantilla de Criterios de Éxito del Proyecto.

CRITERIOS DE ÉXITO DEL PROYECTO			
Responsable	<nombre del analista>	Fecha	<fecha de confección>
ID	<identificador del proyecto>		
ID Criterio	Descripción	Objetivo del Proyecto	Referencia
<nombre del término>	<descripción breve del criterio de éxito>	<objetivo con el cual se relaciona>	<documentos y sesiones de adquisición de conocimientos que proporcionen la información acerca del criterio de éxito>

Tabla 4: Plantilla de Expectativas del Proyecto.

EXPECTATIVAS DEL PROYECTO				
Responsable	<nombre del analista>	Fecha	<fecha de confección>	
ID	<identificador del proyecto>			
ID Expectativa	Descripción	Objetivo del Proyecto	Criterio de Éxito del Proyecto	Referencia
<identificador secuencial para cada expectativa >	< descripción breve de la expectativa >	< objetivo con el cual se relaciona >	< criterio con el cual se relaciona >	<documentos y sesiones de adquisición de conocimientos que proporcionen la información acerca de la expectativa>

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

- **Supuestos del Proyecto:** Se determinan los supuestos que deben ser tenidos en cuenta respecto de los objetivos del proyecto de Explotación de Información. La plantilla que se confecciona para documentar los supuestos del proyecto se exhibe en la Tabla 5.

Tabla 5: Plantilla de Suposiciones del Proyecto.

SUPOSICIONES DEL PROYECTO			
Responsable	<nombre del analista>	Fecha	<fecha de confección>
ID	<identificador del proyecto>		
ID Suposición	Descripción	Objetivo del Proyecto	Referencia
<identificador secuencial para cada suposición>	<descripción breve de la suposición>	<objetivo con el cual se relaciona>	<documentos y sesiones de adquisición de conocimientos que proporcionen la información acerca de la suposición>

- **Restricciones del Proyecto:** Se identifican los límites del proyecto de Explotación de Información con el objetivo de identificar su contexto. Estos límites se relacionan con la organización (tal como políticas o legislaciones establecidas por el negocio); con los datos

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

(tal como fuentes de acceso a la información y calidad de los datos); con los recursos humanos y recursos técnicos (tal como limitaciones de los recursos humanos o del hardware y software a utilizar); con el proyecto y las actividades que afectan al proyecto. La plantilla que se confecciona para documentar las restricciones del proyecto se exhibe en la Tabla 6.

Tabla 6: Plantilla de Restricciones del Proyecto.

RESTRICCIONES DEL PROYECTO				
Responsable	<nombre del analista>	Fecha	<fecha de confección>	
ID	<identificador del proyecto>			
ID Restricciones	Tipo	Descripción	Objetivo del Proyecto	Referencia
< identificador secuencial para cada restricción >	<tipo de restricción>	<descripción breve de la restricción >	<objetivo con el cual se relaciona>	< documentos y sesiones de adquisición de conocimientos que proporcionen la información acerca de la restricción >

3.3. Viabilidad y Estimación de Proyectos de Explotación de Información

En [37], Pytel propone un primer modelo que realiza la evaluación de viabilidad del proyecto de Explotación de la Información. Esto se realiza a fin de determinar de manera temprana los puntos débiles asociados a las características del proyecto que puedan atraer dificultades. Un segundo modelo se utiliza para la Estimación del esfuerzo (tiempo/hombre) requerido para llevar a cabo el proyecto satisfactoriamente y poder así planificar las actividades que deben desarrollarse.

3.3.1. Propuesta del Modelo para la Evaluación de la Viabilidad

Se presenta una propuesta de Modelado para la evaluación de Viabilidad, la cual describe los cinco pasos que se deben realizar para evaluar la viabilidad de un proyecto.

A continuación, se describen brevemente cada uno de los pasos:

Paso 1: Determinar el valor correspondiente para cada una de las características solicitadas por el modelo. Los valores

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

lingüísticos permitidos para las respuestas son ‘nada’, ‘poco’, ‘regular’, ‘mucho’ y ‘todo’. En la Tabla 7 se describen las características a ser evaluadas.

Tabla 7: Características a ser evaluadas por el método propuesto [38].

Categoría	ID	Pregunta asociada a la Característica	Peso	Umbral
Datos	P1	¿En qué medida los repositorios disponibles poseen datos actuales?	8	<i>poco</i>
	P2	¿Qué tan representativos son los datos de los repositorios disponibles para resolver el problema de negocio?	9	<i>poco</i>
	A1	¿En qué medida los repositorios se encuentran disponibles en formato digital?	4	<i>poco</i>
	A2	¿Qué cantidad de atributos y registros tienen los datos disponibles?	7	<i>poco</i>
	A3	¿Cuánta confianza se posee en la credibilidad de los datos disponibles?	8	<i>poco</i>
	E1	¿Cuánto facilita la tecnología de los repositorios disponibles las tareas de manipulación de los datos?	6	<i>nada</i>
Problema de Negocio	P3	¿Cuánto se entiende del problema de negocio?	7	<i>poco</i>
	A4	¿En qué medida el problema de negocio no puede ser resuelto aplicando técnicas estadísticas tradicionales?	10	<i>poco</i>
	A5	¿Qué tan estable es el problema de negocio durante el desarrollo del proyecto?	9	<i>poco</i>
Tipo de Proyecto	E2	¿Cuánto apoyan los interesados (stakeholders) al proyecto?	8	<i>nada</i>
	E3	¿En qué medida la planificación del proyecto permite considerar la realización de buenas prácticas ingenieriles con el tiempo adecuado?	7	<i>nada</i>
Equipo de Trabajo	P4	¿Qué nivel de conocimientos posee el equipo de trabajo sobre explotación de información?	6	<i>poco</i>
	E4	¿Qué nivel de experiencia posee el equipo de trabajo en proyectos similares?	6	<i>nada</i>

Paso 2: Convertir los valores en intervalos difusos. Una vez que se definen los valores lingüísticos para cada característica de la tabla 1, posteriormente se deben traducir en un intervalo difuso expresados por cuatro valores numéricos (entre cero y diez) que representan los

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

puntos de ruptura (o puntos angulares) de su función de pertenencia correspondiente. En la Fig. 4 se ilustran los intervalos, junto con la representación gráfica de la función de pertenencia.

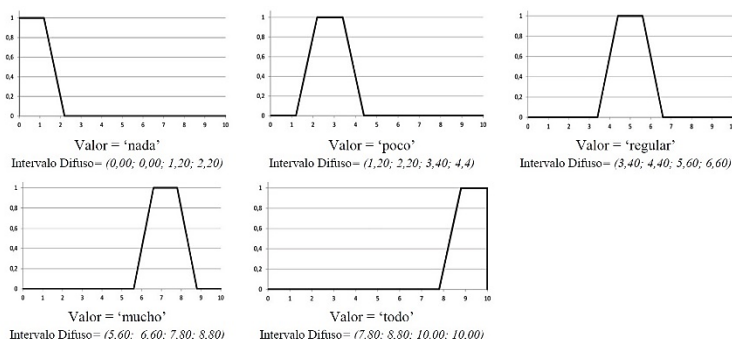


Fig. 4: Representación de la Función de Pertenencia y asignación de Intervalo Difuso para los Valores Lingüísticos. [39].

Paso 3: Calcular la valoración de cada dimensión. Los intervalos difusos convertidos en el paso 2 son ponderados considerando su peso correspondiente el cual se definió en la Tabla 7: Características a ser evaluadas por el método propuesto [38]. Para calcular el intervalo que representa la valoración de cada dimensión se utilizan la fórmula #1 y la fórmula #2 de la Tabla 8: Fórmulas utilizadas por el modelo propuesto [39].

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Paso 4: Calcular la valoración global de la viabilidad del proyecto. Los valores numéricos calculados en el paso anterior son combinados a través de una media aritmética ponderada por medio de la fórmula #3 de la Tabla 8: Fórmulas utilizadas por el modelo propuesto [39]. Con ellos se logra el valor de la viabilidad global del proyecto (EV).

Tabla 8: Fórmulas utilizadas por el modelo propuesto [39].

#	Fórmula	
1	$I_d = \left(\frac{\sum_{i=1}^{n_d} P_{d_i}}{\sum_{i=1}^{n_d} \left(\frac{P_{d_i}}{C_{d_i}} \right)} \right) + \left(\frac{\sum_{i=1}^{n_d} (P_{d_i} \cdot C_{d_i})}{\sum_{i=1}^{n_d} P_{d_i}} \right)$	Donde: I_d : representa el intervalo difuso calculado para la dimensión d (usando como nomenclatura 'P' para plausibilidad, 'A' para adecuación y 'E' para criterio de éxito). P_{d_i} : representa el peso de la característica i perteneciente a la dimensión d . C_{d_i} : representa el intervalo difuso asignado a la característica i perteneciente a la dimensión d . n_d : representa la cantidad de características asociada a la dimensión d .
2	$V_d = \frac{\sum_{i=1}^4 I_{d_i}}{4}$	Donde: V_d : representa el valor numérico calculado para la dimensión d . I_{d_i} : representa el valor de la posición i del intervalo difuso calculado para la dimensión d .
3	$EV = \frac{8 \cdot V_P + 8 \cdot V_A + 6 \cdot V_E}{22}$	Donde: EV : representa el valor global de la viabilidad del proyecto. V_P : representa el valor para la dimensión plausibilidad. V_A : representa el valor para la dimensión adecuación. V_E : representa el valor para la dimensión criterio de éxito.

Paso 5: Interpretar los resultados obtenidos. Para poder interpretar los resultados de la viabilidad de cada dimensión, se recomienda graficar la función de pertenencia del intervalo difuso. Se aceptará la viabilidad de la dimensión, si supera al intervalo del valor 'regular'.

De manera semejante esto se puede determinar si el valor numérico de la dimensión: una dimensión será aceptable si el Valor numérico de la Dimensión d para el Modelo de Viabilidad (Vd) es mayor a 5. Luego, para la viabilidad global del proyecto se utiliza el siguiente criterio: si las tres dimensiones son aceptadas y la valoración global de la viabilidad proyecto (EV) es mayor a 5 entonces el proyecto es viable. En caso contrario, el proyecto no es viable.

3.3.2. Modelo de Estimación para Proyectos de Explotación de Información

Una vez que se determina que el proyecto es viable, es necesario realizar las actividades de planificación del proyecto. En [37] se propone un proceso que permite realizar la estimación de esfuerzo, el cual se compone de tres pasos. En primer lugar, establecer el valor correspondiente para cada uno de los factores de costo del proyecto; en segundo lugar, emplear los valores de los factores de costo en el método seleccionado y en tercer lugar calcular el esfuerzo estimado aplicando la fórmula correspondiente.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Los factores de costos propuestos se encuentran agrupados en tres grupos, relacionados al tipo de proyectos, los datos y recursos disponibles. La fórmula lineal propuesta se ilustra en la Fig. 5:

$$\begin{aligned} \text{PEM}_L = & 0,80 \cdot \text{OBTY} + 1,10 \cdot \text{LECO} - 1,20 \cdot \text{AREP} - 0,30 \cdot \text{QTUM} \\ & - 0,70 \cdot \text{QTUA} + 1,80 \cdot \text{KLDS} - 0,90 \cdot \text{KEXT} \\ & + 1,86 \cdot \text{TOOL} - 3,30 \end{aligned}$$

Fig. 5: Cálculo de estimación de esfuerzo. Fuente [37].

Donde:

- **PEM_L** es el esfuerzo en meses/hombre estimado por la fórmula lineal de estimación.
- **OBTY, LECO, AREP, QTUM, QTUA, KLDS, KEXT y TOOL:** son los valores correspondientes de los factores de costo definidos.

3.4. Propuesta de explotación de información

Los procesos de desarrollo de proyectos de explotación de información existentes, se centran en el descubrimiento de patrones de conocimiento en la masa de datos.

En [18] Britos expone una propuesta de caracterización de los procesos de explotación de información relacionados a

los problemas de inteligencia de negocio en donde define cinco procesos que se adaptan a los posibles diferentes tipos de problemas de negocios. La importancia de estos procesos se centra en que cada uno de ellos representa una problemática diferente, por lo cual la propuesta de solución es diferente para cada caso, para cada una de estos existe un conjunto de pasos y herramientas para la ejecución del mismo.

A continuación, se realiza una breve descripción de los procesos: descubrimiento de reglas de comportamiento, descubrimiento de grupos, descubrimiento de atributos significativos, descubrimiento de reglas de pertenencia a grupos y ponderación de reglas de comportamiento o de pertenencia a grupos y se determina las tecnologías de SI que pueden utilizarse para los procesos caracterizados.

3.4.1. Descubrimiento de Reglas de Comportamiento

El proceso de descubrimiento de reglas de comportamiento se desarrolla cuando es necesario identificar cuál es el comportamiento de un determinado resultado, para ello se

propone la aplicación de algoritmos de inducción TDIDT que permitan identificar dicho comportamiento. Primero se debe seleccionar todas las fuentes de información de interés, ellas pueden ser bases de datos, archivos, fichas, etc. Posteriormente se unifican para lograr un conjunto de datos (Datos Integrados). A partir de allí se selecciona un atributo clase (Atributo de interés para el cual se requiere identificar un comportamiento con base en los datos) y por último se aplica el algoritmo de inducción TDIDT en donde el resultado es un conjunto de reglas que pretende modelar el comportamiento de los datos y el posible resultado de su atributo clase. El esquema procedimental del proceso se muestra en Fig. 6.

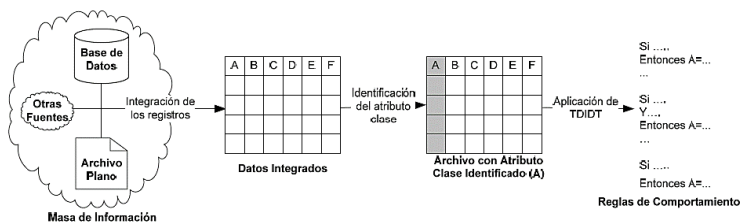


Fig. 6: Esquema y subproductos resultantes de aplicar TDIDT al descubrimiento de reglas de comportamiento: Fuente [18].

3.4.2. Descubrimiento de Grupos

El proceso descubrimiento de grupos se busca identificar dentro de un conjunto de datos y sin empelar ningún criterio previo, subconjuntos del mismo que tengan alguna relación entre sí, para ello se propone la ejecución de Mapas Auto Organizados de Kohonen (SOM). Primero se debe seleccionar todas las fuentes de información de interés, ellas pueden ser bases de datos, archivos, fichas, etc. Posteriormente se aplican mapas auto organizados que generan como resultado diferentes particiones de datos derivadas de los datos integrados. Estas particiones se conocen como grupos identificados. Su esquema procedimental de este proceso se muestra en la Fig. 7.

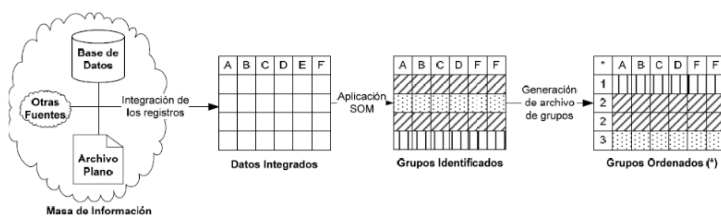


Fig. 7: Esquema de proceso: Descubrimiento de grupos: Fuente [18].

3.4.3. Descubrimiento de Reglas de Pertenencia a Grupos

El proceso Descubrimiento de Reglas de Pertenencia a Grupos se emplea cuando se requiere identificar cuáles son las condiciones de pertenencia a cada una de las clases en una partición desconocida “a priori”, pero presente en la masa de información disponible sobre el dominio de problema. Primero se deben seleccionar todas las fuentes de información de interés, ellas pueden ser bases de datos, archivos, fichas, etc. Posteriormente se aplica SOM que genera como resultado diferentes particiones de datos derivadas de los datos integrados, que se conocen como grupos identificados. Cuando se identifican los grupos, se aplican los algoritmos de inducción (TDIDT) para establecer las reglas de pertenencia a cada uno [36], [40] y [41]. El esquema procedimental de este proceso se muestra en la Fig. 8.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

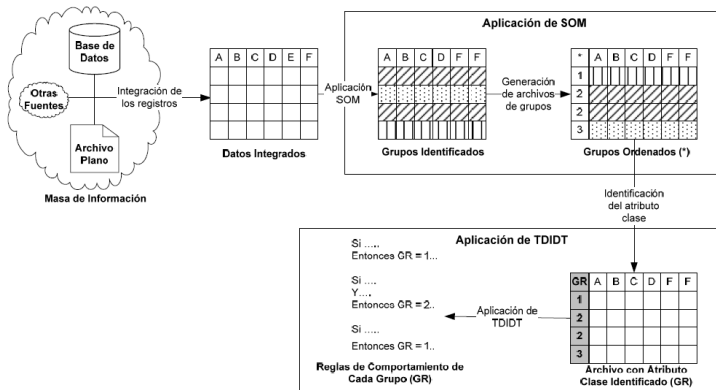


Fig. 8: Esquema y subproductos resultantes de SOM y TDIDT aplicados al descubrimiento de reglas de pertenencia a grupos: Fuente [18].

3.5. Descripción de las herramientas

En el mercado existen entornos que facilitan la ejecución de los algoritmos de minería de datos además de proveer un conjunto más extenso de herramientas que facilitan la gestión de los datos para su posterior uso en diversos algoritmos. Entre las herramientas más conocidas se encuentran Orange, R, Rapidminer, Weka y Tanagra. Para la presente investigación se optó por utilizar las dos últimas.

3.5.1. WEKA

Weka es una plataforma de software destinada a los algoritmos de aprendizaje automático para tareas de minería de datos. Contiene herramientas para la preparación de datos, clasificación, regresión, agrupamiento, minería de reglas de asociación y visualización. Se desarrolló originalmente en el Universidad de Waikato en Nueva Zelanda. WEKA se ha vuelto muy popular entre el sector académico e industrial, investigadores, y también se utiliza ampliamente con fines didácticos [42].

La interfaz de usuario es el explorador, pero la misma funcionalidad puede ser accedida desde la línea de comando o a través del Interfaz de flujo de conocimiento basada en componentes [43]. En la Fig. 9 se ilustra la interfaz de usuario de Weka.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

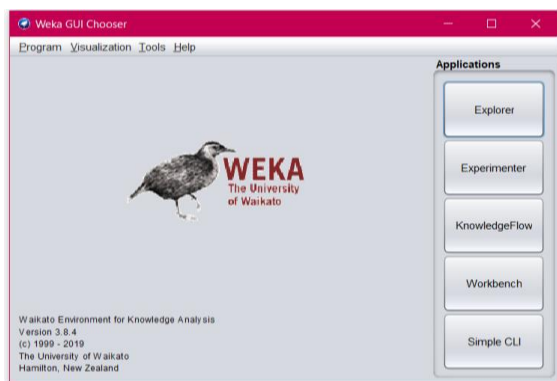


Fig. 9: Interfaz de Usuario. WEKA.

3.5.2. Tanagra

Tanagra es un paquete gratuito de software de aprendizaje automático con fines académicos y de investigación desarrollado por Ricco Rakotomalala en la Universidad Lumière de Lyon 2, Francia. Tanagra admite varias tareas de minería de datos estándar, como: visualización, estadísticas descriptivas, selección de instancias, selección de características, construcción de características, regresión, factor análisis, agrupamiento, clasificación y aprendizaje de reglas de asociación. Tanagra hace un buen compromiso entre los enfoques, los métodos de análisis multivariante y las técnicas de AA [42].

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

El objetivo principal del proyecto Tanagra es brindar a los investigadores y estudiantes un software de minería de datos fácil de usar, conforme a las normas vigentes del software desarrollo en este dominio especialmente en el diseño de su interfaz gráfica de usuario (GUI) y la forma de usarlo, y permite analizar datos reales o sintéticos. El segundo propósito es ofrecer a los investigadores una arquitectura que les permita agregar fácilmente sus propios métodos de minería de datos, para comparar sus actuaciones. Además, actúa más como un experimento plataforma para que puedan ir a lo esencial de su trabajo, dispensándolos para hacer frente a la parte desagradable en la programación de este tipo de herramientas: la gestión de datos [44]. En la Fig. 10 se ilustra la (GUI) de Tanagra.

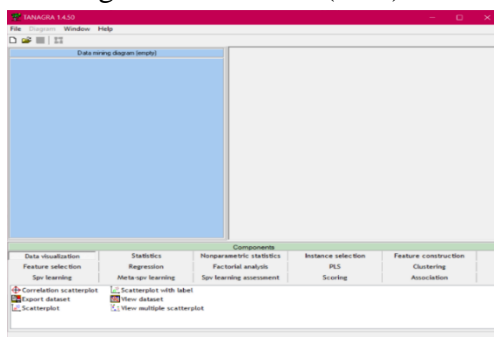


Fig. 10: Tanagra GUI.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Capítulo 4

Modelo de Proceso Propuesto

4. Modelo de Proceso Propuesto

En este capítulo se presenta una propuesta de solución aplicada en forma experimental del MoProPEI a la problemática. El modelo proporciona al equipo de trabajo una guía detallada de los pasos a realizar para el desarrollo de un proyecto de ingeniería de explotación de información en una UCIC, suministrando los elementos de entrada y salida a utilizar para cada una de las actividades requeridas.

4.1. Problemática

La unidad de cuidados intensivos coronarios recibe pacientes desde otros servicios, tales como emergencias, internación, cirugía.

Cuando se solicita a UCIC que reciba un paciente, un médico evalúa la ‘gravedad’ del mismo y determina cómo lo reciben, qué conducta tomar según su estado, por ello los pacientes de mayor vulnerabilidad sanitaria y con riesgo de vida son admitidos en él. Los pacientes de menor gravedad sanitaria son internados en otras dos salas de alta

complejidad: Unidad Intensiva o Terapia Intermedia o son derivados a Emergencia.

También puede suceder que, dada la disponibilidad finita de camas, se deba disponer el traslado a un servicio de menor complejidad, de un internado en UCIC, y permitir la internación de un nuevo ingresante.

Para determinar el estado de un paciente, un médico debe invertir un tiempo importante. Más aún, si no ha tenido contacto con él. Debe analizar la historia clínica electrónica de los que fueron pacientes ya registrados, con toda la información de los eventos sucedidos en la institución. Estos datos involucran los estudios realizados, todas las observaciones en consultas médicas, resultados de laboratorios y de procedimientos previos. Estos datos, consultados para la toma de decisión, no se encuentran ordenados en un único registro o base de datos. Están organizados en interfaces de carga teniendo en cuenta las necesidades de registro según el ámbito en el que la práctica haya sido efectuada.

4.1.1. Pregunta de investigación

Esta tesis pretende dar respuesta a la siguiente pregunta:

- 1) ¿Es posible definir un marco de referencia que permita identificar patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios?

4.1.2. Caso de Estudio

La implementación de las tecnologías de IEI se realizará sobre una base de datos extraída de las HCE de una institución médica especializada en cardiología con el fin de colaborar en la transformación de la información en conocimiento para que contribuya en la toma de decisiones, capacidades operativas e institucionales y la realización de acciones correctivas correspondientes.

4.1.3. Propuesta

En este contexto, se propone la implementación de un Modelo de Procesos para una Ingeniería de Explotación de Información en una unidad de cuidados intensivos coronarios. Como afirma Pollo-Cattaneo en [45], el realizar un planteamiento ingenieril proporciona el

desarrollo de procesos de objetividad, sistematicidad, racionalidad, generalidad y fiabilidad, desarrollando técnicas consistentes para la construcción del conocimiento científico.

Desde una perspectiva general, las investigaciones en Minería de Datos se han centrado en el desarrollo de algoritmos y herramientas. Importa, y por muchas razones, que la atención no esté centrada únicamente en el desarrollo de modelos y algoritmos, sino también en un enfoque metodológico que permita definiciones y especificaciones sistemáticas, organizadas de acuerdo a las necesidades estratégicas para satisfacer los objetivos del proyecto y las restricciones de confiabilidad relativas al resultado [46].

El MoProPEI, como se mencionó en el capítulo 3, se compone de dos subprocesos: Desarrollo, enfocado a las actividades y técnicas, y Gestión, orientado al control y la administración del proyecto. Cada subproceso está compuesto por un conjunto de fases, las cuales agrupan a

las actividades según las finalidades de las mismas, como muestra la Fig. 1.

Con la finalidad de obtener una mayor comprensión de la propuesta y la identificación de la pertenencia de los elementos presentados con respecto a la estructura del proceso, se utiliza una regla nemotécnica, propuesta en [35], donde se indica entre paréntesis a continuación de cada elemento (subproceso, fase, actividad y productos de salida) del proceso.

En las Tabla 9 y Tabla 10 se resume la estructura de MoProPEI a implementar con respecto a los subprocesos, fases, actividades, elementos de entrada y salida.

Tabla 9: G. Estructura del Subproceso Gestión.

FASE	ACTIVIDAD	ENTRADA	SALIDA
Iniciación del Proyecto (G.IP)	Exploración Inicial del Proyecto (G.IP.EIP)	_Fuentes de Información del Cliente. _Objetivo del Proyecto. _Criterios de éxito del proyecto. _Expectativas del proyecto. _Restricciones del proyecto. _Problema del Negocio.	_ Recursos Humanos Involucrados. _Riesgos del Proyecto. _Plan de Contingencias.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

FASE	ACTIVIDAD	ENTRADA	SALIDA
		_Criterios de Éxito del Problema de Negocio.	
	Definición de la Comunicación (G.IP.DCo)	_Recursos Humanos Involucrados.	_Plan de Comunicación.
	Evaluación de la Situación (G.IP.EvS)	_Objetivos del Proyecto _Problema del Negocio _Fuentes de Información del Cliente. _Recursos Humanos Involucrados. _Expectativas del Proyecto. _Suposiciones del Proyecto.	_Reporte de Evaluación de Herramientas. _Reporte de Evaluación de Viabilidad.
	Definición del Ciclo de Vida. (G.IP.DCV)	_Objetivos del Proyecto. _Expectativas del Proyecto. _Problema del Negocio. _Riesgos del Proyecto. _Recursos Humanos Involucrados.	_Modelo de Ciclo de Vida.
Planificación (G.PI)	Planificación de la Mediciones. (G.PI.PIM)	_Objetivos del Proyecto. _Problema del Negocio _Fuentes de Información del Cliente.	_Listado de Métricas. _Estimación del Proyecto.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

FASE	ACTIVIDAD	ENTRADA	SALIDA
		_Recursos Humanos Involucrados _Reporte de Evaluación de Herramientas.	
	Planificación de las Actividades. (G.PI.PIA)	_Modelo de Ciclo de Vida. Estimación del Proyecto. _Objetivos del Proyecto. _Problema del Negocio. _Fuentes de Información del Cliente.	_Mapa de Actividades. _Plan de Acción.
	Planificación de los Recursos. (G.PI.PIR)	_Recursos Humanos Involucrados _Reporte de Evaluación de Herramientas. _Plan de Acción. _Problema del Negocio. _Fuentes de Información del Cliente.	_Plan de Necesidad de Recursos.
	Planificación de las Responsabilidades. (G.PI.PRe)	_Recursos Humanos Involucrados. _Plan de Comunicación. _Plan de Acción. _Plan de Necesidad de Recursos. _Objetivos del Proyecto.	_Matriz de Responsabilidades. _Propuesta del Proyecto.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

FASE	ACTIVIDAD	ENTRADA	SALIDA
		_Criterios de Éxito del Proyecto. Expectativas del Proyecto. _Restricciones del Proyecto. _Problema del Negocio. _Criterios de Éxito del Problema de Negocio. _Riesgos del Proyecto. _Plan de Contingencias.	
Soporte (G.So)	Mediciones del Proyecto (G.So.MeP)	_Listado de Métricas.	_Registro de Mediciones.
	Gestión de la Configuración (G.So.GeC)	_Reglas de Versionado (Externo) _Reporte de Evaluación del Cambio. _Modelo de Ciclo de Vida. _Plan de Acción.	_Reporte de Versionado. _Informe del Estado de la Configuración.
Control (G.Co)	Gestión del Desarrollo (G.Co.GeD)	_Plan de Acción. _Registro de Mediciones.	_Reporte de Estado.
	Control de Actividades (G.Co.CoA)	_Riesgos del Proyecto. _Plan de Contingencias. _Plan de Acción. _Registro de Mediciones. _Reporte de Estado.	_Registro de Riesgos Acontecidos

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

FASE	ACTIVIDAD	ENTRADA	SALIDA
	Gestión del Cambio (G.Co.GCa)	_Solicitud de cambio (Externo). _Propuesta del Proyecto.	Reporte de Evaluación del Cambio
Cierre (G.Ci)	Formalización Externa del Cierre del Proyecto. (G.Ci.FEC)	_Reporte de Evaluación de los Resultados Registro de Riesgos Acontecidos. _Plan de Acción Propuesta del Proyecto.	Documento de Aceptación.
	Formalización Interna del Cierre del Proyecto. (G.Co.FIC)	_Plan de Acción. _Matriz de Responsabilidades. _Registro de Mediciones. _Registro de Riesgos Acontecidos. _Reporte de Evaluación del Cambio. _Reporte de Evaluación de los Resultados. _Documento de Aceptación.	_Reporte de Cierre.

Tabla 10: D. Estructura del Subproceso Desarrollo.

FASE	ACTIVIDAD	ENTRADA	SALIDA
Entendimiento del Negocio (D.EN)	Análisis del Negocio (D.EN.AnN)	_Discursos de los interesados. _Información de la Organización.	_Fuentes de Información del Cliente _Definiciones, Acrónimos y Abreviaturas

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

FASE	ACTIVIDAD	ENTRADA	SALIDA
			<ul style="list-style-type: none"> _Objetivo del Proyecto _Criterios de éxito del proyecto _Expectativas del proyecto _Suposiciones del proyecto _Restricciones del proyecto
	Comprensión del Problema de Negocio (D.EN.CPN)	<ul style="list-style-type: none"> _Discursos de los interesados. Información de la Organización _Definiciones, Acrónimos y Abreviaturas _Objetivo del Proyecto _Criterios de éxito del proyecto _Expectativas del proyecto _Suposiciones del proyecto _Restricciones del proyecto 	<ul style="list-style-type: none"> _Problema del Negocio _Criterios de Éxito del Problema de Negocio
Entendimiento de los Datos (D.ED)	Análisis de los Datos (D.ED.AnD)	<ul style="list-style-type: none"> _Discursos de los interesados (externo) _Fuentes de Información del Cliente _Restricciones del Proyecto 	<ul style="list-style-type: none"> _Diccionario de Fuente de Datos _Campos Relacionados con el Problema de Negocio

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

FASE	ACTIVIDAD	ENTRADA	SALIDA
		Problema del Negocio	
	Exploración de los Datos (D.ED.ExD)	_Suposiciones del Proyecto _Restricciones del Proyecto _Problema del Negocio _Diccionario de Fuente de Datos _Campos Relacionados con el Problema de Negocio _Reporte de Evaluación de Herramientas	_Reporte de Datos Explorados _Fuente Integrada de datos
	Evaluación de los Datos (D.ED.EvD)	_Diccionario de Fuente de Datos _Campos Relacionados con el Problema de Negocio _Reporte de Datos Explorados _Fuente Integrada de datos _Reporte de Evaluación de Herramientas	_Reporte de calidad de los datos
Modelado (D.Mo)	Modelado del problema (D.Mo.MoP)	_Problema del Negocio _Diccionario de Fuente de Datos _Campos Relacionados con	_Diseño del Problema de Explotación

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

FASE	ACTIVIDAD	ENTRADA	SALIDA
		el Problema de Negocio	
	Configuración del Modelo (D.Mo.CoM)	_Criterios de Éxito del Problema de Negocio. _Diccionario de Fuente de Datos. _Campos Relacionados con el Problema de Negocio. _Reporte de Datos Explorados. _Reporte de la Calidad de los Datos. _Diseño del Proceso de Explotación de Información. _Reporte de Evaluación de Herramientas.	_Selección de Algoritmos de Explotación de Información. _Selección de Variables del Modelo. _Estrategias de Evaluación de Modelos.
Preparación de los Datos (D.PD)	Construcción de la Fuente Temporal de Datos (D.PD.CFT)	_Reporte de Datos Explorados. _Estrategias de Evaluación de Modelos. _Fuente Integrada de datos. _Reporte de Evaluación de Herramientas. _Selección de variables del Modelo.	_Fuente Temporal de Datos. _Reporte de Generación de la Fuente Temporal de datos.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

FASE	ACTIVIDAD	ENTRADA	SALIDA
	Adecuación de la Fuente Temporal de Datos (D.PD.AFT)	_Reporte de la Calidad de los Datos. _Selección de variables del Modelo. _Fuente Temporal de Datos. _Reporte de Generación de la Fuente Temporal de datos.	_Reporte de Adecuación de la Fuente Temporal de Datos.
Implementación (D.Im)	Selección del Modelo (D.Im.SeM)	_Selección de Algoritmos de explotación de información. _Reporte de Generación de la Fuente Temporal de datos. _Criterios de Éxito del Problema de Negocio. _Reporte de Evaluación de Herramientas.	_Reporte de Estrategia de Parametrización del Modelo.
	Explotación de la Información. (D.Im.ExI)	_Criterios de Éxito del Problema de Negocio. _Selección de Algoritmos de	_Reporte de Implementación del Modelo. _Patrones de Conocimiento.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

FASE	ACTIVIDAD	ENTRADA	SALIDA
		explotación de información. _Selección de variables del Modelo. _Estrategias de evaluación de modelos. _Reporte de Estrategia de Parametrización del Modelo. _Fuente Temporal de Datos. _Reporte de Generación de la Fuente Temporal de datos.	
Evaluación y Preparación. (D.EP)	Evaluación de los Resultados. (D.EP.EvR)	_Objetivo del Proyecto. _Criterios de éxito del proyecto. _Problema del Negocio. _Criterios de Éxito del Problema de Negocio. _Reporte de Implementación del Modelo. _Patrones de Conocimiento.	_Reporte de Evaluación de los Resultados.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

FASE	ACTIVIDAD	ENTRADA	SALIDA
	Presentación de los Resultados (D.EP.PrR)	_Fuentes de Información del Cliente. _Objetivos del Proyecto. _Criterios de Éxito del Proyecto. _Suposiciones del Proyecto. _Restricciones del Proyecto. _Problema del Negocio. _Patrones de Conocimiento. _Reporte de Generación de la Fuente Temporal de datos. _Reporte de la Calidad de los Datos. _Reporte de Evaluación de los Resultados.	_Reporte del Proyecto.

A continuación, en la sección 4.2: Subproceso de Gestión (G) y la sección 4.3: Subproceso Desarrollo (D), se desarrolla la aplicación de la estructura del MoProPEI y el registro de los formalismos correspondientes al caso de estudio planteado.

4.2. Subproceso de Gestión (G)

El subproceso de gestión se desarrolla de forma transversal a las actividades de desarrollo, su ejecución no es lineal, las tareas se realizan durante todo el progreso del proyecto. Está conformado por cinco fases: Iniciación del Proyecto, Planificación, Soporte, Control y Cierre, como se ilustra a la derecha de la Fig. 1 y la Fig. 2 explican las relaciones entre las fases y las distintas actividades que lo componen.

4.2.1. Iniciación del Proyecto (G.IP)

En esta Fase se realizaron las primeras entrevistas con el médico y los expertos del dominio, las evaluaciones correspondientes a las características del proyecto a desarrollar, se determinaron los recursos humanos involucrados y la estimación de la posibilidad de éxito del proyecto. A continuación, se describen las cuatro actividades de subproceso utilizando el formalismo propuesto, adaptado al caso de estudio.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.2.1.1. Exploración Inicial del Proyecto (G.IP.EIP)

En esta actividad se realizan las tareas necesarias para definir el proyecto, estableciendo las personas necesarias para su desarrollo, analizando los aspectos.

Relacionados con los posibles acontecimientos durante el progreso del proyecto que perjudiquen al desarrollo del mismo y las acciones a realizar en caso de contingencia.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (G.IP.EIP).

Información de Entrada

_Fuentes de Información del Cliente (D.EN.AnN.FuIC). Sección 4.3.1.1.1.

_Objetivo del Proyecto (D.EN.AnN.ObPr). Sección 4.3.1.1.3.

_Criterios de éxito del proyecto (D.EN.AnN.CrEP). Sección 4.3.1.1.4.

Expectativas del proyecto (D.EN.AnN.ExPr). Sección 4.3.1.1.5.

Información de Salida

_Recursos Humanos Involucrados (G.IP.EIP.ReHI). Sección 4.2.1.1.1

_Riesgos del Proyecto (G.IP.EIP.RiPr). Sección 4.2.1.1.2.

_Plan de Contingencias (G.IP.EIP.PICo). Sección 4.2.1.1.3.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

_Restricciones del proyecto
(D.EN.AnN.RePr). Sección
4.3.1.1.7 .

Problema del Negocio
(D.EN.CPN.PrNe). Sección
4.3.1.2.1.

Criterios de Éxito del Problema de
Negocio (D.EN.CPN.CEPN).
Sección 4.3.1.2.2.

4.2.1.1.1. Recursos Humanos Involucrados (G.IP.EIP.ReHI)

En esta etapa se definen los recursos humanos, compuesto por los expertos de dominio, ellos son: el médico de la UCIC con una antigüedad de 25 años trabajando en el Instituto de Cardiología y el representante de la Empresa Aliare S.R.L., (Jefe de Análisis y documentación de Sistema del Instituto de Cardiología) y el Equipo de investigación de FaCENA-UNNE.

En la Tabla 11 se ilustra el formalismo del caso de estudio correspondiente a los recursos humanos involucrados en el proyecto.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 11: G. Recursos Humanos Involucrados.

RECURSOS HUMANOS INVOLUCRADOS					
Responsable		Sofia Vallejos		Fecha	01/06/21
ID		G.IP.EIP.ReHI		Versión	1.0
ID Recurso	Nombre	Rol / posición	Pertenece a	Descripción	Información de contacto
ReHI-01	Fabián Bobadilla	Experto	Empresa Aliare S. R. L.	Experto en el domino de las HCE	fabian@aliare.com.ar
ReHI-02	Jorge Parras	Experto	Instituto de Cardiología	Médico, experto en el domino de las HCE	jorgeparras@gmail.com
ReHI-03	María Inés Pizarrello	Datami-ner del Proyecto	UNNE	Directora de tesis	mainespisar ello@exa.unne.edu.ar
ReHI-04	Sofía Vallejos	Datami-ner	FaCENA UNNE	Tesista	sofiavallejos@gmail.com
ReHI-05	Darío Báez	Datami-ner	FaCENA UNNE	Colaborador	Edbaez1968@gmail.com

4.2.1.1.2. Riesgos del Proyecto (G.IP.EIP.RiPr)

El riesgo del proyecto identificado a partir de la entrevista con los expertos, es un problema frecuente que se presenta en las HCE, es la ausencia de datos del paciente, lo que dificulta conocer la situación de salud del mismo y la disponibilidad del médico.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

En la Tabla 12 se ilustra el formalismo del caso de estudio correspondiente al riesgo del proyecto.

Tabla 12: G. Riesgo del Proyecto.

RIESGOS DEL PROYECTO			
Responsable	Sofía Vallejos	Fecha	02/06/21
ID	G.IP.EIP.RiPr	Versión	1.0
ID Fuente	Descripción	Alcance	Referencia
RiPr-01	Contar con gran cantidad de datos faltantes o incompletos.	Proyecto	Información extraída de la entrevista con los expertos.
RiPr-02	No contar con la presencia del experto (ReHI-02) en etapas críticas de necesidad de interacción con el mismo.	Proyecto	Información extraída de la entrevista con los expertos.

4.2.1.1.3. Plan de Contingencias (G.IP.EIP.PICo)

Luego de plantear el Riesgo del Proyecto es necesario desarrollar un plan de contingencias para mitigar el impacto del mismo. En la Tabla 13 se describen el formalismos del caso de estudio correspondiente al plan de Contingencias.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 13: G. Plan de Contingencias.

PLAN DE CONTINGENCIAS			
Responsable	Sofia Vallejos	Fecha	02/06/21
ID	G.IP.EIP.PICo	Versión	1.0
ID Fuente	Acción	Riesgo asociado	Referencia
PICo-01	Ajustes en los plazos del proyecto.	Contar con datos faltantes o incompletos (RiPr-01).	Información extraída de la entrevista 1 con los expertos.
PICo-02	Ajustes en los plazos del proyecto.	No contar con la presencia del experto (ReHI-02) en etapas críticas de necesidad de interacción con el mismo.	Información extraída 1 de la entrevista con los expertos.

4.2.1.2. Definición de la Comunicación (G.IP.DCo)

En esta actividad se definen las estrategias formales de comunicación de acuerdo a las necesidades e intereses de los miembros.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (G.IP.DCo).

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Información de Entrada

_Recursos Humanos Involucrados
(G.IP.EIP.ReHI). Sección 4.2.1.1.1.

Información de Salida

_Plan de Comunicación
(G.IP.DCo.PCom). Sección
4.2.1.2.1.

**4.2.1.2.1. Plan de Comunicación
(G.IP.DCo.PCom)**

En esta tarea se establecen las comunicaciones a realizar en el proyecto, indicando el responsable y las personas que formarán parte de la misma, la frecuencia con la cual será realizada y el medio de comunicación a utilizar.

Tabla 14: G. Plan de Comunicación.

PLAN DE COMUNICACIÓN				
Responsable	Sofia Vallejos		Fecha	02/06/21
ID	G.IP.DCo.PCom		Versión	1.0
Interesados	Información	Frecuencia	Medio	Responsable
ReHI-01 ReHI-02 ReHI-03 ReHI-04 ReHI-05	Comprensión del Proyecto	Quincenal durante el período de entendimiento del problema.	Correo electrónico / Video Conferencia	Sofia Vallejos (ReHI-04)
ReHI-01 ReHI-03 ReHI-04	Avances del proyecto	Quincenal.	Correo electrónico / Video Conferencia	Sofia Vallejos (ReHI-04)

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

PLAN DE COMUNICACIÓN				
ReHI-01 ReHI-02 ReHI-03 ReHI-04 ReHI-05	Avance de los resultados	Cada vez que se obtienen resultados.	Presencial / Video Conferencia	Sofia Vallejos (ReHI-04)
ReHI-01 ReHI-02 ReHI-03 ReHI-04 ReHI-05	Resultado	Al finalizar el análisis.	Presencial	Sofia Vallejos (ReHI-04)

En la Tabla 14 se ilustra el formalismo del caso de estudio correspondiente al plan de comunicación.

4.2.1.3. Evaluación de la Situación (G.IP.EvS)

En esta actividad se seleccionan las herramientas a utilizar y determina la viabilidad del proyecto.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (G.IP.EvS).

Información de Entrada

_Fuentes de Información del Cliente (D.EN.AnN.FuIC). Sección 4.3.1.1.1.

Información de Salida

_Reporte de Evaluación de Herramientas (G.IP.EvS.EvHe). Sección 4.2.1.3.1

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

_Objetivos del Proyecto
(D.EN.AnN.ObPr). Sección
4.3.1.1.3.

_Reporte de Evaluación de
Viabilidad (G.IP.EvS.REVi).
Sección 4.2.1.3.2

_Problema del Negocio
(D.EN.CPN.PrNe). Sección 4.3.1.2.

_Recursos Humanos Involucrados
(G.IP.EIP.ReHI). Sección 4.2.1.1.1.

_Expectativas del Proyecto
(D.EN.AnN.ExPr). Sección
4.3.1.1.5.

_Suposiciones del Proyecto
(D.EN.AnN.SuPr). Sección
4.3.1.1.6

4.2.1.3.1. Reporte de Evaluación de Herramientas (G.IP.EvS.EvHe)

En esta tarea se identifican como posibles herramientas para utilizar, de acuerdo a las necesidades del cliente y la experiencia de los miembros del proyecto.

Las herramientas a evaluar son Tanagra Versión 1.4.50 y Weka Versión 3.8.4.

Se utiliza el formalismo propuesto en [47] para analizar y seleccionar herramientas para proyectos de explotación de información. Se clasifica cada una de las características de acuerdo a la escala de valores entre 1 y 4 de acuerdo al tipo

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

de pregunta, señalando con “--” aquellos aspectos que no han sido evaluados.

En la Tabla 15 se ilustra el formalismo del caso de estudio correspondiente al reporte de evaluación de herramientas.

Tabla 15: G.Reporte de Evaluación de herramientas.

REPORTE DE EVALUACIÓN DE HERRAMIENTAS				
Responsable	Sofia Vallejos	Fecha	04/06/21	
ID	G.IP.EvS.ReEH	Versión	1.0	
Criterios:				
Evaluación:	1 = Malo, 2 = Débil, 3 = Bueno, 4 = Excelente		1 = No, 4 = SI	
Herramientas		Peso	Tanagra <i>1.450</i>	Weka <i>3.8.4</i>
1. Funcional - Características Técnicas				
Soporte de Metodología / Ciclo de vida	Soporte del proceso	3	2	2
Compatibilidad con fuentes de datos	Base de datos	8	--	--
	Otras fuentes (word, excel, etc.)	8	3	2
Integración	Soporte de distintas técnicas asociadas al proceso de explotación de Información	5	4	4
Multilinguaje	Soporta distintas idiomas	2	1	1
Técnicas	Variedad de técnicas que provee	18	4	4

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

REPORTE DE EVALUACIÓN DE HERRAMIENTAS				
Reporte y visualización	Permite generar reportes y visualizaciones	12	2	2
Multiplataforma	Soporta múltiples plataformas	5	1	4
Instalación remota	La administración y mantenimiento son remotos	5	--	--
Usuarios Múltiples	Posee perfiles de usuarios	2	1	1
Seguridad	Provee seguridad de la información configurada por perfiles	2	1	1
Backup	Metodología de backup	2	1	1
Amigable	Interfaz de usuario	10	4	2
Configuraciones	Permite la configuración del perfil	8	--	--
Documentación	Servicio de soporte y ayuda	5	4	3
Conexión	Soporta conexión por: Inter-net, FTP, ERPs.	2	1	1
Soporte de sistemas de mensaje	Soporta compartir información (por mail u otro medio)	3	1	1
Total			224	206
	Peso del Grupo	40%	89,6	82,4
2. Características del Proveedor				
Características del proveedor	Historia	30	3	3
Crecimiento	Perspectiva a futuro	10	2	3

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

REPORTE DE EVALUACIÓN DE HERRAMIENTAS				
Ubicación Geográfica	Oficinas	30	--	--
Implementación	Otras implementaciones de la misma herramienta	5	--	--
	Contacto con otros clientes	5	--	--
Confidencialidad	Confidencialidad de la información	20	--	--
Total			110	120
Peso del Grupo		25%	27,5	30
3. Características del Servicio				
Garantía del producto	Duración y Alcance	30	--	--
Mejora	Brinda soporte a versiones previas	20	1	1
Licencia	Costo, alcances y soporte postventa	30	--	--
Soporte	Tiempo de respuesta y disponibilidad	20	--	--
Total			20	20
Peso del Grupo		20%	4	4
4. Características Económicas				
Costo del software	Costo de la herramienta	30	--	--
Costo del Hardware	Necesidad de mejorar o comprar nuevo hardware compatible con la herramienta	20	--	--

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

REPORTE DE EVALUACIÓN DE HERRAMIENTAS				
Otros costos software	Costos adicionales al producto (backup, web servers, bases de datos, etc.)	20	--	--
Licencias	Política de licencia	10	--	--
Financiamiento	Existencia	10	--	--
Mejoras	Costo promedio de la mejora del producto	10	--	--
Total				
	Peso del Grupo	-15%	0	0
Final				
1. Funcional - Características Técnicas		40%	89,6	82,4
2. Características del Proveedor		25%	27,5	30
3. Características del Servicio		20%	4	4
4. Características Económicas		-15%	0	0
Total			121,1	116,4

De acuerdo a los resultados obtenidos, se identifica a la herramienta Tanagra como la más adecuada para el proyecto.

4.2.1.3.2. Reporte de Evaluación de Viabilidad (G.IP.EvS.REVi)

Para obtener el Reporte de Evaluación de la Viabilidad se utiliza el formalismo, basado en la técnica de viabilidad propuesta en [37], que se encuentra en el Capítulo 3, sección 3.3.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 16: Preguntas asociadas a la caracterización.

Categoría	ID	Pregunta asociada a la Característica	Valor
Datos	P1	¿En qué medida los repositorios disponibles poseen datos actuales?	Mucho
	P2	¿Qué tan representativos son los datos de los repositorios disponibles para resolver el problema de negocio?	Mucho
	A1	¿En qué medida los repositorios se encuentran disponibles en formato digital?	Mucho
	A2	¿Qué cantidad de atributos y registros tienen los datos disponibles?	Mucho
	A3	¿Cuánta confianza se posee en la credibilidad de los datos disponibles?	Mucho
	E1	¿Cuánto facilita la tecnología de los repositorios disponibles las tareas de manipulación de los datos?	Mucho
Problema de Negocio	P3	¿Cuánto se entiende del problema de negocio?	Mucho
	A4	¿En qué medida el problema de negocio no puede ser resuelto aplicando técnicas estadísticas tradicionales?	Regular
	A5	¿Qué tan estable es el problema de negocio durante el desarrollo del proyecto?	Mucho
Proyecto	E2	¿Cuánto apoyan los interesados al proyecto?	Todo
	E3	¿En qué medida la planificación del proyecto permite considerar la realización de buenas prácticas ingenieriles con el tiempo adecuado?	Mucho

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Categoría	ID	Pregunta asociada a la Característica	Valor
Equipo de Trabajo	P4	¿Qué nivel de conocimientos posee el equipo de trabajo sobre explotación de información?	Regular
	E4	¿Qué nivel de experiencia posee el equipo de trabajo?	Regular

En la Tabla 16 se indican las preguntas para realizar la evaluación de viabilidad y se muestran los valores lingüísticos correspondiente a las respuestas requeridas de las preguntas. Esos valores son convertidos en intervalos difusos para luego calcular el intervalo de cada dimensión como se muestra en la Tabla 18.

Tabla 17: G. Reporte de Evaluación de Viabilidad.

REPORTE DE EVALUACIÓN DE VIABILIDAD													
Responsable		Sofia Vallejos				Fecha		04/06/21					
ID		G.IP.EvS.REVi				Versión		1.0					
Datos						Problema de Negocio			Proyecto		Equipo de Trabajo		
P1	P2	A1	A2	A3	E1	P3	A4	A5	E2	E3	P4	E4	
Mucho	Mucho	Mucho	Mucho	Mucho	Mucho	Mucho	Regular	Mucho	Todo	Mucho	Regular	Regular	
Peso													
8	9	4	7	8	6	7	10	9	8	7	6	6	
Umbral													
Poco	Poco	Poco	Poco	Poco	Nada	Poco	Poco	Poco	Nada	Nada	Poco	Nada	

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 18: Traducción y cálculo de intervalos por dimensión.

Dimensión	ID	Peso	Intervalo				Intervalo Valor Dimensión
Plausibilidad	P1	8	5,6;	6,6;	7,8;	8,8	5,06; 6,08; 7,3; 8,31
	P2	9	5,6;	6,6;	7,8;	8,8	
	P3	7	5,6;	6,6;	7,8;	8,8	
	P4	6	3,4;	4,4;	5,6;	6,6	
Adecuación	A1	4	5,6;	6,6;	7,8;	8,8	4,9; 5,93; 7,15; 8,16
	A2	7	5,6;	6,6;	7,8;	8,8	
	A3	8	5,6;	6,6;	7,8;	8,8	
	A4	10	3,4;	4,4;	5,6;	6,6	
	A5	9	5,6;	6,6;	7,8;	8,8	
Éxito	E1	6	5,6;	6,6;	7,8;	8,8	5,52; 6,56; 7,8; 8,57
	E2	8	7,8;	8,8;	10;	10	
	E3	7	5,6;	6,6;	7,8;	8,8	
	E4	6	3,4;	4,4;	5,6;	6,6	

Por último, se calcula el valor numérico de cada dimensión y la valoración global de la viabilidad del proyecto, los cuales se interpretan como se indica en la Tabla 19.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 19: Evaluación de la situación.

Dimensiones		Viabilidad global	Resultado
Plausibilidad	6,69	6.75	Viable
Adecuación	6,54		
Éxito	7,11		

4.2.1.4. Definición del Ciclo de Vida (G.IP.DCV)

En esta actividad se define el orden, la transición y las iteraciones en el cual se realizarán las fases en el desarrollo del proyecto. A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (G.IP.DCV).

Información de Entrada

_Objetivo del Proyecto
(D.EN.AnN.ObPr). Sección
4.3.1.1.3.

Expectativas del proyecto
(D.EN.AnN.ExPr). Sección
4.3.1.1.5.

Problema del Negocio
(D.EN.CPN.PrNe). Sección
4.3.1.2.1.

_Recursos Humanos Involucrados
(G.IP.EIP.ReHI). Sección 4.2.1.1.1.

_Riesgos del Proyecto
(G.IP.EIP.RiPr). Sección 4.2.1.1.2.

Información de Salida

_ Modelo de Ciclo de Vida
(G.IP.DCV.MCV). Sección
4.2.1.4.1.

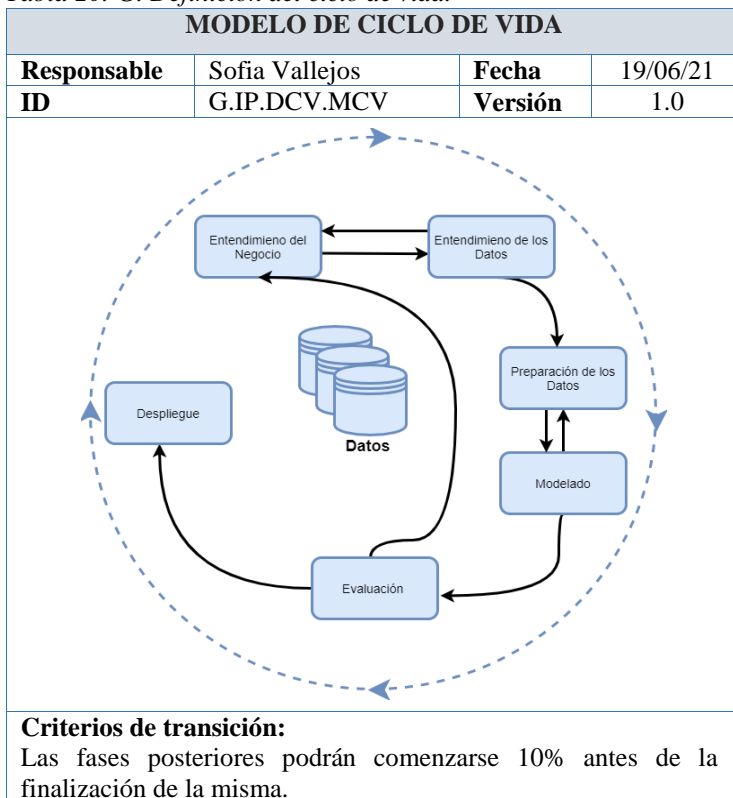
4.2.1.4.1. Modelo de Ciclo de Vida (G.IP.DCV.MCV)

En el Proyecto las necesidades y los alcances son conocidos y se encuentran definidos desde el inicio del proyecto. Es por ello que se determinó utilizar el Ciclo de Vida Genérico para Minería de Datos (Generic Data Mining Life Cycle, DMLC) [48], basado en el definido en CRISP-DM [49]. Para el desarrollo de proyectos de explotación de información este modelo de proceso es muy utilizado. Dicha metodología se desarrolla a partir de un ciclo de vida, compuesto por seis fases: Comprensión del negocio, Comprensión de los datos, Preparación de los datos, Modelado, Evaluación y Despliegue.

En Tabla 20 se ilustra el formalismo desarrollado, presentando la visualización del modelo de ciclo de vida (MCV) y el criterio de transición entre fases.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 20: G. Definición del ciclo de vida.



4.2.2. Planificación (G.PI)

En la fase Planificación se desarrollan actividades relacionadas con la definición del esfuerzo total del proyecto, la definición del sentido de acciones requerido para alcanzar los objetivos y su ajuste (en caso de ser

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

necesario), la planificación de las actividades y documentos involucrados en el proceso.

A continuación, se describen las cuatro actividades de subproceso utilizando el formalismo propuesto, adaptado al caso de estudio.

4.2.2.1. Planificación de la Mediciones (G.PI.PIM)

En esta actividad se determina el tiempo inicial y el conjunto de características que serán medidas y evaluadas en el transcurso del proyecto.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (G.PI.PIM).

Información de Entrada

_ Fuentes de Información del Cliente (D.EN.AnN.FuIC). Sección 4.3.1.1.1.

_ Objetivo del Proyecto (D.EN.AnN.ObPr). Sección 4.3.1.1.3.

Problema del Negocio (D.EN.CPN.PrNe). Sección 4.3.1.2.1.

Información de Salida

_ Listado de Métricas (G.PI.PIM.LiMe). Sección 4.2.2.1.1.

_ Estimación del Proyecto (G.PI.PIM.EsPr). Sección 4.2.2.1.2.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

_Recursos Humanos Involucrados
(G.IP.EIP.ReHI). Sección 4.2.1.1.1.

Reporte de Evaluación de
Herramientas (G.IP.EvS.EvHe).
Sección 4.2.1.3.1.

4.2.2.1.1. Listado de Métricas (G.Pl.PIM.LiMe)

Se propone el uso de dos métricas, una enfocada en el proyecto y la otra en los datos, “Tiempo total requerido para el desarrollo del proyecto” y “Grado de Utilidad de Atributos”.

En la Tabla 21 se ilustra el formalismo del caso de estudio correspondiente al listado de métricas.

Tabla 21: G. Listado de Métricas.

LISTADO DE METRICAS				
Responsable	Sofia Vallejos		Fecha	22/06/21
ID	G.Pl.PIM.LiMe		Versión	1.0
Nombre	Tipo	Fórmula	Comentarios	
Tiempo total requerido para el desarrollo del proyecto	Proyecto	$DRPY = \sum trA$ $trA = \text{tiempo requerido por actividad}$	Sumatoria de los tiempos requeridos para cada actividad del proyecto .	

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Grado de Utilidad de Atributos	Datos	$NA(T) =$ $NASE(T) +$ $NAUD(T) +$ $NANC(T) +$ $NANS(T)$ $NA(T) =$ $NO_$ $UTILES(T) =$ $NANC(T) +$ $NANS(T)$ $NO_$ $UTILES(T) =$	_Nro. de atributos útiles sin errores $[NASE(T)] =$ _Nro. de atributos útiles con defectos $[NAUD(T)] =$ _Nro. de atributos no correctos [$NANC(T)] =$ _Nro. de atributos no significativos $[NANS(T)] =$
--------------------------------	-------	--	---

Los atributos no significativos son los que identifican al paciente, a la historia clínica y a la fecha de ingreso y egreso.

4.2.2.1.2. Estimación del Proyecto (G.PI.PIM.EsPr)

Teniendo en cuenta las características del proyecto y la técnica de estimación visto en el Capítulo 3, sección 3.3.2 “Modelo de Estimación para Proyectos de Explotación de Información” se evaluarán y fijarán los 8 factores de acuerdo a las escalas establecidas en el modelo.

Se realizará una adaptación de las tablas de valoración de costo propuesta por Pytel en [38], en la cual se agregará

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

una tercera columna en donde se indique con una X la valoración que se adapta al caso de estudio.

A continuación, se indican los factores de costo y la valoración correspondiente al caso de estudio:

Tipo de objetivo de explotación de información (OBTY):

Se analiza el objetivo del proyecto a partir del tipo de proceso de EI a ser aplicado, tal como se marca en la Tabla 22.

Tabla 22: Tipo de objetivo de explotación de información.

VALOR	DESCRIPCIÓN	MARCA
1	Se desea conocer los atributos que caracterizan el comportamiento o la descripción de una clase ya conocida.	X
2	Se desea dividir los datos disponibles en grupos sin poseer una clasificación conocida previamente.	
3	Se desea conocer los atributos que caracterizan a grupos sin poseer una clasificación conocida previamente.	
4	Se desea conocer los atributos que poseen mayor frecuencia de incidencia sobre un comportamiento o la identificación de una clase conocida.	
5	Se desea conocer los atributos que poseen mayor frecuencia de incidencia sobre la identificación de una clase desconocida previamente.	

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Grado de apoyo de los miembros de la organización (LECO):

Se debe considerar en qué medida la gerencia media (supervisores y jefes de área) y/o el resto del personal están dispuestos a asistir al equipo de trabajo en el entendimiento del negocio y los datos disponibles. En la Tabla 23 se marca la valoración que se ajusta al caso de estudio.

Tabla 23: Grado de apoyo de los miembros de la organización.

VALOR	DESCRIPCIÓN	MARCA
1	Tanto los directivos como el personal poseen buena disposición para colaborar en el proyecto.	
2	Sólo los directivos poseen buena disposición para colaborar en el proyecto mientras que el personal es indiferente al proyecto.	X
3	Sólo la alta gerencia posee buena disposición para colaborar en el proyecto mientras que la gerencia media y el personal es indiferente.	
4	Sólo la alta gerencia posee buena disposición para colaborar en el proyecto pero la gerencia media no desea colaborar.	

Cantidad y tipo de los repositorios de datos disponibles (AREP):

Se analizan las fuentes de datos disponibles. En la Tabla 24 se marca la valoración que se ajusta al caso de estudio.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 24: Cantidad y tipo de los repositorios de datos disponibles.

VALOR	DESCRIPCIÓN	MARCA
1	Sólo 1 repositorio disponible.	X
2	Entre 2 y 4 repositorios con tecnología compatible para la integración.	
3	Entre 2 y 4 repositorios con tecnología no compatible para la integración.	
4	Más de 5 repositorios con tecnología compatible para la integración.	
5	Más de 5 repositorios con tecnología no compatible para la integración.	

Cantidad de tuplas disponibles en la tabla principal (QTUM):

La cantidad de tuplas correspondiente al caso de estudio es de 690. En la Tabla 25 se marca la valoración que se ajusta al caso de estudio.

Tabla 25: Cantidad de tuplas disponibles en la tabla principal.

VALOR	DESCRIPCIÓN	MARCA
1	Hasta 100 tuplas en la tabla principal.	
2	Entre 101 y 1.000 tuplas en la tabla principal.	X
3	Entre 1.001 y 20.000 tuplas en la tabla principal.	
4	Entre 20.001 y 80.000 tuplas en la tabla principal.	

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Cantidad de tuplas disponibles en las tablas auxiliares (QTUA):

No se utilizan tablas auxiliares. En la Tabla 26 se marca la valoración que se ajusta al caso de estudio.

Tabla 26: Cantidad de tuplas disponibles en las tablas auxiliares.

VALOR	DESCRIPCIÓN	MARCA
1	No se utilizan tablas auxiliares.	X
2	Hasta 1.000 tuplas en las Tablas auxiliares.	
3	Entre 1.001 y 50.000 tuplas en las Tablas auxiliares.	
4	Más de 50.000 tuplas en las Tablas auxiliares.	

Nivel de conocimiento sobre los datos (KLDS):

Se identifica el nivel de documentación existente sobre los repositorios de datos. En la Tabla 27 se marca la valoración que se ajusta al caso de estudio.

Nivel de conocimiento y experiencia del equipo de trabajo (KEXT):

Se analiza la capacidad del equipo de trabajo que se ocupa de llevar adelante el proyecto.

En la Tabla 28 se marca la valoración que se ajusta al caso de estudio.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 27: Nivel de conocimiento sobre los datos.

VALOR	DESCRIPCIÓN	MARCA
1	Todas las tablas y repositorios están correctamente documentados.	
2	Más del 50% de las tablas y repositorios están correctamente documentados y existen expertos en los datos disponibles para explicarlos.	X
3	Menos del 50% de las tablas y repositorios están correctamente documentados, pero existen expertos en los datos disponibles para explicarlos.	
4	Las tablas y repositorios no están documentadas, pero existen expertos en los datos disponibles para explicarlos.	
5	Las tablas y repositorios no están documentados y existen expertos en los datos, pero no están disponibles para explicarlos.	
6	Las tablas y repositorios no están documentados y no existen expertos en los datos para explicarlos.	

Tabla 28: Nivel de conocimiento y experiencia del equipo de trabajo.

VALOR	DESCRIPCIÓN	MARCA
1	El equipo ha trabajado en tipos de organizaciones y con datos similares para obtener los mismos objetivos.	
2	El equipo ha trabajado en tipos de organizaciones similares pero con datos diferentes para obtener los mismos objetivos.	
3	El equipo ha trabajado en otros tipos de organizaciones y con datos similares para obtener los mismos objetivos.	
4	El equipo ha trabajado en otros tipos de organizaciones y con datos diferentes para obtener los mismos objetivos.	
5	El equipo ha trabajado en tipos de organizaciones diferentes, con datos diferentes y otros objetivos.	X

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Funcionalidad de las herramientas disponibles (TOOL):

Se evalúa las características de las herramientas disponibles para ser aplicadas en el proyecto. En la Tabla 29 se marca la valoración que se ajusta al caso de estudio.

Tabla 29: Funcionalidad de las herramientas disponibles .

VALOR	DESCRIPCIÓN	MARCA
1	La herramienta posee funciones tanto para el formateo e integración de los datos (permitiendo importar más de una Tabla de datos) como para aplicar las técnicas de minería de datos.	
2	La herramienta posee funciones tanto para el formateo como para aplicar las técnicas de minería de datos, y permite importar más de una Tabla de datos en forma independiente.	
3	La herramienta posee funciones tanto para el formateo como para aplicar las técnicas de minería de datos, pero sólo permite importar una Tabla de datos.	X
4	La herramienta posee funciones sólo para aplicar las técnicas de minería de datos, y permite importar más de una Tabla de datos.	
5	La herramienta posee funciones sólo para aplicar las técnicas de minería de datos y sólo permite importar una Tabla de datos.	

Una vez definido los valores de los factores, se debe aplicar la formula lineal vista en el la Fig. 5: Cálculo de estimación de esfuerzo. Fuente [37].

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

$$PEM_L = 0,80*OBTY + 1,10*LECO - 1,20*AREP - 0,30*QTUM - 0,70*QTUA + 1,80*KLDS - 0,90*KEXT + 1,86*TOOL - 3,30$$

$$PEM_L = (0,80*1) + (1,10*2) - (1,20*1) - (0,30*2) - (0,70*1) + (1,80*2) - (0,90*5) + (1,86*3 - 3,30)$$

$$PEM_L = 1,88 \text{ meses/hombre}$$

Se obtiene como esfuerzo total del subproceso desarrollo: 1,58 meses/hombre y de acuerdo al porcentaje estimado para tareas de gestión (15%), se determina como esfuerzo para dicha fase: 0,28 meses/hombre, siendo el esfuerzo total del proyecto igual a: 2,16 meses/hombre. En la Tabla 30 se indica la estimación realizada.

Tabla 30: G. Estimación del Proyecto

ESTIMACIÓN DEL PROYECTO										
Responsable		Sofia Vallejos					Fecha		23/06/21	
ID		G.PI.PIM.EsPr					Versión		1.0	
Esfuerzo										
O B T Y	L E C O	A R E P	Q T U M	Q T U A	K L D S	K E X T	T O O L	Total Desa- rrollado	Total Gestión	Total
1	2	1	2	1	2	5	3	1,88	0,28	2,16

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.2.2.2. Planificación de las Actividades (G.PI.PIA)

En esta actividad se describen las acciones a realizar durante el transcurso del proyecto, sus alcances y se define la ejecución de las actividades (tiempo).

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (G.PI.PIA).

Información de Entrada

_Modelo de Ciclo de Vida
(G.IP.DCV.MCV). Sección
4.2.1.4.1.

_Estimación del Proyecto
(G.PI.PIM.EsPr). Sección 4.2.2.1.2.

_Fuentes de Información del
Cliente (D.EN.AnN.FuIC). Sección
4.3.1.1.1.

_Objetivo del Proyecto
(D.EN.AnN.ObPr). Sección
4.3.1.1.3.

_Problema del Negocio
(D.EN.CPN.PrNe). Sección
4.3.1.2.1.

Información de Salida

_Mapa de Actividades
(G.PI.PIA.MaAc). Sección 4.2.2.2.1

_Plan de Acción (G.PI.PIA.PIAc).
Sección 4.2.2.2.2

4.2.2.2.1. Mapa de Actividades (G.PI.PIA.MaAc)

El formalismo consiste en listar las actividades a realizar en filas incluyendo las fases a las que pertenecen a modo

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

de referencia en la estructura del proyecto y en las columnas las etapas del modelo de ciclo de vida (Ver: Sección 4.2.1.4.1: *Modelo de Ciclo de Vida (G.IP.DCV.MCV)*). En los cuadrantes de intersección entre las actividades y las etapas del MCV, se indica con una equis (“X”), si la actividad se debe realizar durante dicha fase. El desarrollo completo de la actividad puede requerir más de una etapa. En la Tabla 31 se ilustra el formalismo del caso de estudio correspondiente al Mapa de Actividades.

Tabla 31: G. Mapa de Actividades.

MAPA DE ACTIVIDADES							
Responsable	Sofia Vallejos	Fecha:	19/06/21				
ID	G.Pl.PIA.MaAc	versión:	1.0				
Esfuerzo							
ID	Fase / Actividad	E. N.	E. D.	P. D.	M.	E v.	D.
G.IP	Iniciación						
G.IP.EIP	Exploración Inicial del Proyecto	x					
G.IP.DCo	Definición de la Comunicación	x					
G.IP.EvS	Evaluación de la Situación	x					
G.IP.DCV	Definición del Ciclo de Vida	x					

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

MAPA DE ACTIVIDADES							
Responsable	Sofia Vallejos	Fecha:		19/06/21			
ID	G.Pl.PIA.MaAc	versión:		1.0			
Esfuerzo							
ID	Fase / Actividad	E. N.	E. D.	P. D.	M.	E v.	D.
G.Pl	Planificación						
G.Pl.PIM	Planificación de la Mediciones	x	x				
G.Pl.PIA	Planificación de las Actividades	x	x				
G.Pl.PIR	Planificación de los Recursos	x	x				
G.Pl.PRe	Planificación de las Responsabilidades	x	x				
G.So	Soporte						
G.So.MeP	Mediciones del Proyecto	x	x	x	x	x	x
G.So.GeC	Gestión de la Configuración	x	x	x	x	x	x
G.Co	Control						
G.Co.GeD	Gestión del Desarrollo	x	x	x	x	x	x
G.Co.CoA	Control de las Actividades	x	x	x	x	x	x
G.Co.GCa	Gestión del Cambio	x	x	x	x	x	x
G.Ci	Cierre						
G.Ci.FEC	Formalización Externa del Cierre del Proyecto						x

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

MAPA DE ACTIVIDADES							
Responsable	Sofia Vallejos	Fecha:	19/06/21				
ID	G.Pl.PIA.MaAc	versión:	1.0				
Esfuerzo							
ID	Fase / Actividad	E. N.	E. D.	P. D.	M.	E v.	D.
G.Ci.FIC	Formalización Interna del Cierre del Proyecto						x
D.EN	Entendimiento del Negocio						
D.EN.AnN	Análisis del Negocio	x					
D.EN.CPN	Comprensión del Problema de Negocio	x					
D.ED	Entendimiento de los Datos						
D.ED.AnD	Análisis de los Datos	x					
D.ED.ExD	Exploración de los Datos	x					
D.ED.EvD	Evaluación de los Datos	x					
D.Mo	Modelado						
D.Mo.MoP	Modelado del problema			x			
D.Mo.CoM	Configuración del Modelo			x			
D.PD	Preparación de los Datos						
D.PD.CFT	Construcción de la Fuente			x			

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

MAPA DE ACTIVIDADES							
Responsable	Sofia Vallejos	Fecha:		19/06/21			
ID	G.PI.PIA.MaAc	versión:		1.0			
Esfuerzo							
ID	Fase / Actividad	E. N.	E. D.	P. D.	M.	E v.	D.
	Temporaria de Datos						
D.PD.AFT	Adecuación de la Fuente Temporaria de Datos			x			
D.Im	Implementación						
D.Im.SeM	Selección del Modelo				x		
D.Im.ExI	Explotación de Información				x		
D.EP	Evaluación y Presentación						
D.EP.EvR	Evaluación de los Resultados					x	
D.EP.PrR	Presentación de los Resultados						x

4.2.2.2.2. Plan de Acción (G.PI.PIA.PIaC)

Una vez definido la estimación de tiempos y la selección de las actividades a realizar en cada etapa del modelo de ciclo de vida (ver Mapa de Actividades (G.PI.PIA.MaAc)), se asigna la duración y rango de fechas de ejecución de cada una de las actividades. Por ello, se utiliza la propuesta

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

de [50] con respecto a la base las mediciones de una estimación empírica del porcentaje del tiempo del proyecto de explotación de información que insume la ejecución de cada una de las actividades.

En la Tabla 32 se ilustra el formalismo del caso de estudio correspondiente al plan de acción, para optimizar los espacios en la tabla se le retiró el año el cual corresponde al 2021.

Tabla 32: G. Plan de Acción.

PLAN DE ACCIÓN								
Responsable		Sofia Vallejos			Fecha:		08/08/21	
ID		G.Pl.PIA.PIAc			versión:		1.1	
ID	Actividad	Inicio Estimado	Inicio Real	Fin Estimado	Fin Real	Esfuerzo Estimado hs.	Esfuerzo Real hs.	Comentario
G.IP	Iniciación	1/06	1/06	19/06	19/06	12	12	
G.IP. EIP	Exploración Inicial del Proyecto	1/06	1/06	15/06	15/06	4	5	
G.IP. DCo	Definición de la Comunicación	1/06	1/06	15/06	15/06	2	2	
G.IP. EvS	Evaluación de la Situación	1/06	1/06	15/06	15/06	4	4	

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

PLAN DE ACCIÓN								
Responsable		Sofia Vallejos			Fecha:		08/08/21	
ID		G.Pl.PIA.PlAc			versión:		1.1	
ID	Actividad	Inicio Estimado	Inicio Real	Fin Estimado	Fin Real	Esfuerzo Estimado hs.	Esfuerzo Real hs.	Comentario
G.IP. DCV	Definición del Ciclo de Vida	16/06	16/06	19/06	19/06	2	2	
G.Pl	Planificación	1/06	1/06	26/06	26/06	10	10	
G.Pl. PIM	Planificación de las Mediciones	1/06	1/06	26/06	26/06	2	2	
G.Pl. PIA	Planificación de las Actividades	1/06	1/06	26/06	26/06	2	2	
G.Pl. PIR	Planificación de los Recursos	1/06	1/06	26/06	26/06	2	2	
G.Pl. PRe	Planificación de las Responsabilidades	1/06	1/06	26/06	26/06	4	4	
G.So	Soporte	18/06	18/06	08/08	08/08	8	14	
G.So. MeP	Mediciones del Proyecto	18/06	18/06	08/08	08/08	4	4	
G.So. GeC	Gestión de la Configuración	18/06	18/06	08/08	08/08	4	10	
G.Co	Control	18/06	18/06	06/08	06/08	12	10	

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

PLAN DE ACCIÓN								
Responsable		Sofia Vallejos			Fecha:		08/08/21	
ID		G.Pl.PIA.PIAc			versión:		1.1	
ID	Actividad	Inicio Estimado	Inicio Real	Fin Estimado	Fin Real	Esfuerzo Estimado hs.	Esfuerzo Real hs.	Comentario
G.Co .GeD	Gestión del Desarrollo	18/06	18/06	06/08	06/08	6	4	
G.Co .CoA	Control de las Actividades	18/06	18/06	06/08	06/08	4	4	
G.Co .Gca	Gestión del Cambio	18/06	18/06	06/08	06/08	2	2	
G.Ci	Cierre	4/08	4/08	08/08	08/08	6	6	
G.Ci. FEC	Formalización Externa del Cierre del Proyecto	4/08	4/08	06/08	06/08	2	2	
G.Ci. FIC	Formalización Interna del Cierre del Proyecto	07/08	07/08	08/08	08/08	4	4	
D.EN	Entendimiento del Negocio	1/06	1/06	19/06	19/06	48	40	
D.EN .AnN	Análisis del Negocio	1/06	1/06	19/06	19/06	30	30	
D.EN .CPN	Comprensión del Problema de Negocio	15/06	15/06	19/06	19/06	18	10	
D.ED	Entendimiento de los Datos	22/06	22/06	03/07	03/07	60	58	

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

PLAN DE ACCIÓN								
Responsable		Sofia Vallejos			Fecha:		08/08/21	
ID		G.Pl.PIA.PlAc			versión:		1.1	
ID	Actividad	Inicio Estimado	Inicio Real	Fin Estimado	Fin Real	Esfuerzo Estimado hs.	Esfuerzo Real hs.	Comentario
D.ED .AnD	Análisis de los Datos	22/06	22/06	26/06	26/06	22	20	
D.ED .ExD	Exploración de los Datos	25/06	25/06	03/07	03/07	20	20	
D.ED .EvD	Evaluación de los Datos	29/06	29/06	03/07	03/07	18	16	
D.Mo .Mo	Modelado	06/07	06/07	10/07	10/07	22	18	
D.Mo .MoP	Modelado del problema	06/07	06/07	07/07	07/07	12	10	
D.Mo .CoM	Configuración del Modelo	08/07	08/07	10/07	10/07	10	8	
D.PD .PD	Preparación de los Datos	13/07	13/07	17/07	17/07	32	26	
D.PD .CFT	Construcción de la Fuente Temporal de Datos	13/07	13/07	15/07	15/07	12	6	
D.PD .AFT	Adecuación de la Fuente Temporal de Datos	14/07	14/07	17/07	17/07	20	20	
D.Im .Im	Implementación	19/07	19/07	23/07	23/07	32	30	

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

PLAN DE ACCIÓN								
Responsable		Sofia Vallejos			Fecha:		08/08/21	
ID		G.PI.PIA.PIAc			versión:		1.1	
ID	Actividad	Inicio Estimado	Inicio Real	Fin Estimado	Fin Real	Esfuerzo Estimado hs.	Esfuerzo Real hs.	Comentario
D.Im .SeM	Selección del Modelo	19/07	19/07	20/07	20/07	12	10	
D.Im .ExI	Explotación de Información	20/07	20/07	23/07	23/07	20	20	
D.EP	Evaluación y Presentación	26/07	26/07	06/08	06/08	24	22	
D.EP .EvR	Evaluación de los Resultados	26/07	26/07	30/07	30/07	12	12	
D.EP .PrR	Presentación de los Resultados	28/07	28/07	06/08	06/08	12	10	

4.2.2.3. Planificación de los Recursos (G.PI.PIR)

En esta actividad se determina los recursos, tanto materiales como humanos para el desarrollo del proyecto.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (G.PI.PIR).

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Información de Entrada

_Recursos Humanos Involucrados (G.IP.EIP.ReHI). Sección 4.2.1.1.1.

_Reporte de Evaluación de Herramientas (G.IP.EvS.EvHe). Sección 4.2.1.3.1.

_Plan de Acción (G.PI.PIA.PIAC). Sección 4.2.2.2

_Problema del Negocio (D.EN.CPN.PrNe). Sección 4.3.1.2.1.

_Fuentes de Información del Cliente (D.EN.AnN.FuIC). Sección 4.3.1.1.1.

Información de Salida

_Plan de Necesidad de Recursos (G.PI.PIR.PINR). Sección 4.2.2.3.1.

4.2.2.3.1. Plan de Necesidad de Recursos (G.PI.PIR.PINR)

Se registran la cantidad de recursos, la fecha de inicio y fin en el que son requeridos, y la descripción de los mismos.

Se planifica contar con un líder de proyecto quien es la directora de la tesis, dos especialistas en minería de datos: uno corresponde a la Tesista y el otro forma parte del equipo de investigación de la Facultad, y dos expertos en el dominio del problema: el Jefe de Análisis y documentación de Sistema y le médico de la UCIC del Instituto de Cardiología.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

En la Tabla 33 se ilustra el formalismo del caso de estudio correspondiente al plan de necesidades de recursos.

Tabla 33: G. Plan de Necesidad de Recursos.

PLAN DE NECESIDAD DE RECURSOS					
Responsable	Sofía Vallejos		Fecha	01/08/21	
ID	G.PI.PIR.PINR		Versión	1.1	
Recursos Humanos					
ID	Recurso	Cant.	F. Inicio	F. Fin	Descripción
ReHI-01 y ReHI-02	Expertos	2	01/06/21	10/08/21	Experto en el Dominio del Problema
ReHI-03	Data miner	1	01/06/21	10/08/21	Experta en Explotación de la Información
ReHI-01 y ReHI-02	Tesista y colaborador	2	01/06/21	10/08/21	Junio en Explotación de la Información
Recursos Materiales					
ID	Recurso	Cant.	F. Inicio	F. Fin	Descripción
ReMa	Computadora Personal	2	01/06/21	16/08/21	SO Windows 10 RAM 4gb o más 10GB o más en disco

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.2.2.4. Planificación de las Responsabilidades (G.PI.PRe)

En esta actividad se definen las responsabilidades y obligaciones de los involucradas en el proyecto.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (G.PI.PRe).

Información de Entrada

_Recursos Humanos Involucrados (G.IP.EIP.ReHI). Sección

4.2.1.1.1.

_Plan de Comunicación

(G.IP.DCo.PCom). Sección

4.2.1.2.1.

_Plan de Acción (G.PI.PIA.PIAc).

Sección 4.2.2.2

_Plan de Necesidad de Recursos

(G.PI.PIR.PINR). Sección

4.2.2.3.1.

_Objetivo del Proyecto

(D.EN.AnN.ObPr). Sección

4.3.1.1.3.

_Criterios de éxito del proyecto

(D.EN.AnN.CrEP). Sección

4.3.1.1.4.

_Expectativas del proyecto

(D.EN.AnN.ExPr). Sección

4.3.1.1.5.

_Restricciones del proyecto

(D.EN.AnN.RePr). Sección

4.3.1.1.7.

Información de Salida

_Plan de Necesidad de Recursos (G.PI.PIR.PINR). Sección

4.2.2.3.1.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

_Problema del Negocio
(D.EN.CPN.PrNe). Sección

4.3.1.2.1.

_Criterios de Éxito del Problema
de Negocio (D.EN.CPN.CEPN).
Sección 4.3.1.2.2.

_Riesgos del Proyecto
(G.IP.EIP.RiPr). Sección

4.2.1.1.2.

_Plan de Contingencias
(G.IP.EIP.PiCo). Sección

4.2.1.1.3.

4.2.2.4.1. Matriz de Responsabilidades (G.PI.PRe.MaRe)

En esta matriz se asigna el rol a cada recurso humano que debe emplear para desempeñar cada actividad dada. Es necesario que la actividad tenga un responsable (R), Un mismo recurso puede tener más de un rol para una tarea. En la Tabla 34 se ilustra el formalismo del caso de estudio correspondiente a la matriz de responsabilidades.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 34: G. Matriz de Responsabilidades.

MATRIZ DE RESPONSABILIDADES						
Responsable	Sofia Vallejos	Fecha	22/06/21			
ID	G.Pl.PRe.MaRe	versión	1.0			
Descripción						
Niveles de participación:						
(R) Responsable: encargado de las tareas asociadas a la actividad.						
(E) Ejecución: asignado tareas asociadas a la actividad.						
(A) Aprobación: aceptación Final del resultado de la actividad.						
(C) Consultado: posee conocimiento relevante para el desarrollo de la actividad.						
(I) Informado: requiere estar alerta del progreso de la actividad.						
ID	Actividad	Re HI- 01	Re HI- 02	Re HI- 03	Re HI- 04	Re HI- 05
G.IP	Iniciación					
G.IP.EIP	Exploración Inicial del Proyecto			R	E	I
G.IP.DCo	Definición de la Comunicación	I	I	R	E	I
G.IP.EvS	Evaluación de la Situación	I	I	R	E	I
G.IP.DCV	Definición del Ciclo de Vida			R	E	I
G.Pl	Planificación					
G.Pl.PIM	Planificación de la Mediciones			R	E	I
G.Pl.PIA	Planificación de las Actividades			R	E	I
G.Pl.PIR	Planificación de los Recursos	I	I	R	E	I
G.Pl.PRe	Planificación de las Responsabilidades	A	A	R	E	I
G.So	Soporte					

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

MATRIZ DE RESPONSABILIDADES						
Responsable	Sofia Vallejos	Fecha		22/06/21		
ID	G.Pl.PRe.MaRe	versión		1.0		
Descripción						
Niveles de participación:						
(R) Responsable: encargado de las tareas asociadas a la actividad.						
(E) Ejecución: asignado tareas asociadas a la actividad.						
(A) Aprobación: aceptación Final del resultado de la actividad.						
(C) Consultado: posee conocimiento relevante para el desarrollo de la actividad.						
(I) Informado: requiere estar alerta del progreso de la actividad.						
ID	Actividad	Re HI- 01	Re HI- 02	Re HI- 03	Re HI- 04	Re HI- 05
G.So.MeP	Mediciones del Proyecto			R	E	I
G.So.GeC	Gestión de la Configuración			R	E	I
G.Co	Control					
G.Co.GeD	Gestión del Desarrollo	I	I	R	E	I
G.Co.CoA	Control de las Actividades	I	I	R	E	I
G.Co.GCa	Gestión del Cambio	I	I	A	R	I
G.Ci	Cierre					
G.Ci.FEC	Formalización Externa del Cierre del Proyecto	A	A	R	E	I
G.Ci.FIC	Formalización Interna del Cierre del Proyecto	C	C	R	E	I
D.EN	Entendimiento del Negocio					
D.EN.AnN	Análisis del Negocio	C	R	E	E	I

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

MATRIZ DE RESPONSABILIDADES						
Responsable	Sofia Vallejos	Fecha	22/06/21			
ID	G.Pl.PRe.MaRe	versión	1.0			
Descripción						
Niveles de participación:						
(R) Responsable: encargado de las tareas asociadas a la actividad.						
(E) Ejecución: asignado tareas asociadas a la actividad.						
(A) Aprobación: aceptación Final del resultado de la actividad.						
(C) Consultado: posee conocimiento relevante para el desarrollo de la actividad.						
(I) Informado: requiere estar alerta del progreso de la actividad.						
ID	Actividad	Re HI- 01	Re HI- 02	Re HI- 03	Re HI- 04	Re HI- 05
D.EN.CPN	Comprensión del Problema de Negocio	C	R	E	E	I
D.ED	Entendimiento de los Datos					
D.ED.AnD	Análisis de los Datos	C	R	E	E	C
D.ED.ExD	Exploración de los Datos	R	C	E	E	C
D.ED.EvD	Evaluación de los Datos	C	I	R	E	C
D.Mo	Modelado					
D.Mo.MoP	Modelado del problema			A	R	C
D.Mo.CoM	Configuración del Modelo			A	R	C
D.PD	Preparación de los Datos					
D.PD.CFT	Construcción de la Fuente Temporal de Datos	C/A	C/A	C	R	C

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

MATRIZ DE RESPONSABILIDADES						
Responsable	Sofia Vallejos	Fecha		22/06/21		
ID	G.Pl.PRe.MaRe	versión		1.0		
Descripción						
Niveles de participación:						
(R) Responsable: encargado de las tareas asociadas a la actividad.						
(E) Ejecución: asignado tareas asociadas a la actividad.						
(A) Aprobación: aceptación Final del resultado de la actividad.						
(C) Consultado: posee conocimiento relevante para el desarrollo de la actividad.						
(I) Informado: requiere estar alerta del progreso de la actividad.						
ID	Actividad	Re HI- 01	Re HI- 02	Re HI- 03	Re HI- 04	Re HI- 05
D.PD.AFT	Adecuación de la Fuente Temporal de Datos	C/A	C/A	C	R	C
D.Im	Implementación					
D.Im.SeM	Selección del Modelo	I	I	A	R	I
D.Im.ExI	Explotación de Información	I	I	A	R	I
D.EP	Evaluación y Presentación					
D.EP.EvR	Evaluación de los Resultados	C/I	C/I	R	E	C
D.EP.PrR	Presentación de los Resultados	C/I	C/I	R	E	I

4.2.2.4.2. Propuesta del Proyecto (G.Pl.PRe.PrPr)

Se registran los aspectos a acordar entre las partes para dar inicio formal al desarrollo del proyecto. En la Tabla 35 se ilustra el formalismo del caso de estudio correspondiente a la propuesta del proyecto.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 35: G. Propuesta del Proyecto.

PROPUESTA DEL PROYECTO			
Responsable	Sofia Vallejos	Fecha	26/06/21
ID	G.Pl.PRe.PrPr	Versión	1.0
Esfuerzo			
Alcance	Se establece como objetivo del proyecto implementar una metodología de explotación de información para la identificación de patrones de comportamiento de pacientes de la unidad de cuidados intensivos coronarios.		
Obligaciones y Responsabilidades	<p>Los expertos en el dominio del problema se comprometen brindar la información requerida en los tiempos establecidos.</p> <p>La contraparte, asumen el compromiso con la responsabilidad de brindar solución a las problemáticas planteadas por el cliente en el período convenido, además se compromete a informar sobre los avances del proyecto.</p> <p>Se acuerda como fecha de finalización del proyecto: 04 / 08 / 21</p>		
Firma del Contratante		Firma de la Contraparte	
Aclaración:		Aclaración:	

4.2.3. Soporte (G.So)

Las actividades correspondientes a esta fase se realizan de manera transversal al ciclo de vida.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

A continuación, se describen las dos actividades del subproceso utilizando el formalismo propuesto, adaptados al caso de estudio.

4.2.3.1. Mediciones del Proyecto (G.So.MeP)

En esta actividad se realizan los cálculos de métricas necesarias durante el desarrollo del proyecto, en donde se deja registro del progreso de los indicadores. Por ello es de suma importancia para la colaboración en la toma de decisiones del líder del proyecto.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (G.So.MeP).

Información de Entrada

_Listado de Métricas
(G.Pl.PIM.LiMe). Sección
4.2.2.1.1.

Información de Salida

_Registro de Mediciones
(G.So.MeP.ReMe). Sección
4.2.3.1.1.

4.2.3.1.1. Registro de Mediciones (G.So.MeP.ReMe)

En esta actividad se registran los resultados obtenidos de las métricas.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

En la Tabla 36 se ilustra el formalismo del caso de estudio correspondiente al registro de mediciones.

Tabla 36: G. Registro de Mediciones.

REGISTRO DE MEDICIONES			
Responsable	Sofia Vallejos	Fecha	08/08/21
ID	G.So.MeP.ReMe	Versión	1.1
Indicador	Medición	Descripción	
Tiempo total requerido para el desarrollo del proyecto	DRPY = 2,16	Tdesarrollo = 1,88	
		Tgestión = 0,28	
Grado de Utilidad de Atributos	GUA = 0,846	NA (T)= 26	
		NANC (T) = 0	
		NANS (T) = 4	
		NAUD (T) = 0	

Grado de Utilidad de Atributos: se considera que cuando la medida de la métrica toma un valor ALTO (un valor superior a 0.70) significa que la mayoría de los atributos son de utilidad y significativos para el proyecto y no se requieren demasiadas correcciones [51]. Por ello y tomando el valor de GUA obtenido en el caso de estudio se puede decir que los atributos son significativos.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.2.3.2. Gestión de la Configuración (G.So.GeC)

En esta actividad se integran los resultados al documento del proyecto y se determina el estado actual del sistema permitiendo la trazabilidad del mismo.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (G.So.GeC).

Información de Entrada

_Reglas de versionado (externo).
_Reporte de Evaluación del Cambio (G.Co.Gca.RECa). Sección 4.2.4.3.1.
_Modelo de Ciclo de Vida (G.IP.DCV.MCV). Sección 4.2.1.4.1.
_Plan de Acción (G.Pl.PIA.PIAC). Sección 4.2.2.2.2.

Información de Salida

_Reporte de Versionado (G.So.GeC.ReVe). Sección 4.2.3.2.1.
_Informe del Estado de la Configuración (G.So.GeC.InEC). Sección 4.2.3.2.2.

4.2.3.2.1. Reporte de Versionado (G.So.GeC.ReVe)

Se deja registro de los cambios de versión en los distintos productos internos del proyecto, se registra la fecha en la cual se realiza el registro de la versión, el elemento afectado, el número de versión previa (VP) y versión actual (VA), los cambios asociados y la descripción.

Reglas de Versionado:

Se utilizan dos dígitos para reflejar el progreso de los productos a lo largo del proyecto X.Y. El primer elemento (X) indica alteraciones en las necesidades o estrategias de ejecución del proyecto, mientras que el segundo (Y) representa iteraciones en el ciclo de vida. X representa al valor mayor del documento, se incrementan de a uno cuando se evidencia alteraciones o eliminación que puede sufrir un elemento del producto, Y representa la versión menor del documento, se incrementa de a uno cada vez que se evidencian incorporaciones o cambios que no modifiquen la alteración del producto. En caso que el producto se encuentre en un estadio temprano, el cual no puede ser utilizado para su uso como entrada en otras tareas, este dígito debe ser indicado con cero, cuando se modifique (X), (Y) retomará al valor cero.

Para el registro del estado del proyecto, se utiliza la misma lógica (X.Y). En la Tabla 37 se ilustra el formalismo del caso de estudio correspondiente a Reporte de Versionado.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 37: G. Reporte de Versionado.

REPORTE DE VERSIONADO					
Responsable		Sofia Vallejos		Fecha	06/08/21
ID		G.So.GeC.ReVe			
Fecha	Elemento	VP	VA	Cambio Asociado	Descripción
01/08/21	Plan de acción	1.0	1.1	Ajuste por reporte de estado y Reporte de Evaluación del Cambio(RECa-1)	
01/08/21	Registro de Mediciones	1.0	1.1	Ajuste por reporte de estado	
01/08/21	Plan de Necesidad de Recursos	1.0	1.1	Ajuste de la necesidad de recursos causados por riesgo (RiPr-02).	Imposibilidad de realizar la presentación formal de los resultados en la fecha estipulada, por ausencia del experto.

4.2.3.2.2. Informe del Estado de la Configuración (G.So.GeC.InEC)

Por no haber surgido cambios en los objetivos y en las iteraciones en el ciclo de vida del proyecto, se identifica una única versión, detallando la versión de cada producto interno. En la Tabla 38 se ilustra el formalismo del caso de estudio correspondiente al informe del estado de la configuración.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 38: G. Informe del Estado de la Configuración.

INFORME DEL ESTADO DE LA CONFIGURACIÓN			
Responsable	Sofia Vallejos	Fecha	06/08/21
ID	G.So.GeC.ReVe		
ID	Actividad	Elemento	Versión del Proyecto
G.IP	Iniciación		
G.IP.EIP	Exploración Inicial del Proyecto	Recursos Humanos Involucrados	1.0
		Riesgos del Proyecto	1.0
		Plan de Contingencias	1.0
G.IP.DCo	Definición de la Comunicación	Plan de Comunicación	1.0
G.IP.EvS	Evaluación de la Situación	Reporte de Evaluación de Herramientas	1.0
		Reporte de Evaluación de Viabilidad	1.0
G.IP.DCV	Definición del Ciclo de Vida	Modelo de Ciclo de Vida	1.0
G.PI	Planificación		
G.PI.PIM	Planificación de la Mediciones	Listado de Métricas	1.0
		Estimación del Proyecto	1.0
G.PI.PIA	Planificación de las Actividades	Mapa de Actividades	1.0
		Plan de Acción	1.1
G.PI.PIR	Planificación de los Recursos	Plan de Necesidad de Recursos.	1.1
G.PI.PRe	Planificación de las Responsabilidades	Matriz de Responsabilidades	1.0
		Propuesta del Proyecto	1.0
G.So	Soporte		
G.So.MeP	Mediciones del Proyecto	Registro de Mediciones	1.1
G.So.GeC		Reporte de Versionado	-

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

INFORME DEL ESTADO DE LA CONFIGURACIÓN			
Responsable	Sofia Vallejos	Fecha	06/08/21
ID	G.So.GeC.ReVe		
ID	Actividad	Elemento	Versión del Proyecto
	Gestión de la Configuración	Informe del Estado de la Configuración	-
G.Co	Control		
G.Co.GeD	Gestión del Desarrollo	Reporte de Estado	-
G.Co.CoA	Control de las Actividades	Registro de Riesgos Acontecidos	-
G.Co.GCa	Gestión del Cambio	Reporte de Evaluación del Cambio	-
G.Ci	Cierre		
G.Ci.FEC	Formalización Externa del Cierre del Proyecto	Documento de aceptación	-
G.Ci.FIC	Formalización Interna del Cierre del Proyecto	Reporte de Cierre	-
D.EN	Entendimiento del Negocio		
D.EN.AnN	Análisis del Negocio	Fuentes de Información del Cliente	1.0
		Definiciones, Acrónimos y Abreviaturas	1.0
		Objetivo del Proyecto	
		Criterios de éxito del proyecto	1.0
		Expectativas del proyecto	1.0

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

INFORME DEL ESTADO DE LA CONFIGURACIÓN			
Responsable	Sofia Vallejos	Fecha	06/08/21
ID	G.So.GeC.ReVe		
ID	Actividad	Elemento	Versión del Proyecto
		Suposiciones del proyecto	1.0
		Restricciones del proyecto	1.0
D.EN.CPN	Comprensión del Problema de Negocio	Problema del Negocio.	1.0
		Criterios de Éxito del Problema de Negocio.	1.0
D.ED	Entendimiento de los Datos		
D.ED.AnD	Análisis de los Datos	Diccionario de Fuente de Datos.	1.0
		Campos Relacionados con el Problema de Negocio	1.0
D.ED.ExD	Exploración de los Datos	Reporte de Datos Explorados	1.0
		Fuente Integrada de datos	-
D.ED.EvD	Evaluación de los Datos	Reporte de calidad de los datos	1.0
D.Mo	Modelado		
D.Mo.MoP	Modelado del problema	Diseño del Problema de Explotación	1.0
D.Mo.CoM	Configuración del Modelo	Selección de Algoritmos de Explotación de Información	1.0
		Selección de Variables del Modelo	1.0
		Estrategias de Evaluación de Modelos	1.0

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

INFORME DEL ESTADO DE LA CONFIGURACIÓN			
Responsable	Sofia Vallejos	Fecha	06/08/21
ID	G.So.GeC.ReVe		
ID	Actividad	Elemento	Versión del Proyecto
D.PD	Preparación de los Datos		
D.PD.CFT	Construcción de la Fuente Temporal de Datos	Reporte de Generación de la Fuente Temporal de datos	1.0
		Fuente Temporal de Datos	-
D.PD.AFT	Adecuación de la Fuente Temporal de Datos	Reporte de Adecuación de la Fuente Temporal de Datos	1.0
D.Im	Implementación		
D.Im.SeM	Selección del Modelo	Reporte de Estrategia de Parametrización del Modelo	1.0
D.Im.ExI	Explotación de Información	Reporte de Implementación del Modelo	1.0
D.EP	Evaluación y Presentación		
D.EP.EvR	Evaluación de los Resultados	Reporte de Evaluación de los Resultados	1.0
D.EP.PrR	Presentación de los Resultados	Reporte del Proyecto	1.0

4.2.4. Control (G.Co)

En la fase Control se verifica que el desarrollo del proyecto se realice de acuerdo a lo planificado y pactado con el

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

cliente, en él se asientan cambios y posibles riesgos que pudieran ocurrir.

4.2.4.1. Gestión del Desarrollo (G.Co.GeD)

En esta actividad se realiza el registro del progreso del proyecto. En él se puede observar el cumplimiento de lo planificado y reajustar en el caso que sea necesario. previstas.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (G.Co.GeD).

Información de Entrada

_Plan de Acción (G.Pl.PIA.PIAC).
Sección 4.2.2.2.2.
_Registro de Mediciones
(G.So.MeP.ReMe). Sección
4.2.3.1.1.

Información de Salida

_Reporte de Estado
(G.Co.GeD.ReEs). Sección
4.2.4.1.1 .

4.2.4.1.1. Reporte de Estado (G.Co.GeD.ReEs)

En esta etapa se utiliza la técnica que proporciona un resumen del desempeño del proyecto para comunicar de manera eficiente el progreso del mismo propuesto en [52].

En la Tabla 39 se ilustra el formalismo del caso de estudio correspondiente al reporte de estado.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 39: G. Reporte de Estado.

REPORTE DE ESTADO			
Responsable	Sofia Vallejos	Fecha	01/08/21
ID	G.Co.GeD.ReEs		
Evaluación del Programa			
Global	- 22,93% (por debajo de lo planificado)		
Desarrollo		Gestión	0,00%
Entendimiento del Negocio	-16,67%	Iniciación	0,00%
Entendimiento de los Datos	-3,33%	Planificación	0,00%
Modelado	-18,18	Soporte	-
Preparación de los Datos	-18,75%	Control	-
Implementación	-6,25%	Cierre	-
Evaluación y Presentación	-54,16%		
Descripción:			
Situaciones identificadas que requieren de seguimiento: -.			
Cambios durante el periodo (alcances, tiempos): -.			
Logros principales durante el periodo: Definición de los alcances del proyecto y los problemas de negocio.			

4.2.4.2. Control de las Actividades (G.Co.CoA)

En esta etapa se realiza un seguimiento, control y registro de acontecimientos, así como de las acciones realizadas.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (G.Co.CoA).

Información de Entrada

_Riesgos del Proyecto
(G.IP.EIP.RiPr). Sección 4.2.1.1.2.
_Plan de Contingencias
(G.IP.EIP.PiCo)Sección 4.2.1.1.3.

Información de Salida

_Registro de Riesgos Acontecidos
(G.Co.CoA.ReRA). Sección
4.2.3.1.1.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

_Plan de Acción (G.Pl.PIA.PIAC).
Sección 4.2.2.2.2.

_Registro de Mediciones
(G.So.MeP.ReMe). Sección
4.2.3.1.1.

_Registro de Mediciones
(G.So.MeP.ReMe). Sección
4.2.3.1.1.

_Reporte de Estado
(G.Co.GeD.ReEs).Sección
4.2.4.1.1.

4.2.4.2.1. Registro de Riesgos Acontecidos (G.Co.CoA.ReRA)

Se registran las situaciones riesgosas que ocurrieron en el proyecto. En donde se especifica: la fecha del acontecimiento, el identificador del riesgo, el responsable de la resolución, descripción del riesgo ocurrido, las acciones a realizar para solucionar la eventualidad y la fecha en la cual la misma fue resuelta.

En la Tabla 40 se ilustra los riesgos acontecidos durante desarrollo del proyecto.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 40: G. Registro de Riesgos Acontecidos.

REGISTRO DE RIESGOS ACONTECIDOS					
Responsable		Sofia Vallejos		Fecha	01/08/21
ID		G.Co.CoA.ReRA			
Fecha	Riesgo	Respon- sable	Descripción	Acción	Fecha Resolución
01/08/21	RiPr-02	Médico	Imposibilidad de realizar la presentación de los resultados en la fecha estipulada, por ausencia del experto.	Posponer 2 días las actividades pendientes	1/08/21

4.2.4.3. Gestión del Cambio (G.Co.GCa)

Se realiza la evaluación formal de las peticiones de cambio, determinando como resultado la procedencia o no del mismo y sus efectos asociados.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (G.Co.GCa).

Información de Entrada

_Solicitud de Cambio (Externo)
_Propuesta del Proyecto
(G.PI.PRe.PrPr). Sección 4.2.2.4.2.

Información de Salida

_Reporte de Evaluación del Cambio
(G.Co.Gca.RECa). Sección
4.2.4.3.1.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.2.4.3.1. Reporte de Evaluación del Cambio (G.Co.Gca.RECa)

En esta etapa se registra las solicitudes de cambios realizadas. El formalismo está compuesto por la fecha de solicitud del cambio, la solicitud, la/s persona/s que lo solicitan, una descripción, el impacto del cambio que causa sobre el proyecto, la evaluación y el responsable del mismo.

Tabla 41: G. Reporte de Evaluación del Cambio

REPORTE DE EVALUACIÓN DEL CAMBIO						
Responsable		Sofia Vallejos		Fecha		1/08/21
ID		G.Co.Gca.RECa				
Fecha	Cambio	Pedido por	Descripción	Impacto	Estado	Responsable
1/08/21	Ajuste de plazos de hitos del proyecto (RECa-1)	Médico	Imposibilidad de realizar la presentación formal de los resultados en la fecha estipulada, por ausencia del experto.	Modificación en la fecha de las actividades pendientes en dos días. La fecha de Formalización Externa del Cierre del Proyecto, se cambia al 06/08/21.	Aprobado	Sofia Vallejos

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.2.5. Cierre (G.Ci)

En esta fase se deja registro formal de la culminación de todas las actividades relacionadas con el proyecto, evaluando los resultados obtenidos y el proceso.

4.2.5.1. Formalización Externa del Cierre del Proyecto (G.Ci.FEC)

En esta actividad se formaliza la conformidad del cliente, respecto de los compromisos asumidos en la propuesta del proyecto, dejando registro de la culminación del mismo.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (G.Ci.FEC).

Información de Entrada

_Reporte de Evaluación de los Resultados (D.EP.EvR.ReER). Sección 4.3.6.1.1.

_Registro de Riesgos Acontecidos (G.Co.CoA.ReRA). Sección 4.2.4.2.1.

_Plan de Acción (G.Pl.PIA.PIAC). Sección 4.2.2.2.2.

_Propuesta del Proyecto (G.Pl.PRe.PrPr). Sección 4.2.2.4.2.

Información de Salida

_ Documento de Aceptación (G.Ci.FEC.DoAc). Sección 4.2.5.1.1.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.2.5.1.1. Documento de Aceptación (G.Ci.FEC.DoAc)

Se describen los resultados derivados del desarrollo del proyecto. En la Tabla 42 se ilustra el formalismo del caso de estudio correspondiente al Documento de Aceptación.

Tabla 42: G. Documento de Aceptación.

DOCUMENTO DE ACEPTACIÓN			
Responsable	Sofia Vallejos	Fecha	06/08/21
ID	G.Ci.FEC.DoAc		
Objetivos	El objetivo planteado consiste en analizar las historias clínicas electrónicas de una unidad de cuidado intensivos coronarios, con el propósito de identificar reglas de comportamiento que permitan comprender a los pacientes que requieren el grado de cuidados intensivos coronarios. Cabe resaltar que la evaluación realizada por el experto (médico) determinó que las reglas de comportamiento identificadas colaboran con la toma de decisiones de los médicos de la UCIC. En efecto, permite determinar el estado de gravedad de un paciente y conducta a tomar según su estado, reduciendo sustancialmente el tiempo que un médico debe invertir para identificarlo.		
Programa	Se realizaron dos informes del progreso de las actividades de manera mensual. La fecha de entrega del proyecto se pactó el día 04/08/21, la cual debió ser pospuesta al 6/08/2021 por pedido del médico.		
Conclusión	El proyecto fue finalizado con un desvío de dos días dentro de las fechas estipuladas.		

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

	Podríamos resumir a continuación que por medio de este documento se deja constancia que se ha cumplimentado exitosamente los requerimientos realizados, dando por finalizado el proyecto.	
Firma		
Aclaración	J. Parras (Médico)	M. I. Pisarello (Dataminer del Proyecto)

4.2.5.2. Formalización Interna del Cierre del Proyecto (G.Ci.FIC).

En esta etapa se evalúa el desarrollo del proyecto, dejando registro de aquellos aspectos relevantes que puedan ser tenidos en cuenta en proyectos futuros.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (G.Ci.FIC).

Información de Entrada

_Plan de Acción (G.Pl.PIA.PIAc).
Sección 4.2.2.2.2.

_Matriz de Responsabilidades
(G.Pl.PRe.MaRe). Sección
4.2.2.4.1.

_Registro de Mediciones
(G.So.MeP.ReMe). Sección
4.2.3.1.1.

_Registro de Riesgos Acontecidos
(G.Co.CoA.ReRA). Sección
4.2.4.2.1.

Información de Salida

_Reporte de Cierre
(G.Ci.FIC.ReCi). Sección 4.2.5.2.1.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

_ Reporte de Evaluación del Cambio
(G.Co.Gca.RECa). Sección
4.2.4.3.1.

_ Reporte de Evaluación de los
Resultados (D.EP.EvR.ReER).
Sección 4.3.6.1.1.

_ Documento de Aceptación
(G.Ci.FEC.DoAc). Sección
4.2.5.1.1.

4.2.5.2.1. Reporte de Cierre (G.Ci.FIC.ReCi)

En esta etapa se exhibe un resumen formal del proyecto, apoyado en los propuestos en [52]. Se desarrollan los objetivos pactados con el cliente y los resultados obtenidos, el tiempo estimado y los valores reales para cada etapa del proyecto, realizando un análisis del porcentaje de diferencia entre ambos y registrando el motivo que derivó la diferencia identificada. Los hitos a evaluar en la sección “Evaluación del tiempo” y por último se listan los principales desafíos y aprendizajes realizados durante el proyecto. En la Tabla 43 se ilustra el formalismo del caso de estudio correspondiente al Reporte de Cierre.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 43: G. Reporte de Cierre

REPORTE DE CIERRE				
Responsable	Sofia Vallejos	Fecha	08/08/21	
ID	G.Ci.FIC.ReCi			
Objetivos del Proyecto				
Objetivos		Resultados		
Identificar patrones de comportamiento de pacientes que requieren el grado de cuidados intensivos coronarios.		Identificación de reglas de comportamiento de interés para el experto (médico). En este sentido se evidencia que las reglas permiten comprender los aspectos generales de los pacientes en la unidad de cuidados intensivos coronarios.		
Evaluación del Tiempo (en HS.)				
Hito	Estimado	Real	%desvío	Motivo
Gestión	48	52	8,33%	
Iniciación	12	12	0,00%	
Planificación	10	10	0,00%	
Soporte	8	14	57,14%	
Control	12	10	-20,00%	
Cierre	6	6	0,00%	
Desarrollo	218	194	-11,01%	
Entendimiento del Negocio	48	40	-16,67%	
Entendimiento de los Datos	60	58	-3,33%	
Modelado	22	18	-18,18	
Preparación de los Datos	32	26	-18,75%	
Implementación	32	30	-6,25%	
Evaluación y Presentación	24	22	-8,33%	
TOTAL	266	246	-7,52%	
Principales Desafíos				
Estimación de los tiempos: el modelo utilizado para el proceso de desarrollo sobreestimó en aproximadamente un 11% el esfuerzo requerido en la carga de trabajo, en cambio, en el proceso gestión subestimo los tiempos requeridos cercano al 7,5%.				
Lecciones Aprendidas: -.				

4.3. Subproceso Desarrollo (D)

El subproceso de desarrollo está compuesto por un conjunto de fases y actividades orientadas a la generación del producto resultante del proyecto, en donde, se identifican de patrones relevantes, novedosos y de calidad, así como su análisis y comprensión para la generación de conocimiento validables y de interés, que sean relevantes para el proceso de toma de decisión.

4.3.1. Entendimiento del Negocio (D.EN)

En la fase Entendimiento del Negocio o también conocido como Entendimiento del Dominio se inicia identificando las características propias del proyecto tomando como referencia la gerencia de proyectos y no tanto la minería de datos.

Los ítems que a continuación se desarrollan teniendo en cuenta los esquemas de las plantillas exhibidas en la sesión 3.2, página 34.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.3.1.1. Análisis del negocio (D.EN.AnN)

En esta actividad se identifica y comprende las metas del proyecto, en base a las necesidades del solicitante y los interesados del proyecto.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (D.EN.AnN).

Información de Entrada

- _ Discursos de los interesados.
- _ Información de la Organización.

Información de Salida

- _ Fuentes de Información del Cliente (D.EN.AnN.FuIC). Sección 4.3.1.1.1.
- _ Definiciones, Acrónimos y Abreviaturas (D.EN.AnN.DAAb). Sección 4.3.1.1.2.
- _ Objetivo del Proyecto (D.EN.AnN.ObPr). Sección 4.3.1.1.3.
- _ Criterios de éxito del proyecto (D.EN.AnN.CrEP). Sección 4.3.1.1.4.
- _ Expectativas del proyecto (D.EDo.ANe.EPr). Sección 4.3.1.1.5.
- _ Suposiciones del proyecto (D.EDo.ANe.SPr). Sección 4.3.1.1.6.
- _ Restricciones del proyecto (D.EDo.ANe.RPr). Sección 4.3.1.1.7.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.3.1.1.1. Fuentes de Información del Cliente (D.EN.AnN.FuIC)

La historia clínica electrónica debe contener suficiente información para identificar al paciente, justificar el diagnóstico y tratamiento, y documentar los resultados de la asistencia. Además, es fruto de un equipo por lo que, habitualmente, en la confección de la misma intervienen diferentes y diversos profesionales (médicos, enfermeras, fisioterapeutas, farmacéuticos, y otros).

Tabla 44: D. Fuentes de Información del Cliente.

FUENTES DE INFORMACIÓN DEL CLIENTE				
Responsable	Sofia Vallejos		Fecha	12/06/21
ID	D.EN.AnN.FuIC		Versión	1.0
ID	Nombre	Categoría	Responsable	Descripción
FICI-01	DataSet1	Bases de Datos.	Fabián Bobadilla (RHIn-01)	Datos almacenados en distintas bases correspondiente a las HCE. Periodo 2019

Por otra parte, para cumplir con la documentación derivada de la atención al paciente, en ocasiones se presta en circunstancias poco favorables. Por ello la documentación

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

presente en la historia clínica en ocasiones no es la adecuada o no está completa [53].

En la Tabla 44 se ilustra el formalismo del caso de estudio correspondiente a las Fuentes de Información del Cliente.

4.3.1.1.2. Definiciones, Acrónimos y Abreviaturas (D.EN.AnN.DAAb)

Se registra las terminologías específicas del dominio que no sean familiares para el equipo de trabajo.

En la Tabla 45 se ilustra el formalismo del caso de estudio correspondiente a las definiciones, acrónimos y abreviaturas utilizadas en el proyecto.

Tabla 45: D. Definiciones, Acrónimos y Abreviaturas.

DEFINICIONES, ACRONIMOS Y ABREVIATURAS			
Responsable	Sofia Vallejos	Fecha	12/06/21
ID	D.EN.AnN.DAAb	Versión	1.0
Termino	Descripción	Tipo	Referencia
Dataset	El conjunto de datos que conforman la tabla que recoge información de distintas bases de datos.	Definición	Documento para la utilización de la base de datos de pacientes.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

DEFINICIONES, ACRONIMOS Y ABREVIATURAS			
Responsable	Sofia Vallejos	Fecha	12/06/21
ID	D.EN.AnN.DAAb	Versión	1.0
<i>Termino</i>	<i>Descripción</i>	<i>Tipo</i>	<i>Referencia</i>
Piso	Lugar en donde el paciente debe permanecer en vigilancia médica, teniendo en cuenta que se encuentra pendiente la evolución de su enfermedad o el resultado de pruebas diagnósticas.	Definición	Información extraída de la entrevista con los expertos.
Potasio	El potasio es un mineral del cuerpo que colabora a que los nervios y los músculos funcionen.	Definición	Información extraída de la entrevista con los expertos.
Hemoglobina	Los niveles altos de hemoglobina ocasionan que el cuerpo forme demasiados glóbulos rojos, lo que hace que la sangre sea más espesa de lo usual. Esto puede ocasionar coágulos, ataques cardíacos y etc.	Definición	Información extraída de la entrevista con los expertos.
Creatinina	Es una prueba que mide los niveles de creatinina en la sangre o en la orina. Es un producto de desecho	Definición	Información extraída de la entrevista

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

DEFINICIONES, ACRONIMOS Y ABREVIATURAS			
Responsable	Sofia Vallejos	Fecha	12/06/21
ID	D.EN.AnN.DAAb	Versión	1.0
<i>Termino</i>	<i>Descripción</i>	<i>Tipo</i>	<i>Referencia</i>
	presente en la sangre que proviene de los músculos. Los riñones sanos filtran la creatinina de la sangre y la pasan a la orina.		con los expertos.
ECO FEy	Ecocardiograma de fracción de eyección: es una medida del porcentaje de sangre que sale del corazón cada vez que se contrae. El corazón se contrae y se relaja. Cuando se contrae, expulsa sangre de las dos cámaras de bombeo (los ventrículos). Cuando el corazón se relaja, los ventrículos vuelven a llenarse de sangre. Sin importar cuán fuerte sea la contracción, el corazón nunca puede expulsar toda la sangre del ventrículo. La fracción de eyección se refiere al porcentaje de sangre que se bombea de	Definición	Información extraída de la entrevista con los expertos.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

DEFINICIONES, ACRONIMOS Y ABREVIATURAS			
Responsable	Sofia Vallejos	Fecha	12/06/21
ID	D.EN.AnN.DAAb	Versión	1.0
Termino	Descripción	Tipo	Referencia
	un ventrículo lleno con cada latido del corazón.		
Ecocardiografía Doppler tisular	La ecocardiografía Doppler tisular aporta información útil y precisa para la valoración de la función sistólica y diastólica de los ventrículos.	Definición	Información extraída de la entrevista con los expertos.
ECO Aurícula izquierda	La aurícula izquierda recibe sangre rica en oxígeno procedente de los pulmones y vacía la sangre en el ventrículo izquierdo. La información que aporta la ecocardiografía permite una exacta cuantificación del tamaño auricular.	Definición	Información extraída de la entrevista con los expertos.

4.3.1.1.3. Objetivo del Proyecto (D.EN.AnN.ObPr)

La definición de un objetivo para el proyecto permite determinar y tener como referencia el sentido del mismo ya que las siguientes actividades que se definan, estarán enmarcadas al cumplimiento del mismo. Como objetivo principal se plantea identificar los patrones de

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

comportamiento y las características distintivas de pacientes de la unidad de cuidados intensivos coronarios. En la Tabla 46 se ilustra el formalismo del caso de estudio correspondiente al objetivo del proyecto.

Tabla 46: D. Objetivos del Proyecto.

OBJETIVOS DEL PROYECTO			
Responsable	Sofia Vallejos	Fecha	12/06/21
ID	D.EN.AnN.ObPr	Versión	1.0
ID Objetivo	Descripción	Referencia	
ObPr-01	Identificar patrones de comportamiento de pacientes que requieren el grado de cuidados intensivos coronarios.	Análisis de la información recopilada de la entrevista.	

4.3.1.1.4. Criterios de éxito del proyecto (D.EN.AnN.CrEP)

Una vez finalizadas las entrevistas con el médico, se definieron los criterios de éxito del proyecto. Esta definición se realizó analizando y unificando la información relevada de esta última reunión. Para ello, esta tarea aplica parte del proceso de *Criterios de Éxito y Expectativas del Proyecto* correspondiente a la formalización de los criterios de éxitos y expectativas, que se detalla más adelante en la sección.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

En D.EN.AnN.CrEP se estableció un caso de éxito. Se trata del descubrimiento de patrones de comportamiento. En la Tabla 47 se ilustra el formalismo del caso de estudio correspondiente a los criterios de éxito del proyecto.

Tabla 47: D. Criterio de éxito del Proyecto.

CRITERIOS DE ÉXITO DEL PROYECTO			
Responsable	Sofia Vallejos	Fecha	12/06/21
ID	D.EN.AnN.CrEP	Versión	1.0
ID Criterio	Descripción	Objetivo del Proyecto	Referencia
CrEP -01	Identificar aspectos que permitan comprender el comportamiento de los pacientes de la UCIC.	ObPr-01	Análisis de la información recopilada de la entrevista.

4.3.1.1.5. Expectativas del proyecto (D.EN.AnN.ExPr)

La expectativa del proyecto se enmarca en lo que pueda ocurrir después de la ejecución del mismo, por ello se espera la adquisición de conocimiento relevante que permita la generación de acciones necesarias que ayuden a identificar los patrones de comportamiento y las características distintivas de pacientes de la UCIC.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Además de lo expuesto, se pretende generalizar el caso de estudio y con la posibilidad de poder aplicarlo en distintos escenarios. En la Tabla 48 se ilustra el formalismo del caso de estudio correspondiente a las expectativas del proyecto.

Tabla 48: D. Expectativa del Proyecto.

EXPECTATIVAS DEL PROYECTO					
Responsable		Sofia Vallejos		Fecha	12/06/21
ID		D.EN.AnN.ExPr		Versión	1.0
<i>ID</i>	<i>Descripción</i>	<i>Objetivo del Proyecto</i>	<i>Criterio de Éxito del Proyecto</i>	<i>Referencia</i>	
ExPr-02	Colaborar con la extracción de conocimiento de las HCE para que se tomen las medidas necesarias dentro de la UCIC.	ObPr-01	CrEP-01	Análisis de la información recopilada de la entrevista.	

4.3.1.1.6. Suposiciones del proyecto (D.EN.AnN.SuPr)

En esta etapa se menciona un conjunto de hipótesis y conjeturas que pudieran estar relacionadas con los objetivos de negocio del proyecto, estas hipótesis se establecieron a partir de un conocimiento previo del

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

entorno donde se desarrolla el problema de negocio e impactan de forma general sobre los problemas de la unidad de cuidados intensivos coronarios, los datos y los resultados derivados de los mismos. En la Tabla 49 se ilustra el formalismo del caso de estudio correspondiente a Suposición del Proyecto.

Tabla 49: D. Suposición del Proyecto.

SUPOSICIONES DEL PROYECTO			
Responsable	Sofia Vallejos	Fecha	17/06/21
ID	D.EN.AnN.SuPr	Versión	1.0
ID Suposición	Descripción	Objetivo del Proyecto	Referencia
SuPr-01	Los datos almacenados en las bases de datos están correctos, y se los puede acceder sin inconvenientes.	ObPr-01	Análisis de la información recopilada de la entrevista.
SuPr-02	Existen características distintivas en las HCE que pueden impactar en la toma de decisiones de los médicos de UCIC.	ObPr-01	Análisis de la información recopilada de la entrevista.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.3.1.1.7. Restricciones del proyecto (D.EN.AnN.RePr)

En esta etapa se identifican los límites de requisitos los cuales puedan retrasar, alterar o impedir la consecución de los objetivos del negocio. Ellos podrían estar asociados al recurso humano, a los datos, a cuestiones técnicas del proyecto y/o de la organización. En la Tabla 50 se ilustra el formalismo del caso de estudio correspondiente a las Restricciones del Proyecto.

Tabla 50: D. Restricciones del Proyecto.

RESTRICCIONES DEL PROYECTO					
Responsable		Sofia Vallejos		Fecha	17/06/21
ID		D.EN.AnN.RePr		Versión	1.0
ID Restricciones	Tipo	Descripción	Objetivo del Proyecto	Referencia	
RePr-01	Datos	La información está compuesta por diferentes interfaces de carga teniendo en cuenta las necesidades de registro según el ámbito en el que se implemente.	ObPr-01	Análisis de la información recopilada de la entrevista.	

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

RESTRICCIONES DEL PROYECTO				
RePr-02	Alcance	El proyecto no pretende abarcar la totalidad de los datos contenidos en las HCE, sino solo aquellos que tengan incidencia en la UCIC.	ObPr-01	Análisis de la información recopilada de la entrevista.

4.3.1.2. Comprensión del Problema de Negocio (D.EN.CPN)

En esta actividad se describen los aspectos centrales que permitan la visualización de información relevante para apoyar la toma de decisiones. Se determinarán los problemas de negocio a partir de las características del proyecto y los criterios por los cuales se van a considerar exitosos los resultados obtenidos en el problema.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (D.EN.CPN).

Información de Entrada

- _ Discursos de los interesados.
- _ Información de la Organización.

Información de Salida

- _ Problema del Negocio (D.EN.CPN.PrNe). Sección 4.3.1.2.1.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

_ Definiciones, Acrónimos y Abreviaturas (D.EN.AnN.DAAb). Sección 4.3.1.1.2.

_Objetivo del Proyecto (D.EN.AnN.ObPr). Sección 4.3.1.1.3.

_Criterios de éxito del proyecto (D.EN.AnN.CrEP)Sección 4.3.1.1.4.

_Expectativas del proyecto (D.EN.AnN.ExPr)Sección 4.3.1.1.5.

_Suposiciones del proyecto (D.EN.AnN.SuPr).Sección 4.3.1.1.6.

_Restricciones del proyecto (D.EN.AnN.RePr). Sección 4.3.1.1.7.

_Criterios de Éxito del Problema de Negocio (D.EN.CPN.CEPN). Sección 4.3.1.2.2.

4.3.1.2.1. Problema del Negocio (D.EN.CPN.PrNe)

Las UCIC, requieren una infraestructura y equipamiento complejos, así como una dotación de personal superior a la de una sala de hospitalización. Se plantea la necesidad no sólo de dotar a los hospitales de equipamiento para atender correctamente a los pacientes con enfermedad coronaria, sino también la obligación de gestionar con eficiencia los recursos. Dentro de este orden de ideas, se buscan estrategias más óptimas que permitan analizar en

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

profundidad esta problemática a partir de la información que poseen las HCE del Instituto de cardiología. A partir de lo expuesto, se definen las problemáticas a tratar teniendo presente que dichas problemáticas se derivan del objetivo del proyecto. En la Tabla 51 se ilustra el formalismo del caso de estudio correspondiente al Problema del Negocio.

Tabla 51: D. Problema del Negocio.

PROBLEMA DEL NEGOCIO			
Responsable	Sofía Vallejos	Fecha	17/06/21
ID	D.EN.CPN.PrNe	Versión	1.0
ID	Descripción	Referencia	
PrNe-01	Cuáles son las condiciones de los pacientes que requieren el grado de cuidados intensivos coronarios.	Análisis de la información recopilada de la entrevista 1.	

4.3.1.2.2. Criterios de Éxito del Problema de Negocio (D.EN.CPN.CEPN)

En esta tarea se determina bajo qué criterio se consideran exitosos los resultados obtenidos para el problema de Negocio. En la Tabla 52 se exhibe el formalismo desarrollado para asentar el registro y control del Criterios de Éxito del Problema de Negocio.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 52: D. Criterio de Éxito del Problema del Negocio.

CRITERIOS DE ÉXITO DEL PROBLEMA DE NEGOCIO			
Responsable	Sofia Vallejos	Fecha	17/06/21
ID	D.EN.CPN.CEPN	Versión	1.0
ID Criterio	Descripción	Problema Asociado	Referencia
CEPN-01	Identificar aspectos que permitan comprender el comportamiento de los pacientes que se encuentran en la unidad de cuidados intensivos coronarios.	PrNe-01	Análisis de la información recopilada de la entrevista.

4.3.2. Entendimiento de los Datos (D.ED)

En la fase de D.ED se identifican y reúnen los datos iniciales de interés para el problema de negocio y que se utilizarán dentro del proyecto. Se realizará un análisis detallado de las variables para comprender el significado haciendo hincapié en la calidad, la comprensión del significado, sus valores y exploración de los datos.

4.3.2.1. Análisis de los Datos (D.ED.AnD)

En esta actividad se describen las fuentes de datos existentes en el Instituto de Cardiología, se evalúan las variables disponibles en las distintas fuentes de

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

información, para comprender sus significados, valoraciones y se indaga cómo registran los datos. De este modo, se utiliza la técnica “Identificación de atributos relacionados con el *Problema de Negocio* definida en [36].

Los datos se extrajeron de las historias clínicas electrónicas para fundamentar con bases y recursos sólidos el estudio y las investigaciones. A efectos de la validación se seleccionó un dominio de las historias clínicas electrónicas de los pacientes del Instituto de Cardiología.

Una vez analizados los datos junto con el médico y el responsable del área de informática encargado de administrar las fuentes de información del Instituto de Cardiología se identificaron las variables que el modelo contendrá.

Es importante remarcar que la descripción de los campos a utilizar, puede verse restringida por el alcance del problema de negocio o las limitaciones del proyecto.

Los resultados que se describen a continuación se obtuvieron de las entrevistas con los actores involucrados

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

en el proyecto, en particular con el médico de la UCIC (ReHI-02), los miembros vinculados con los problemas de negocio identificados, los expertos a cargo y el conocimiento vinculado con las fuentes de información identificadas en la fase de entendimiento del negocio, así como información del proyecto que tenga impacto en los datos a utilizar (restricciones del proyecto y problema del negocio) y se analizan las fuentes de información disponibles (formalismo Fuentes de Información del Cliente) [35].

A continuación, se describen los insumos de entrada y los elementos de salida que se requieren en D.ED.AnD.

Información de Entrada

- _ Discursos de los interesados.
- _ Información de la Organización.
- _ Fuentes de Información del Cliente (D.EN.AnN.FuIC). Sección 4.3.1.1.1.
- _ Restricciones del proyecto (D.EN.AnN.RePr) Sección 4.3.1.1.7.
- _Problema del Negocio (D.EN.CPN.PrNe). Sección 4.3.1.2.1.

Información de Salida

- _Diccionario de Fuentes de Datos (D.ED.AnD.DiFD). Sección 4.3.2.1.1.
- _Campos Relacionados con el Problema de Negocio (D.ED.AnD.CRPN). Sección 4.3.2.1.2.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.3.2.1.1. Diccionario de Fuentes de Datos (D.ED.AnD.DiFD)

En esta tarea se describen detalladamente como se estructuran los campos identificados en las fuentes de datos, garantizando la correcta comprensión del significado de las variables y de sus valores posibles.

En la Tabla 53 se ilustra el subconjunto seleccionado de los atributos disponibles de las HCE.

Tabla 53: D. Diccionario de Fuente de Datos.

DICCIONARIO DE FUENTES DE DATOS			
Responsable	Sofia Vallejos	Fecha	25/06/21
ID	D.ED.AnD.DiFD	Versión	1.0
Fuente de Información:		FICI-01	
Campo	Tipo	Descripción	
LRHC	Numérico	Identificador único de Historia Clínica Electrónica.	
InCabFIng	Numérico	Ultima fecha de ingreso del paciente a la internación.	
InCabFEgr	Numérico	Ultima fecha de egreso del paciente a la internación.	
PacieNro	Numérico	Identificador único del paciente.	
Edad	Numérico	Edad en años cumplidos del paciente.	
PacieSexo	Texto	Sexo del paciente.	

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

DICCIONARIO DE FUENTES DE DATOS			
Responsable	Sofia Vallejos	Fecha	25/06/21
ID	D.ED.AnD.DiFD	Versión	1.0
Fuente de Información:		FICI-01	
Campo	Tipo	Descripción	
InterEnUCIC	Numérico	Cantidad de veces que el paciente ha sido internado en UCIC.	
InterEnEMER	Numérico	Cantidad de veces que el paciente ha sido internado en Emergencias.	
InterEnPISO	Numérico	Cantidad de veces que el paciente ha sido internado en Piso.	
InterEnRCVA	Numérico	Cantidad de veces que el paciente ha sido internado en el área de Recuperación Cardiovascular Adultos.	
InterEnRCVP	Numérico	Cantidad de veces que el paciente ha sido internado en el área de Recuperación Cardiovascular Pediátricos.	
InterEnUTI	Numérico	Cantidad de veces que el paciente ha sido internado en el área de Unidad de Terapia Intensiva.	
Potasio	Numérico	Último valor de laboratorio de Potasio.	
GlobRojos	Numérico	Último valor de laboratorio de conteo de glóbulos rojos.	
Hemoglobina	Numérico	Último valor de laboratorio de hemoglobina.	

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

DICCIONARIO DE FUENTES DE DATOS			
Responsable	Sofia Vallejos	Fecha	25/06/21
ID	D.ED.AnD.DiFD	Versión	1.0
Fuente de Información:		FICI-01	
Campo	Tipo	Descripción	
Glucemia	Numérico	Último valor de laboratorio de glucemia.	
Hematocrito	Numérico	Último valor del laboratorio de hematocrito.	
CreatiSerica	Numérico	Último valor del laboratorio de creatinina sérica.	
ECO_FEY	Numérico	Ultimo valor de la Ecografía de la fracción de eyección.	
ECO_IndDeMa	Numérico	Peso del corazón.	
ECO_Tisular	Numérico	Velocidad de las fibras cardiacas. Músculos cardíacos para bombear la sangre.	
ECO_VolAurIzq	Numérico	Tamaño de la aurícula izquierda.	

4.3.2.1.2. Campos Relacionados con el Problema de Negocio (D.ED.AnD.CRPN)

En esta etapa junto a los expertos se analizaron las variables relevantes para el problema del negocio y las que se generaron a partir de otros datos.

En la Tabla 54 se ilustra el formalismo del caso de estudio correspondiente al Problema de Negocio.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 54: D. Campos Relacionados con el Problema de Negocio.

CAMPOS RELACIONADOS CON EL PROBLEMA DE NEGOCIO			
Responsable	Sofia Vallejos	Fecha	26/06/21
ID	D.ED.AnD.CRPN	Versión	1.0
Problema del Negocio	PrNe-01: Cuáles son las condiciones de los pacientes que requieren el grado de cuidados intensivos coronarios.		
Nombre	Generar	Descripción	Referencia
PacieNro	Numérico	Identificador único del paciente.	FIC1-01
EdadRango	X	1_(20-29) 2_(30-39) 3_(40-49) 4_(50-59) 5_(60-69) 6_(70-79) 7_(80+)	FIC1-01
PacieSexo		Sexo del Paciente	FIC1-01
InterEnUCIC_G	X	Determina si el paciente ha sido internado en UCIC. _Si $InterEnUCIC = 0$ entonces $InterEnUCIC_G = N$ Caso contrario $InterEnUCIC_G = S$	FIC1-01
Potasio	Numérico	Último valor de laboratorio de Potasio.	FIC1-01
GlobRojos	Numérico	Último valor de laboratorio de conteo de glóbulos rojos.	FIC1-01
Hemoglobina	Numérico	Último valor de laboratorio de hemoglobina.	FIC1-01

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

CAMPOS RELACIONADOS CON EL PROBLEMA DE NEGOCIO			
Responsable	Sofía Vallejos	Fecha	26/06/21
ID	D.ED.AnD.CRPN	Versión	1.0
Problema del Negocio	PrNe-01: Cuáles son las condiciones de los pacientes que requieren el grado de cuidados intensivos coronarios.		
Nombre	Generar	Descripción	Referencia
Glucemia	Numérico	Último valor de laboratorio de glucemia.	FIC1-01
Hematocrito	Numérico	Último valor del laboratorio de hematocrito.	FIC1-01
CreatiSerica	Numérico	Último valor del laboratorio de creatinina sérica.	FIC1-01
ECO_FEY	Numérico	Último valor de la Ecografía de la fracción de eyección.	FIC1-01
ECO_IndDeMa	Numérico	Peso del corazón.	FIC1-01
ECO_Tisular	Numérico	Velocidad de las fibras cardíacas. Músculos cardíacos para bombear la sangre.	FIC1-01
ECO_VolAur Izq	Numérico	Tamaño de la aurícula izquierda.	FIC1-01
ECGAnormal	Numérico	Representa la actividad eléctrica del corazón	FIC1-01

4.3.2.2. Exploración de los Datos (D.ED.ExD)

De acuerdo al análisis realizado se identificó al campo de interés, con el objetivo de comprender las características

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

del caso de estudio, por ellos se identificaron relaciones entre las distintas variables estudiadas.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad D.ED.ExD.

Información de Entrada

_Suposiciones del proyecto
(D.EN.AnN.SuPr). Sección
4.3.1.1.6.

_Restricciones del proyecto
(D.EN.AnN.RePr). Sección
4.3.1.1.7.

_Problema del Negocio
(D.EN.CPN.PrNe). Sección
4.3.1.2.1.

_Diccionario de Fuentes de Datos
(D.ED.AnD.DiFD)Sección
4.3.2.1.1.

_Campos Relacionados con el
Problema de Negocio
(D.ED.AnD.CRPN). Sección
4.3.2.1.2.

_Reporte de Evaluación de
Herramientas (G.IP.EvS.EvHe).
Sección 4.2.1.3.1.

Información de Salida

_Reporte de Datos Explorados
(D.ED.ExD.ReDE). Sección
4.3.2.2.2.

_Fuente Integrada de datos
(D.ED.ExD.FuID). Sección
4.3.2.2.1.

4.3.2.2.1. Fuente Integrada de datos (D.ED.ExD.FuID)

En esta actividad se procede a integrar los campos a utilizar en una única fuente de información, la cual posee el conjunto de registros de interés con los campos sin alterar. Es una práctica recomendada, utilizar un campo identificador que permita mantener la trazabilidad de los registros con respecto a la fuente de almacenamiento original. En el caso de estudio, si bien los datos se encontraban integrados en un único elemento, se generan la fuente integrada de datos correspondiente a los problemas de negocio, en donde se redujo el conjunto de datos a 14.

La Fig. 11 se ilustra la representación en un Diagrama Entidad-Relación (DER) de la estructura de la fuente de integrada de datos correspondiente al problema de negocio.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

	Fuente Integrada de datos
	PacieNro: int (6) EdadRango: char (15) PacieSexo: char (1) InterEnUCIC_D: char (1) Potasio: decimal (2:1) Hemoglobina: decimal (2:1) GlobRojos: decimal (2:2) Glucemia: int (3) Hematocrito: int (2) CreatiSerica: decimal (2:2) ECO_FEY: int (2) ECO_IndDeMa: int (3) ECO_Tisular: int (2) ECO_VolAurlzq: int (2)

Fig. 11: Fuente de Diagrama de Datos.

4.3.2.2.2. Reporte de Datos Explorados (D.ED.ExD.ReDE)

Una vez definida la fuente integrada de datos, se describe la distribución de valores para cada atributo significativo para el problema de negocio.

En la Tabla 55 se ilustra el formalismo del caso de estudio correspondiente al Reporte de Datos Explorados.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 55: D. Reporte de Datos Explorados

REPORTE DE DATOS EXPLORADOS				
Responsable	Sofia Vallejos		Fecha	03/07/21
ID	D.ED.ExD.ReDE		Versión	1.0
Problema del Negocio	PrNe-01: Cuáles son las condiciones de los pacientes que requieren el grado de cuidados intensivos coronarios.			
Atributos Cualitativos				
Nombre	Valor		Distribución	
EdadRango	1_(20-29)		1_(20-29), 6	
	2_(30-39)		2_(30-39), 24	
	3_(40-49)		3_(40-49), 63	
	4_(50-59)		4_(50-59), 140	
	5_(60-69)		5_(60-69), 194	
	6_(70-79)		6_(70-79), 171	
	7_(80+)		7_(80+), 21	
PacieSexo	F		F, 473	
	M		M, 217	
InterEnUCIC_G	N		N, 548	
	S		S, 142	
Atributos Cuantitativos				
Nombre	Min	Max	Tendencia Central	Dispersión
Potasio	2,4	7,3	Media = 4,1	Des. Est. = 0,54
GlobRojos	1,69	7,61	Media = 4,21	Des. Est. = 0,82
Hemoglobina	5	18,7	Media = 12,4	Des. Est. = 2,34
Glucemia	59	545	Media = 144	Des. Est. = 70,27
Hematocrito	16	55	Media = 37	Des. Est. =

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

REPORTE DE DATOS EXPLORADOS				
				6,60
CreatiSerica	0,4	8,45	Media = 1,21	Des. Est. = 0,87
ECO_FEY	5	86	Media = 52	Des. Est. = 14,30
ECO_IndDeMa	2	370	Media = 144 (F)	Des. Est. = 54,74
			Media = 119 (M)	Des. Est. = 41,67
ECO_Tisular	4	46	Media = 14	Des. Est. = 6,87
ECO_VolAurIzq	13	99	Media = 41	Des. Est. = 17,83

4.3.2.3. Evaluación de los Datos (D.ED.EvD)

Se registran las variables y/o valores que posean datos incorrectos, nulos, outsiders o anómalos.

Información de Entrada

_Diccionario de Fuentes de Datos (D.ED.AnD.DiFD). Sección 4.3.2.1.1.

_Campos Relacionados con el Problema de Negocio (D.ED.AnD.CRPN). Sección 4.3.2.1.2.

_Reporte de Datos Explorados (D.ED.ExD.ReDE). Sección 4.3.2.2.2.

Información de Salida

_Reporte de calidad de los datos (D.ED.EvD.ReCD). Sección 4.3.2.3.1.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

_Fuente Integrada de datos
(D.ED.ExD.FuID). Sección
4.3.2.2.1.

_Reporte de Evaluación de
Herramientas (G.IP.EvS.EvHe).
Sección 4.2.1.3.1.

4.3.2.3.1. Reporte de calidad de los datos (D.ED.EvD.ReCD)

Se logró identificar características irregulares que pudieran afectar los datos. En la Tabla 56 se describen los campos identificados.

Tabla 56: D. Reporte de calidad de los datos.

REPORTE DE CALIDAD DE LOS DATOS			
Responsable	Sofia Vallejos	Fecha	03/07/21
ID	D.ED.EvD.ReCD	Versión	1.0
Problema del Negocio	PrNe-01: Cuáles son las condiciones de los pacientes que requieren el grado de cuidados intensivos coronarios.		
Nombre	Registros	Tipo	Descripción
GlobRojos	15	Nulos	No se realizó el examen
Hemoglobina	12	Nulos	No se realizó el examen
Glucemia	61	Nulos	No se realizó el examen
Hematocrito	12	Nulos	No se realizó el examen
CreatiSerica	1	Nulos	No se realizó el examen
ECO_IndDeMa	263	Nulos	No se realizó el examen

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

ECO_Tisular	55	Nulos	No se realizó el examen
ECO_VolAurIzq	192	Nulos	No se realizó el examen

4.3.3. Modelado de los Datos (D.Mo)

En esta fase se seleccionan y construyen los modelos más apropiados para dar respuesta al problema propuesto. Se identifica y definió los modelos a utilizar para la extracción de conocimiento.

4.3.3.1. Modelado del problema (D.Mo.MoP)

Se identifican y definen cuáles son los procesos de explotación de información apropiados para extraer el conocimiento requerido para el caso de estudio planteado.

Información de Entrada

_Problema del Negocio
(D.EN.CPN.PrNe). Sección
4.3.1.2.1.

_Diccionario de Fuentes de Datos
(D.ED.AnD.DiFD). Sección
4.3.2.1.1.

_Campos Relacionados con el
Problema de Negocio
(D.ED.AnD.CRPN). Sección
4.3.2.1.2.

Información de Salida

_Diseño del Proceso de Explotación de
Información (D.Mo.MoP.DPEI).
Sección 4.3.3.1.1.

4.3.3.1.1. Diseño del Proceso de Explotación de Información (D.Mo.MoP.DPEI)

En esta etapa se utiliza la Red Semántica del Problema de Explotación de Información desarrollada en [54].

En el caso de estudio a partir del problema de negocio, se reconoce al concepto “paciente”. La acción de interés es la de caracterizar al paciente, en relación a las condiciones de los pacientes que requieren el grado de cuidados intensivos coronarios. En este sentido, se definen como características a considerar, aquellas establecidas en el formalismo, es decir, los campos relacionados con el problema de negocio.

En [18] recomienda que cuando se requiere identificar cuáles son las condiciones para obtener un determinado resultado en el dominio del problema, se plantea la utilización de algoritmos de inducción TDIDT para descubrir las reglas de comportamiento de cada atributo clase.

Unas de las técnicas más comunes de minería de datos son los árboles de decisión perteneciente a la familia de los

TDIDT. Son utilizados para descubrir conocimientos en formato de regla que constituye un modelo que representa el dominio del conocimiento subyacente a los ejemplos disponibles del mismo.

El árbol de decisión puede ser de dos tipos: Árboles de regresión o Árboles de clasificación. La clasificación es la separación de datos de acuerdo a sus características similares uno de otro. Se divide en Clasificación supervisada y clasificación no supervisada. La clasificación implica predecir un cierto resultado basado en una entrada determinada. Los datos se cotejan con el atributo que se encuentra almacenado previamente y en base a supuestos los datos se clasifican. El algoritmo analiza los datos y predice los resultados. Los algoritmos de clasificación incluyen a los árboles de decisión, redes neuronales, redes bayesianas, entre otros [55].

En la Tabla 57 se ilustra el formalismo del caso de estudio correspondiente a Diseño del Proceso de Explotación. Se observa la Red Semántica del Problema de Explotación de Información y se identifica al proceso de descubrimiento

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

de reglas de comportamiento el cual indica la utilización de algoritmos de inducción TDIDT.

Tabla 57: D. Diseño del Proceso de Explotación de Información

DISEÑO DEL PROCESO DE EXPLOTACIÓN DE INFORMACIÓN			
Responsable	Sofía Vallejos	Fecha	07/07/21
ID	D.Mo.MoP.DPEI	Versión	1.0
Problema del Negocio	PrNe-01: Cuáles son las condiciones de los pacientes que requieren el grado de cuidados intensivos coronarios.		
Proceso de Explotación de Información	Descubrimiento de Reglas de Comportamiento		

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.3.3.2. Configuración del Modelo (D.MO.CoM)

En esta actividad se definen los elementos que conforman la estrategia de implementación y evaluación de los distintos modelos para la extracción de patrones vinculados con el problema de negocio. Se selecciona y define el orden en el que se aplicarán los algoritmos de EI, se elige el formato de las variables, del proceso de aplicación para la obtención y evaluación de los resultados obtenidos. A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (D.MO.CoM).

Información de Entrada

_Criterios de éxito del proyecto
(D.EN.AnN.CrEP). Sección
4.3.1.1.4.

_Diccionario de Fuentes de Datos
(D.ED.AnD.DiFD). Sección
4.3.2.1.1.

_Campos Relacionados con el
Problema de Negocio
(D.ED.AnD.CRPN). Sección
4.3.2.1.2.

_Reporte de Datos Explorados
(D.ED.ExD.ReDE). Sección
4.3.2.2.2.

Información de Salida

_Selección de Algoritmos de
Explotación de Información
(D.Mo.CoM.SAEI). Sección
4.3.3.2.1.

_Selección de Variables del Modelo
(D.Mo.CoM.SeVM). Sección
4.3.3.2.2.

_Estrategias de Evaluación de
Modelos (D.Mo.CoM.EsEM).
Sección 4.3.3.2.3.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

__Reporte de calidad de los datos
(D.ED.EvD.ReCD). Sección
4.3.2.3.1.

_Diseño del Proceso de Explotación
de Información (D.Mo.MoP.DPEI).
Sección 4.3.3.1.1.

_Reporte de Evaluación de
Herramientas (G.IP.EvS.EvHe).
Sección 4.2.1.3.1.

4.3.3.2.1. Selección de Algoritmos de Explotación de Información (D.Mo.CoM.SAEI)

En esta etapa se describe el algoritmo de EI que se utilizará para extraer el conocimiento requerido de acuerdo al problema planteado. El campo InterUCIC_G representa la clase “minoritaria”, esto ocasiona un desbalanceo en los datos que utilizaremos para el entrenamiento. Por ello, se empleará un algoritmo de aprendizaje sensible a los costos para la clasificación desequilibrada.

La mayoría de los algoritmos de aprendizaje automático asumen que todos los errores de clasificación errónea cometidos por un modelo son iguales. Este no suele ser el caso de los problemas de clasificación desequilibrada en los que omitir un caso positivo o de clase minoritaria es peor que clasificar incorrectamente un ejemplo de la clase

negativa o mayoritaria. Hay muchos ejemplos del mundo real, como la detección de correo no deseado, el diagnóstico de una afección médica o la identificación de fraudes. En todos estos casos, un falso negativo (falta un caso) es peor o más costoso que un falso positivo. Se plantea entonces, la utilización del algoritmo de aprendizaje sensible a los costos. Este tipo de algoritmo tiene en cuenta los costos de los errores de predicción (y potencialmente otros costos) al entrenar un modelo de aprendizaje automático. Es un campo de estudio que está estrechamente relacionado con el campo del aprendizaje desequilibrado que se ocupa de la clasificación en conjuntos de datos con una distribución de clases sesgada [56].

En [57] se evaluó el desempeño de siete algoritmos (ID3, C4.5, CR-T, Random Tree, Naive Bayes, CS-CRT y CS-MC4) de clasificación de árboles de decisión con un conjunto de datos de hepatitis de la UCI. El estudio determinó que los algoritmos utilizados impactan favorablemente para realizar la categorización de clases

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

binarias de datos clínicos. La exactitud los algoritmos se evaluó mediante una validación cruzada de orden 10.

Dentro de este orden de ideas, conjuntamente con el interés centrado en la familia de los métodos de TDIDT con base a lo desarrollado en *Diseño del Proceso de Explotación de Información (D.Mo.MoP.DPEI)* y de acuerdo a la herramienta seleccionada en la etapa *Reporte de Evaluación de Herramientas (G.IP.EvS.EvHe)*, se identifican dos algoritmos de árboles de decisión, ellos son el C4.5 y el CS-MC4 (Cost sensitive decision tree algorithm). Se seleccionó uno que contempla el costo sensitivo y el otro no, con el fin de evaluar cual obtiene mejor resultado.

En la Tabla 58 se ilustra el formalismo del caso de estudio correspondiente a la Selección de Algoritmos de Explotación de Información.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 58: D. Selección de Algoritmos de Explotación de Información

SELECCIÓN DE ALGORITMOS DE EXPLOTACIÓN DE INFORMACIÓN				
Responsable	Sofia Vallejos		Fecha	08/07/21
ID	D.Mo.CoM.SAEI		Versión	1.0
Problema del Negocio	PrNe-01: Cuáles son las condiciones de los pacientes que requieren el grado de cuidados intensivos coronarios.			
Algoritmo	Input	Target	Estrategia	Descripción
C4.5	Discretos /Continuos	Discretos	1	Algoritmo de la familia de TDIDT, que mediante la inducción, permite descubrir reglas para clasificar a los pacientes de la UCIC.
CS-MC4	Discretos /Continuos	Discretos	1	Algoritmo de la familia de TDIDT, que mediante la inducción, permite descubrir reglas para clasificar a los pacientes de la UCIC.

4.3.3.2.2. Selección de Variables del Modelo (D.Mo.CoM.SeVM)

Se registran los campos que utilizarán los algoritmos. En la Tabla 59 se ilustra el formalismo del caso de estudio correspondiente a Selección de Variables del Modelo.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 59: D. Selección de Variables del Modelo.

SELECCIÓN DE VARIABLES DEL MODELO			
Responsable	Sofia Vallejos	Fecha	10/07/21
ID	D.Mo.CoM.SeVM	Versión	1.0
Problema del Negocio	PrNe-01: Cuáles son las condiciones de los pacientes que requieren el grado de cuidados intensivos coronarios.		
Campo		C4.5/CS-MC4	
		Tipo	Conversión
EdadRango		Input	
PacieSexo		Input	
InterEnUCIC_D		Target	
Potasio		Input	Discretizar
GlobRojos		Input	Discretizar
Hemoglobina		Input	Discretizar
Glucemia		Input	Discretizar
Hematocrito		Input	Discretizar
CreatiSerica		Input	Discretizar
ECO_FEY		Input	Discretizar
ECO_IndDeMa		Input	Discretizar
ECO_Tisular		Input	Discretizar
ECO_VolAurIzq		Input	Discretizar

4.3.3.2.3. Estrategias de Evaluación de Modelos (D.Mo.CoM.EsEM)

Se registran las técnicas a utilizar con el fin de medir y valorar la mejora de los resultados obtenidos por los modelos. Una vez obtenidos los resultados se evalúa la calidad de los procesos con el experto en explotación de información.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 60: D. Estrategias de Evaluación de Modelos

ESTRATEGIAS DE EVALUACIÓN DE MODELOS			
Responsable	Sofia Vallejos	Fecha	10/07/21
ID	D.Mo.CoM.EsEM	Versión	1.0
Problema del Negocio	PrNe-01: Cuáles son las condiciones de los pacientes que requieren el grado de cuidados intensivos coronarios.		
Técnica	Alcance	Descripción	
Cross-Validation	C4.5	10 fold Cross con 10 repeticiones	
Cross-Validation	CS-MC4	10 fold Cross con 10 repeticiones	

En la Tabla 60 se ilustra el formalismo del caso de estudio correspondiente a Estrategias de Evaluación de Modelos.

4.3.4. Preparación de los Datos (D.PD)

En esta fase se transforman los datos preparándolos para su correcta aplicación en los algoritmos de EI. Para ello, incluye el formateo y limpieza de los datos de acuerdo a las restricciones de los modelos a utilizar, la definición de las fuentes de datos a utilizar para las distintas instancias de aplicación y evaluación del modelo y la validación de la composición de la misma por parte de los expertos del negocio.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.3.4.1. Construcción de la Fuente Temporal de Datos (D.PD.CFT)

En esta actividad se generan las fuentes temporales de datos necesarias para las distintas etapas de implementación del modelo. A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (D.PD.CFT).

Información de Entrada

__Reporte de Datos Explorados (D.ED.ExD.ReDE). Sección 4.3.2.2.2.

__Estrategias de Evaluación de Modelos (D.Mo.CoM.EsEM). Sección 4.3.3.2.3.

_Fuente Integrada de datos (D.ED.ExD.FuID). Sección 4.3.2.2.1.

_Reporte de Evaluación de Herramientas (G.IP.EvS.EvHe). Sección 4.2.1.3.1.

__Selección de Variables del Modelo (D.Mo.CoM.SeVM). Sección 4.3.3.2.2.

Información de Salida

_Fuente Temporal de Datos (D.PD.CFT.FuTD). Sección 4.3.4.1.1.

_Reporte de Generación de la Fuente Temporal de datos (D.PD.CFT.RGFT). Sección 4.3.4.1.2.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.3.4.1.1. Fuente Temporal de Datos (D.PD.CFT.FuTD)

Se desarrolla una única fuente temporal de datos. En la Fig. 12 se ilustra la fuente temporal de datos representada mediante una entidad del formalismo DER.

	Fuente Temporal de datos
	PacieNro: int (6) EdadRango: char (15) PacieSexo: char (1) InterEnUCIC_G: char (1) Potasio: char (12) Hemoglobina: char (12) GlobRojos: char (12) Glucemia: char (12) Hematocrito: char (12) CreatiSerica: char (12) ECO_FEY: char (12) ECO_IndDeMa: char (12) ECO_Tisular: char (12) ECO_VolAurlzq: char (12)

Fig. 12: Fuente Temporal de datos.

4.3.4.1.2. Reporte de Generación de la Fuente Temporal de datos (D.PD.CFT.RGFT)

En esta fase se describen las fuentes temporales de datos desarrollada para el problema de negocio, en donde se

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

realiza una descripción estadística de las variables diferenciadas en cualitativo o cuantitativo.

En la Tabla 61 se ilustra el formalismo del caso de estudio correspondiente al Reporte de Generación de la Fuente Temporaria de datos.

Tabla 61: D. Reporte de Generación de la Fuente Temporaria de datos.

REPORTE DE GENERACIÓN DE LA FUENTE TEMPORARIA DE DATOS			
Responsable	Sofia Vallejos	Fecha	15/07/21
ID	D.PD.CFT.RGFT	Versión	1.0
Problema del Negocio	PrNe-01: Cuáles son las condiciones de los pacientes que requieren el grado de cuidados intensivos coronarios.		
Fuente Temporaria de datos	D.PD.CFT.FuTD		
Atributos Cualitativos			
Nombre	Valor	Distribución	
EdadRango	1_(20-29)	1_(20-29), 6	
	2_(30-39)	2_(30-39), 24	
	3_(40-49)	3_(40-49), 63	
	4_(50-59)	4_(50-59), 140	
	5_(60-69)	5_(60-69), 194	
	6_(70-79)	6_(70-79), 171	
	7_(80+)	7_(80+), 21	
PacieSexo	F	F, 473	
	M	M, 217	
InterEnUCIC_G	N	N, 548	
	S	S, 142	
Atributos Cuantitativos			

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

REPORTE DE GENERACIÓN DE LA FUENTE TEMPORARIA DE DATOS				
Nombre	Min	Max	Tendencia Central	Dispersión
Potasio	2,4	7,3	Media = 4,1	Des. Est. = 0,54
GlobRojos	1,69	7,61	Media = 4,21	Des. Est. = 0,82
Hemoglobina	5	18,7	Media = 12,4	Des. Est. = 2,34
Glucemia	59	545	Media = 144	Des. Est. = 70,27
Hematocrito	16	55	Media = 37	Des. Est. = 6,60
CreatiSerica	0,4	8,45	Media = 1,21	Des. Est. = 0,87
ECO_FEY	5	86	Media = 52	Des. Est. = 14,30
ECO_IndDeMa	2	370	Media = 144 (F)	Des. Est. = 54,74
			Media = 119 (M)	Des. Est. = 41,67
ECO_Tisular	4	46	Media = 14	Des. Est. = 6,87
ECO_VolAurIzq	13	99	Media = 41	Des. Est. = 17,83

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.3.4.2. Adecuación de la Fuente Temporal de Datos (D.PD.AFT)

En esta actividad se adecuan los datos para la extracción de patrones de conocimiento. Se identifica y realiza actividades de conversión y ajuste de los registros.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (D.PD.AFT).

Información de Entrada

__Reporte de calidad de los datos
(D.ED.EvD.ReCD). Sección
4.3.2.3.1.

__Selección de Variables del Modelo
(D.Mo.CoM.SeVM). Sección
4.3.3.2.2.

_Fuente Temporal de Datos
(D.PD.CFT.FuTD). Sección 4.3.4.1.1.

_Reporte de Generación de la Fuente
Temporal de datos
(D.PD.CFT.RGFT). Sección
4.3.4.1.2.

Información de Salida

_ Reporte de Adecuación de la Fuente
Temporal de Datos
(D.PD.AFT.RAFT). Sección 4.3.4.2.1.

4.3.4.2.1. Reporte de Adecuación de la Fuente Temporal de Datos (D.PD.AFT.RAFT)

En esta etapa se realizan los ajustes a la fuente de información que integra todas las variables de interés para

el problema de negocio, conforme a las restricciones de los algoritmos a utilizar.

Los valores faltantes corresponden al caso en que el paciente no realizó el estudio. El manejo de los valores faltantes o perdidos, generalmente está orientado a implementar dos tipos de estrategias. La más sencilla consiste en filtrar los valores faltantes, ignorando por completo el ejemplo donde aparece al menos un valor faltante, sin embargo, esta estrategia claramente sesgaría los datos. Para el caso de estudio esta estrategia no es conveniente ya que los valores faltantes en el caso de los campos ECO_IndDeMa y ECO_VolAurIzq es alto, esto podría resultar una estrategia inapropiada. De este modo, una forma más robusta para tratar los valores faltantes, consiste en remplazarlos.

La media se utiliza en diversas situaciones, es fácil ejecutarla, es una de las opciones que ofrecen la mayoría del software estadístico de análisis de datos.

En el método de imputación con la media, se estima la media absoluta de los registros presentes en la base de

datos completa para la variable a imputar, por lo tanto, el valor resultante (media absoluta) sería el valor donante para los registros con datos faltantes para cada una de las variables que presenten al menos un registro ausente.

De acuerdo a lo expuesto anteriormente y a la sugerencia del médico, se propone reemplazar los valores con imputación por la media de la población de los valores del atributo que corresponde al valor faltante.

En el formalismo de la selección de variables del modelo, se identifica la necesidad de convertir a discretas diez variables. La discretización se realizará mediante una clasificación por categoría, en donde se tendrá en cuenta si los valores de las variables se encuentran fuera del rango normal de acuerdo a las características que presenta cada una.

En la Tabla 62 se ilustra el formalismo del caso de estudio correspondiente al Reporte de Adecuación de la Fuente Temporaria de Datos.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 62: D. Reporte de Adecuación de la Fuente Temporal de Datos.

REPORTE DE ADECUACIÓN DE LA FUENTE TEMPORARIA DE DATOS			
Responsable	Sofia Vallejos	Fecha	16/07/21
ID	D.PD.AFT.RAFT	Versión	1.0
Problema del Negocio	PrNe-01: Cuáles son las condiciones de los pacientes que requieren el grado de cuidados intensivos coronarios.		
Nombre	Acción	Efecto	Descripción
GlobRojos	Imputación de la media	15 registros afectados	Los valores null han sido sustituidos por las media de la variable.
Hemoglobina	Imputación de la media	12 registros afectados	Los valores null han sido sustituidos por las media de la variable.
Glucemia	Imputación de la media	61 registros afectados	Los valores null han sido sustituidos por las media de la variable.
Hematocrito	Imputación de la media	12 registros afectados	Los valores null han sido sustituidos por las media de la variable.
CreatiSerica	Imputación de la media	1 registros afectados	Los valores null han sido sustituidos por las media de la variable.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

REPORTE DE ADECUACIÓN DE LA FUENTE TEMPORARIA DE DATOS			
ECO_IndDeMa	Imputación de la media	263 registros afectados	Los valores null han sido sustituidos por las media de la variable.
ECO_Tisular	Imputación de la media	55 registros afectados	Los valores null han sido sustituidos por las media de la variable.
ECO_VolAurIzq	Imputación de la media	192 registros afectados	Los valores null han sido sustituidos por las media de la variable.
Potasio	Conversión a discreto	2 valores Identificados	Transformar la variable en una categoría. _Si Potasio es mayor o igual a 3,5 o Potasio es menor o igual 5 entonces <i>Potasio = 1_Normal</i> caso contrario <i>Potasio = 2_FueradeRango</i>
GlobRojos	Conversión a discreto	2 valores Identificados	Transformar la variable en una categoría. _Si GlobRojos es mayor o igual

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

REPORTE DE ADECUACIÓN DE LA FUENTE			
TEMPORARIA DE DATOS			
			<p>a 4 o GlobRojos es menor o igual 5,2 entonces <i>GlobRojos = 1_Normal</i> caso contrario GlobRojos = <i>2_FueradeRango</i></p>
Hemoglobina	Conversión a discreto	2 valores Identificados	<p>Transformar la variable en una categoría. _Si Hemoglobina es mayor o igual a 12 o Hemoglobina es menor o igual 16 entonces Hemoglobina = <i>1_Normal</i> caso contrario Hemoglobina = <i>2_FueradeRango</i></p>
Glucemia	Conversión a discreto	2 valores Identificados	<p>Transformar la variable en una categoría. _Si Glucemia es mayor o igual a 70 o Glucemia es menor o igual 100</p>

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

REPORTE DE ADECUACIÓN DE LA FUENTE TEMPORARIA DE DATOS			
			<p>entonces Glucemia = <i>1_Normal</i></p> <p>caso contrario Glucemia = <i>2_FueradeRango</i></p>
Hematocrito	Conversión a discreto	2 valores identificados	<p>Transformar la variable en una categoría. _Si Hematocrito es mayor o igual a 36 o Hematocrito es menor o igual 46</p> <p>entonces Hematocrito = <i>1_Normal</i></p> <p>caso contrario Hematocrito = <i>2_FueradeRango</i></p>
CreatiSerica	Conversión a discreto	2 valores identificados	<p>Transformar la variable en una categoría. _Si CreatiSerica es mayor o igual a 0,5 o CreatiSerica es menor o igual 0,9</p> <p>entonces CreatiSerica = <i>1_Normal</i></p>

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

REPORTE DE ADECUACIÓN DE LA FUENTE			
TEMPORARIA DE DATOS			
			caso contrario CreatiSerica = <i>2_FueradeRango</i>
ECO_FEY	Conversión a discreto	2 valores identificados	Transformar la variable en una categoría. _Si ECO_FEY es mayor o igual a 50 o ECO_FEY es menor o igual 70 entonces ECO_FEY = <i>1_Normal</i> caso contrario ECO_FEY = <i>2_FueradeRango</i>
ECO_IndDeMa	Conversión a discreto	2 valores identificados	Transformar la variable en una categoría. _Si (ECO_IndDeMa es mayor o igual a 95 y ECO_IndDeMa es igual a “F”) o (ECO_IndDeMa es mayor o igual a 115 y ECO_IndDeMa es igual a “M”) entonces

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

REPORTE DE ADECUACIÓN DE LA FUENTE TEMPORARIA DE DATOS			
			ECO_IndDeMa = 1_Normal caso contrario ECO_IndDeMa = 2_FueradeRango
ECO_Tisular	Conversión a discreto	2 valores identificados	Transformar la variable en una categoría. _Si ECO_Tisular es mayor o igual a 15 entonces ECO_Tisular = 1_Normal caso contrario ECO_Tisular = 2_FueradeRango
ECO_VolAurIzq	Conversión a discreto	2 valores identificados	Transformar la variable en una categoría. _Si ECO_VolAurIzq es mayor o igual a 34 entonces ECO_VolAurIzq = 1_Normal caso contrario ECO_VolAurIzq = 2_FueradeRango

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.3.5. Implementación (D.Im)

En esta fase se realizan las tareas necesarias para determinar el modelo y la extracción de los patrones de conocimiento de la fuente de información identificadas.

4.3.5.1. Selección del Modelo (D.Im.SeM)

En esta actividad se definen el criterio y la forma mediante la cual se determinarán los algoritmos o combinación de algoritmos con el fin de lograr adquirir con mayor exactitud los patrones ocultos en los datos. A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (D.Im.SeM).

Información de Entrada

_Selección de Algoritmos de Explotación de Información (D.Mo.CoM.SAEI). Sección 4.3.3.2.1.

_Reporte de Generación de la Fuente Temporal de datos (D.PD.CFT.RGFT). Sección 4.3.4.1.2.

_Criterios de Éxito del Problema de Negocio (D.EN.CPN.CEPN). Sección 4.3.1.2.2.

_Reporte de Evaluación de Herramientas (G.IP.EvS.EvHe). Sección 4.2.1.3.1.

Información de Salida

_Reporte de Estrategia de Parametrización del Modelo (D.Im.SeM.REPM). Sección 4.3.5.1.1.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.3.5.1.1. Reporte de Estrategia de Parametrización del Modelo (D.Im.SeM.REPM)

En esta etapa se describen las estrategias a utilizar para seleccionar el mejor modelo. Teniendo en cuenta el problema que se pretende resolver (descubrimiento de reglas de comportamiento), los algoritmos de clasificación a utilizar (C4.5 y CS-MC4) y la herramienta seleccionada (Tanagra), se definen los parámetros y rangos de valores, los cuales serán definidos mediante la estrategia de búsqueda aleatoria.

Tabla 63: D. Reporte de Estrategia de Parametrización del Modelo

REPORTE DE ESTRATEGIA DE PARAMETRIZACIÓN DEL MODELO						
Responsable		Sofia Vallejos		Fecha		21/07/21
ID		D.Im.SeM.REPM		Versión		1.0
Problema del Negocio		PrNe-01: Cuáles son las condiciones de los pacientes que requieren el grado de cuidados intensivos coronarios.				
ID	Algoritmo E.I.	Estrategia de Configuración	Criterio	Rango Inferior	Rango Superior	Comentarios
REPM - 01	C4.5	Radom-Search	Tasa de Error	Min size of leaves: 5	Min size of leaves: 10	Confidence Level: 0,25
REPM - 02	CS-MC4	Radom-Search	Tasa de Error	Min size of leaves: 5	Min size of leaves: 10	Lamba for laplacian: 3

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

En la Tabla 63 se ilustra el formalismo del caso de estudio correspondiente al Reporte de Estrategia de Parametrización del Modelo.

4.3.5.2. Explotación de la Información (D.Im.ExI)

En esta actividad se aplican los algoritmos de explotación de información sobre las fuentes de información con el fin de extraer los patrones de conocimientos ocultos, registrando los resultados obtenidos de la aplicación. A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (D.Im.ExI).

Información de Entrada

_Criterios de Éxito del Problema de Negocio (D.EN.CPN.CEPN). Sección 4.3.1.2.2.

_Selección de Algoritmos de Explotación de Información (D.Mo.CoM.SAEI) . Sección 4.3.3.2.1.

_Selección de Variables del Modelo (D.Mo.CoM.SeVM). Sección 4.3.3.2.2.

_Estrategias de Evaluación de Modelos (D.Mo.CoM.EsEM). Sección 4.3.3.2.3.

Información de Salida

_Reporte de Implementación del Modelo (D.Im.ExI.ReIM). Sección 4.3.5.2.1.

_Patrones de Conocimiento (D.Im.ExI.PaCo). Sección 4.3.5.2.2

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

_Reporte de Estrategia de Parametrización del Modelo (D.Im.SeM.REPM). Sección 4.3.5.1.1

_Fuente Temporal de Datos (D.PD.CFT.FuTD). Sección 4.3.4.1.1.

_Reporte de Generación de la Fuente Temporal de datos (D.PD.CFT.RGFT). Sección 4.3.4.1.2.

4.3.5.2.1. Reporte de Implementación del Modelo (D.Im.ExI.ReIM)

En esta etapa se describe la obtención de los parámetros de la mejor configuración para los algoritmos a analizar.

Tabla 64: D. Reporte de Implementación del Modelo

REPORTE DE IMPLEMENTACIÓN DEL MODELO			
Responsable	Sofia Vallejos	Fecha	24/07/21
ID	D.Im.SeM.REPM	Versión	1.0
Problema del Negocio	PrNe-01: Cuáles son las condiciones de los pacientes que requieren el grado de cuidados intensivos coronarios.		
Algoritmo E.I.	Estrategia	Configuración	Descripción
C4.5	REPM – 01 Radom-Search	Min size of leaves: 5 Confidence Level: 0,25	Número de nodos: 25 Número de Hojas: 16 Tasa de error (promedio): 0.1870
CS-MC4	REPM – 02 Radom-Search	Min size of leaves: 5 Lamba for laplacian: 3	Número de nodos: 23 Número de Hojas: 12 Tasa de error (promedio): 0.1841

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

En la Tabla 64 se ilustra el formalismo del caso de estudio correspondiente al Reporte de Implementación del Modelo.

4.3.5.2.2. Patrones de Conocimiento (D.Im.ExI.PaCo)

En esta fase se describen los patrones obtenidos de acuerdo al modelo que obtuvo mejores resultados. De acuerdo a la información obtenida en *Reporte de Implementación del Modelo* el modelo que obtuvo mejores resultados es el CS-MC4.

En las Fig. 13 y Fig. 14 se presentan las reglas de comportamiento obtenidas en la herramienta Tanagra a partir del caso de estudio. Dentro de este marco, en la Tabla 65 se presentan la interpretación de las reglas de comportamiento generadas.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Results						
Classifier performances			Error rate			
			0,1841			
Values prediction			Confusion matrix			
Value	Recall	1-Precision	N	S	Sum	
N	0,9854	0,1806	N	540	8	548
S	0,1620	0,2581	S	119	23	142
			Sum	659	31	690

Classifier characteristics

Data description

Target attribute InterEnUCIC_G (2 values)

descriptors 12

Tree description

Number of nodes 23

Number of leaves 12

Fig. 13: Patrones de Conocimiento (Tanagra).A.

Decision tree

- Potasio in [1_Normal]
 - GlobRojos in [2_FueraRango] then InterEnUCIC_G = N (87,02 % of 285 examples)
 - GlobRojos in [1_Normal]
 - ECO_Tisular in [1_Normal]
 - CreatSerica in [2_FueraRango]
 - ECO_FEY in [1_Normal] then InterEnUCIC_G = N (66,67 % of 18 examples)
 - ECO_FEY in [2_FueraRango]
 - Glucemia in [2_FueraRango]
 - Hematocrito in [2_FueraRango] then InterEnUCIC_G = S (57,14 % of 7 examples)
 - Hematocrito in [1_Normal] then InterEnUCIC_G = N (59,09 % of 22 examples)
 - Glucemia in [1_Normal] then InterEnUCIC_G = S (75,00 % of 8 examples)
 - CreatSerica in [1_Normal] then InterEnUCIC_G = N (73,68 % of 38 examples)
 - ECO_Tisular in [2_FueraRango] then InterEnUCIC_G = N (82,35 % of 238 examples)
- Potasio in [2_FueraRango]
 - PacieSexo_D in [F] then InterEnUCIC_G = N (76,74 % of 43 examples)
 - PacieSexo_D in [M]
 - Glucemia in [2_FueraRango]
 - CreatSerica in [2_FueraRango]
 - Hematocrito in [2_FueraRango] then InterEnUCIC_G = N (60,00 % of 10 examples)
 - Hematocrito in [1_Normal] then InterEnUCIC_G = S (70,00 % of 10 examples)
 - CreatSerica in [1_Normal] then InterEnUCIC_G = S (100,00 % of 6 examples)
 - Glucemia in [1_Normal] then InterEnUCIC_G = N (80,00 % of 5 examples)

Fig. 14: Patrones de Conocimiento. B.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 65: D. Patrones de Conocimiento

<p>SI Potasio es Normal Y los Glóbulos Rojos está Fuera del Rango ENTONCES No necesita estar en UCIC (87,02 % de 285)</p>	<p>SI Potasio esta Fuera del Rango Y el Sexo es femenino ENTONCES No necesita estar en UCIC (76,74 % de 43)</p>
<p>SI Potasio es Normal Y los Glóbulos Rojos es Normal Y ECO_Tisular es Normal Y Creatinina Sérica está Fuera del Rango Y el ECO FEY es Normal. ENTONCES No necesita estar en UCIC (66,67 de 18)</p>	<p>SI Potasio esta Fuera del Rango Y el Sexo es masculino Y la Glucemia esta Fuera del Rango Y la CreatiSerica está Fuera del Rango Y el Hematocrito está Fuera del Rango ENTONCES No necesita estar en UCIC (60,00 % de 10)</p>
<p>SI Potasio es Normal Y los Glóbulos Rojos es Normal Y ECO_Tisular es Normal Y Creatinina Sérica está Fuera del Rango Y el ECO FEY está Fuera del Rango Y Glucemia está Fuera del Rango Y el Hematocrito está Fuera del Rango ENTONCES Necesita estar en UCIC (57,14 % de 7)</p>	<p>SI Potasio esta Fuera del Rango Y el Sexo es masculino Y la Glucemia esta Fuera del Rango Y la CreatiSerica está Fuera del Rango Y el Hematocrito es Normal ENTONCES Necesita estar en UCIC (70,00 % de 10)</p>
<p>SI Potasio es Normal Y los Glóbulos Rojos es Normal Y ECO_Tisular es Normal</p>	<p>SI Potasio esta Fuera del Rango Y el Sexo es masculino Y la Glucemia esta Fuera del Rango Y la CreatiSerica es Normal ENTONCES Necesita estar en UCIC (100,00 % de 6)</p>
	<p>SI Potasio esta Fuera del Rango</p>

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

<p> Y Creatinina Sérica está Fuera del Rango Y el ECO FEY está Fuera del Rango Y Glucemia está Fuera del Rango Y el Hematocrito es Normal ENTONCES No necesita estar en UCIC (59,09 % de 22) </p> <p> SI Potasio es Normal Y los Glóbulos Rojos es Normal Y ECO_Tisular es Normal Y Creatinina Sérica está Fuera del Rango Y el ECO FEY está Fuera del Rango Y Glucemia es Normal ENTONCES Necesita estar en UCIC (75,09 % de 8) </p> <p> SI Potasio es Normal Y los Glóbulos Rojos es Normal Y ECO_Tisular es Normal Y Creatinina Sérica es Normal ENTONCES No necesita estar en UCIC (73,68 % de 38) </p> <p> SI Potasio es Normal Y los Glóbulos Rojos es Normal Y ECO_Tisular es Normal ENTONCES No necesita estar en UCIC (82,35 % de 238) </p>	<p> Y el Sexo es masculino Y la Glucemia es Normal ENTONCES No necesita estar en UCIC (80,00 % de 5) </p>
--	--

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

4.3.6. Evaluación y Presentación (D.EP)

En esta fase se realizan las validaciones finales del producto y la presentación de los resultados obtenidos, ajustando la misma para la comprensión del procedimiento y las conclusiones procedentes para la contribución de la toma de decisión.

4.3.6.1. Evaluación de los Resultados (D.EP.EvR)

En esta actividad se realiza una evaluación de la autenticidad de los patrones de conocimiento obtenidos para el dominio de negocio y en particular para la problemática de negocio planteado. Además, se debe corroborar la satisfacción de la problemática o identificación de nuevas necesidades a resolver.

A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (D.EP.EvR).

Información de Entrada

_Objetivo del Proyecto
(D.EN.AnN.ObPr). Sección
4.3.1.1.3

Información de Salida

_Reporte de Evaluación de los
Resultados (D.EP.EvR.ReER).
Sección 4.3.6.1.1.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

_Criterios de éxito del proyecto
(D.EN.AnN.CrEP). Sección
4.3.1.1.4.

_Problema del Negocio
(D.EN.CPN.PrNe). Sección
4.3.1.2.1.

_Criterios de Éxito del Problema de
Negocio (D.EN.CPN.CEPN).
Sección 4.3.1.2.2.

_Reporte de Implementación del
Modelo (D.Im.ExI.ReIM). Sección
4.3.5.2.1.

_Patrones de Conocimiento
(D.Im.ExI.PaCo). Sección 4.3.5.2.2.

4.3.6.1.1. Reporte de Evaluación de los Resultados (D.EP.EvR.ReER)

En esta etapa se validan los resultados, con respecto a su representatividad, validez y novedad para los expertos del negocio.

En función de lo planteado en En la *Tabla 64* se ilustra el formalismo del caso de estudio correspondiente al Reporte de Implementación del Modelo.

Patrones de Conocimiento (D.Im.ExI.PaCo), y la evaluación realizada con el experto (Médico de UCIC), se determinó la veracidad de los resultados. En resumida cuenta se estableció que los mismos son satisfactorio y

Implementación de IEL para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

cumplen con los *Objetivos del Proyecto*, el *Criterios de Éxito del Proyecto*, el *Problema del Negocio* y el *Criterio de Éxito del Problema de Negocio*.

Algunas de sus manifestaciones:

“Una de las cosas que van junto con la Creatinina es el Potasio y el potasio por fuera de un determinado valor induce arritmia cardíaca, el paciente puede entrar en paro, etc., los dos aportes están muy correlacionados y a su vez aportan información independiente entre ellos hay una correlación muy fuerte (potasio y creatinina). Es increíble lo que puede hacer el algoritmo, algo que no tiene la posibilidad de pensar medicamento, que cosa increíble”

“La posibilidad de implementar esto en un sistema informático y que venga un paciente nuevo se pueda contrastar, me va a terminar diciendo este nuevo paciente es probable que le vaya de esta manera porque con todos los datos que ya se analizó en este algoritmo sabemos que la probabilidades son estas... eso es lo que aporta fundamentalmente, en general nosotros

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

tendemos a tratar de facilitar para el análisis médico... cuando yo hice la residencia lo que teníamos en el examen físico, es el estetoscopio y hablar con el paciente, el electrocardiograma y el análisis de laboratorio y se terminó. Ahora los chicos que están haciendo la residencia ocupan menos el estetoscopio, cada vez revisan menos al paciente, cada vez hablan menos con el paciente y lo inmediatamente que hacen antes que suban, es entrar al sistema informático, miran la radiografía, el electro, los análisis de laboratorio, los antecedentes, la fracción de eyección y potasio, entonces recién ahí toman una decisión. Lo que ayudaría esto (Implementar EI) es que todo ese trabajo se automatiza y poder dedicar ese tiempo a otra cosa. *Este universo de datos junto con el algoritmo permita analizar rápidamente, esto es lo que hay que hacer, es un paso adelante”*

En la Tabla 66 se ilustra el formalismo del caso de estudio correspondiente al Reporte de Evaluación de los Resultados.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Tabla 66: D. Reporte de Evaluación de los Resultados

REPORTE DE EVALUACIÓN DE LOS RESULTADOS			
Responsable	Sofia Vallejos	Fecha	30/07/21
ID	D.EP.EvR.ReER	Versión	1.0
Problema de Negocio	Criterio de Éxito	Resultado	Descripción
Cuáles son las condiciones de los pacientes que requieren el grado de cuidados intensivos coronarios.	Identificar aspectos que permitan comprender el comportamiento de los pacientes que se encuentran en la unidad de cuidados intensivos coronarios.	Satisfactorio	Las reglas identificadas permiten comprender el comportamiento de los pacientes que requieren un grado de cuidado intensivo.

4.3.6.2. Presentación de los Resultados (D.EP.PrR)

En esta actividad se realizan las tareas finales del proyecto, con respecto a la documentación los resultados obtenidos y la presentación de los mismos a los interesados. Además, se debe tener en cuenta la correcta comunicación de los patrones obtenidos y el conocimiento extraído para contribuir como soporte al proceso de toma de decisiones del cliente. A continuación, se detallan los elementos de entrada necesarios y los de salida generados de la actividad (D.EP.PrR).

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Información de Entrada

_Fuentes de Información del Cliente (D.EN.AnN.FuIC). Sección 4.3.1.1.1.

_Objetivo del Proyecto (D.EN.AnN.ObPr). Sección 4.3.1.1.3

_Criterios de éxito del proyecto (D.EN.AnN.CrEP). Sección 4.3.1.1.4.

_Suposiciones del proyecto (D.EN.AnN.SuPr). Sección 4.3.1.1.6.

_Restricciones del proyecto (D.EN.AnN.RePr). Sección 4.3.1.1.7.

_Problema del Negocio (D.EN.CPN.PrNe). Sección 4.3.1.2.1.

_Patrones de Conocimiento (D.Im.ExI.PaCo). Sección 4.3.5.2.2.

_Reporte de Generación de la Fuente Temporal de datos (D.PD.CFT.RGFT). Sección 4.3.4.1.2.

_Reporte de calidad de los datos (D.ED.EvD.ReCD). Sección 4.3.2.3.1.

_Reporte de Evaluación de los Resultados (D.EP.EvR.ReER). Sección 4.3.6.1.1.

Información de Salida

_Reporte del Proyecto (D.EP.PrR.RepP). Sección 4.3.6.2.1.

4.3.6.2.1. Reporte del Proyecto (D.EP.PrR.RepP)

En esta etapa se desarrolla un reporte direccionado a los expertos del dominio del negocio con el objetivo de

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

contribuir como soporte al proceso de toma de decisiones. En la Tabla 67 se ilustra el formalismo del caso de estudio correspondiente al Reporte del Proyecto.

Tabla 67: D. Reporte del Proyecto.

REPORTE DEL PROYECTO			
Responsable	Sofia Vallejos	Fecha	4/08/21
ID	D.EP.PrR.RepP	Versión	1.0
DESCRIPCIÓN DEL PROBLEMA	La unidad de cuidados intensivos coronarios recibe pacientes desde otros servicios, tales como emergencias, internación, cirugía. Cuando se solicita a UCIC que reciba un paciente, un médico evalúa la ‘gravedad’ del mismo y determina cómo lo reciben, qué conducta tomar según su estado, por ellos los pacientes de mayor vulnerabilidad sanitaria y con riesgo de vida son admitidos en él. Para determinar el estado de un paciente, un médico debe invertir un tiempo importante. Más aún, si no ha tenido contacto con él. Debe analizar la historia clínica electrónica de los que fueron pacientes ya registrados, con toda la información de los eventos sucedidos en la institución.		
DESCRIPCIÓN DE LOS DATOS	El análisis fue realizado sobre 689 pacientes correspondiente al año 2019. Para ello se tuvieron en cuenta las siguientes variables: Edad PacieSexo InterEnUCIC Potasio Glóbulos Rojos Hemoglobina Glucemia Hematocrito		

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

	<p>Creatinina Sérica ECO FEY ECO de Índice de Masa ECO Tisular ECO del Volumen de la Aurícula Izquierda</p>
<p>RESULTADOS DE EXPLOTACIÓN DE INFORMACIÓN</p>	<p>SI Potasio es Normal Y los Glóbulos Rojos está Fuera del Rango ENTONCES No necesita estar en UCIC (87,02 % de 285)</p> <p>SI Potasio es Normal Y los Glóbulos Rojos es Normal Y ECO_Tisular es Normal Y Creatinina Sérica está Fuera del Rango Y el ECO FEY es Normal. ENTONCES No necesita estar en UCIC (66,67 de 18).</p> <p>SI Potasio es Normal Y los Glóbulos Rojos es Normal Y ECO_Tisular es Normal Y Creatinina Sérica está Fuera del Rango Y el ECO FEY está Fuera del Rango Y Glucemia está Fuera del Rango Y el Hematocrito está Fuera del Rango</p> <p>SI Potasio esta Fuera del Rango Y el Sexo es femenino ENTONCES No necesita estar en UCIC (76,74 % de 43)</p> <p>SI Potasio esta Fuera del Rango Y el Sexo es masculino Y la Glucemia esta Fuera del Rango Y la CreatiSerica está Fuera del Rango Y el Hematocrito está Fuera del Rango ENTONCES No necesita estar en UCIC (60,00 % de 10).</p> <p>SI Potasio esta Fuera del Rango Y el Sexo es masculino Y la Glucemia esta Fuera del Rango Y la CreatiSerica está Fuera del Rango Y el Hematocrito es Normal ENTONCES Necesita estar en UCIC (70,00 % de 10)</p> <p>SI Potasio esta Fuera del Rango Y el Sexo es masculino Y la Glucemia esta Fuera del Rango Y la CreatiSerica es Normal</p>

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

	<p>ENTONCES Necesita estar en UCIC (57,14 % de 7)</p> <p>SI Potasio es Normal</p> <p>Y los Glóbulos Rojos es Normal</p> <p>Y ECO_Tisular es Normal</p> <p>Y Creatinina Sérica está Fuera del Rango</p> <p>Y el ECO FEY está Fuera del Rango</p> <p>Y Glucemia está Fuera del Rango</p> <p>Y el Hematocrito es Normal</p> <p>ENTONCES No necesita estar en UCIC (59,09 % de 22)</p> <p>SI Potasio es Normal</p> <p>Y los Glóbulos Rojos es Normal</p> <p>Y ECO_Tisular es Normal</p> <p>Y Creatinina Sérica está Fuera del Rango</p> <p>Y el ECO FEY está Fuera del Rango</p> <p>Y Glucemia es Normal</p> <p>ENTONCES Necesita estar en UCIC (75,09 % de 8)</p> <p>SI Potasio es Normal</p> <p>Y los Glóbulos Rojos es Normal</p> <p>Y ECO_Tisular es Normal</p> <p>Y Creatinina Sérica es Normal</p>	<p>ENTONCES Necesita estar en UCIC (100,00 % de 6)</p> <p>SI Potasio esta Fuera del Rango</p> <p>Y el Sexo es masculino</p> <p>Y la Glucemia es Normal</p> <p>ENTONCES No necesita estar en UCIC (80,00 % de 5)</p>
--	--	--

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

	<p>ENTONCES No necesita estar en UCIC (73,68 % de 38)</p> <p>SI Potasio es Normal</p> <p>Y los Glóbulos Rojos es Normal</p> <p>Y ECO_Tisular es Normal</p> <p>ENTONCES No necesita estar en UCIC (82,35 % de 238).</p>
<p>EVALUACIÓN DE LOS RESULTADOS</p>	<p>Tras el análisis, podemos deducir que las reglas de comportamiento identificadas colaboran con la toma de decisiones de los médicos de la UCIC. En efecto, permite determinar el estado de gravedad de un paciente y conducta tomar según su estado, reduciendo sustancialmente el tiempo que un médico debe invertir para identificarlo.</p>
<p>DIFICULTADES Y RECOMENDACIONES</p>	<p>Se señala como posibles trabajos a realizar, el análisis de los casos de los óbitos y el desarrollo de una interface de visualización de los resultados. Además, se sugiere ampliar el estudio realizado incorporando datos de los pacientes de otros años.</p>

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Capítulo 5

Conclusión y Futuras Líneas de Investigación

5. Conclusión y Futuras Líneas de Investigación

En este capítulo se presentan las principales conclusiones obtenidas tras la realización del proyecto. Además, se aportan algunas posibles líneas de investigación futuras que pueden enriquecer el modelo desarrollado.

5.1. Conclusiones

En función de lo planteado en el capítulo anterior y haciendo hincapié en los resultados obtenidos podemos concluir que la implementación del Modelo de Procesos para una Ingeniería de Explotación de la Información en una unidad de cuidado intensivos coronarios provee un enfoque acertado para resolver problemas de este tipo.

En el caso específico de las tareas de diagnóstico/pronóstico médico de las unidades de cuidado intensivos coronarios, la implementación de MoProPEI permitió identificar reglas de comportamiento, convirtiéndose en una alternativa válida y posible para la solución de problemas en el gran campo de la Informática Médica.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

Debe señalarse que no es suficiente con obtener patrones de comportamiento, uno de los componentes más importante es el de validar los resultados con los expertos y alcanzar los objetivos con el menor costo y plazos posibles. Gracias a la incorporación de modelo de procesos al problema planteado permitió una sistematización al proyecto, enfocado en la planificación, adaptación, gestión y control.

La explotación de la información ha tenido muchos avances en relación a la aplicación y desarrollo de modelos de procesos en el área de la salud. Sin embargo, no se han encontrado trabajos que apliquen el MoProPEI particularmente en la UCIC con el objetivo de determinar el grado de gravedad de un paciente. En vista de que la implementación del modelo fue exitosa, esta propuesta podría aplicarse a otras instituciones con características similares.

5.2. Futuras Líneas de Investigación

Una vez implementada la IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios, se han planteado nuevas líneas de investigación con las que se puede mejorar y ampliar el trabajo de investigación. A continuación, se enumeran algunas de ellas:

1. En este proyecto no se abarcó la totalidad de los datos contenidos en las HCE, sino solo aquellos que contienen incidencia en la UCIC. Podrían surgir nuevas investigaciones, teniendo en cuenta la caracterización de los otros tipos de internaciones o la caracterización de los óbitos.
2. Desarrollar una interface de visualización de datos en donde los resultados se representen de forma gráfica con el objetivo de proporcionar una manera accesible a los médicos de ver y comprender tendencias, valores atípicos y patrones en los datos.

Implementación de IEI para la identificación de patrones que contribuyan a la toma de decisiones en una unidad de cuidados intensivos coronarios.

3. Desarrollar una herramienta software de apoyo al modelo de proceso propuesto con el objetivo de facilitar la comunicación entre los recursos humanos involucrados en el proyecto para el seguimiento del desarrollo de las fases, actividades y etapas. De este modo es conveniente automatizar los formalismos, en donde se registre el estado, distinga las diferentes versiones, es decir diferentes vistas, quién, cuándo, y cómo se modificó.
4. Implementar otro modelo de proceso con el objetivo de analizar, verificar y evaluar cuál de ellos se ajusta mejor a estos tipos de problemas de dominio.

6. Referencias

- [1] S. Martins, P. Pesado y R. García-Martínez. “Propuesta de Modelo de Procesos para una Ingeniería de Explotación de Información: MoProPEI”. *Revista Latinoamericana de Ingeniería de Software*, 2014, vol. 2, no. 5, pp. 313-332.
- [2] M. O. Parra. “Historia y evolución de la medicina crítica: de los cuidados intensivos a la terapia intensiva y cuidados críticos”, 2017, doi.org/10.1016/j.acci.2017.08.006.
- [3] Society of critical care medicine. Disponible en: <http://www.myicucare.org/Pages/default.aspx>. [Accedido: 29, Jun, 2020].
- [4] B. A. Lara et al. “Medicina de urgencia y unidades de cuidados intensivos: Una alianza necesaria en busca de la mejoría de la atención de pacientes críticos”. [En línea]. Disponible en: <https://dx.doi.org/10.4067/S0034-98872016000700014>. [Accedido: 29, Jun, 2020].
- [5] J. Pérez Arrarte, “De las unidades de cuidado coronario a las unidades de cuidado intensivo cardiovascular: una evolución inevitable”, *Revista Uruguay de Cardiología*, Uruguay, Montevideo, 2017, vol.32, no.3, pp. 229-233, <http://dx.doi.org/10.29277/ruc/32.3.2>.
- [6] A. Juan, E. Benjamio, *et al.*, “Impacto de la implementación de medidas de gestión hospitalaria para aumentar la eficiencia en la gestión de camas y disminuir la saturación del servicio de urgencias”. *Emergencias: Revista de la Sociedad Española de*

- Medicina de Urgencias y Emergencias, 2010, vol. 22, no. 4, pp. 249-253.
- [7] R. Sabartés Fortuny, “Historia clínica electrónica en un departamento de obstetricia, ginecología y reproducción: desarrollo e implementación”, tesis doctoral, Universidad Autónoma de Barcelona, 2013.
- [8] S. Borah, V. E. Balas, and Z. Polkowski, “Advances in data science and management”, Proceedings of ICDSM, 2019, [Online]. Available: <https://link.springer.com/book/10.1007%2F978-981-15-0978-0>. [Accessed: Dec 1, 2021]
- [9] P. Amirian, F. Van Loggerenberg and T. Lang, “Data Science and Analytics”, Big Data in Healthcare, SpringerBriefs in Pharmaceutical Science & Drug Development. Springer, Cham, pp. 15-37, 2017, https://doi.org/10.1007/978-3-319-62990-2_2
- [10] M. Li, et al., “Medical informatics: an essential tool for health sciences research in acute care” Bosnian journal of basic medical sciences, vol. 9, no Suppl 1, 2009, [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5655170/>. [Accessed: Oct 1, 2021]
- [11] N. Hasan and Y. Bao, “Understanding current states of machine learning approaches in medical informatics: a systematic literature review”, Health and Technology, vol. 11, pp. 471–482, 2021, <https://doi.org/10.1007/s12553-021-00538-6>.

- [12] P. Amirian, F. Van Loggerenberg and T. Lang, “Big Data and Big Data Technologies”, *Big Data in Healthcare, SpringerBriefs in Pharmaceutical Science & Drug Development*. Springer, Cham, pp. 39-58, 2017, https://doi.org/10.1007/978-3-319-62990-2_3.
- [13] M. Spruit and M. Lytras, “Applied data science in patient-centric healthcare: Adaptive analytic systems for empowering physicians and patients”, *Telematics and Informatics*, vol 35, pp. 643-653, 2018, <https://doi.org/10.1016/j.tele.2018.04.002>.
- [14] L. Ehwerhemuepha, et al., “HealtheDataLab – a cloud computing solution for data science and advanced analytics in healthcare with application to predicting multi-center pediatric readmissions”, *BMC Medical Informatics and Decision Making*, vol. 20, no. 115, 2020, <https://doi.org/10.1186/s12911-020-01153-7>.
- [15] Z. Abedjan et al., “Data Science in Healthcare: Benefits, Challenges and Opportunities”, *Data Science for Healthcare*. Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-05249-2_1.
- [16] C. Auffray, et al., “Making sense of big data in health research: Towards an EU action plan”, *Genome Medicine*, vol. 8, no. 71, 2016, <https://doi.org/10.1186/s13073-016-0323-y>.
- [17] S. Martins, P. Pesado y R. Garcia-Martinez, “Propuesta de proceso de ingeniería de explotación de información centrado en control y gestión del proyecto” 2014. [En línea]. Disponible en:

- <https://digital.cic.gba.gob.ar/handle/11746/2065>.
[Último acceso: 04 2020].
- [18] P. Britos y R. García Martínez, “Propuesta de Procesos de Explotación de Información”, SEDICI Repositorio institucional de la UNLP, octubre 2009. [En línea]. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/21206>. [Último acceso: 06 2020].
- [19] B., Curtis, M. Kellner y J. Over, “Process Modelling” Communications of the ACM, 1992, vol. 35, no. 9, pp. 75-90. doi.org/10.1145/130994.130998.
- [20] F. Pollo-Cattaneo, H. Amatriain, D. Rodríguez, P. Pytel, E. Ciccolella, C. Vegega, M. Dearriba, M. Aubert, M. F. Bose, M. L. Giordano, P. Britos y R. Garcia Martínez, “Ingeniería de proyectos de explotación de la información”. 2010.
- [21] R. S. Michalski, “A Theory and Methodology of Inductive Learning”, 1983, pp. 83-134, doi.org/10.1016/B978-0-08-051054-5.50008-X.
- [22] N. Mohapatra, K. Shreya, and A. Chinmay, “Optimization of the random forest algorithm”, Advances in data science and management Springer, Singapore, pp. 201-208, 2020, https://doi.org/10.1007/978-981-15-0978-0_19.
- [23] V. Kumar, et al., “Prediction of Malignant and Benign Breast Cancer: A Data Mining Approach in Healthcare Applications”, Advances in Data Science and Management, Lecture Notes on Data Engineering and Communications Technologies,

- Springer, Singapore, vol. 37, 2020, https://doi.org/10.1007/978-981-15-0978-0_43.
- [24] S.Maji, S. Arora, “ Decision Tree Algorithms for Prediction of Heart Disease”, Information and Communication Technology for Competitive Strategies, Lecture Notes in Networks and Systems, Springer, Singapore, vol 40, pp. 447-454, 2019, https://doi.org/10.1007/978-981-13-0586-3_45.
- [25] S. Negash y P. Gray, “Business Intelligence”. En Handbook on Decision Support Systems 2, ed. F. Burstein y C. Holsapple Heidelberg, Springer, pp. 175-193, 2008.
- [26] A., Abran, J. W. Moore, P. Bourque, R. Dupuis, L. Tripp, “Guide to the Software Engineering Body of Knowledge” (2004 version). IEEE Computer Society Press. ANSI/IEEE, 2007, Draft IEEE Standard for software and system test documentation. ANSI/IEEE Std P829-2007, 2004.
- [27] R. García Martínez, E. Diez, R. García, S. Martins y E. Baldizzoni, “Modelos de proceso para ingeniería de explotación de información para pymes: abordaje ágil y abordaje robusto”, En XVII Workshop de Investigadores en Ciencias de la Computación, Argentina, Salta, 2015.
- [28] R. García-Martínez, P. Britos, P. Pesado, R. Bertone, F. Pollo-Cattaneo, D. Rodríguez, P. Pytel y J. Vanrell, “Towards an Information Mining Engineering. En Software Engineering, Methods, Modeling and Teaching”, Sello Editorial Universidad de Medellín, 2011, pp. 83-99.

- [29] J. Vanrell, R. Bertone y R. García-Martínez, “Modelo de Proceso de Operación para Proyectos de Explotación de Información”, *Anales del XVI Congreso Argentino de Ciencias de la Computación*, 2010, pp. 674-682.
- [30] J. Vanrell, R. Bertone y R. García-Martínez, “Un Modelo de Procesos para Proyectos de Explotación de Información”, *Proceedings Latin American Congress on Requirements Engineering and Software Testing*, 2012, pp. 46-52.
- [31] J. Langseth y N. Vivatrat, “Why Proactive Business Intelligence is a Hallmark of the Real- Time Enterprise: Outward Bound”, *Intelligent Enterprise*, 2003, vol. 5, no. 18, pp. 34-41.
- [32] J. Srivastava, R. Cooley, M. Deshpande y P. Tan, “Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data”, *SIGKDD Explorations*, 2000, vol. 1, no. 2, pp. 12-23.
- [33] A. Abraham, “Business Intelligence from Web Usage Mining”, *Journal of Information & Knowledge Management*, 2003, vol. 2, no. 4, pp. 375-390.
- [34] R. Cooley, “The Use of Web Structure and Content to Identify Subjectively Interesting Web Usage Patterns”, *ACM Transactions on Internet Technology*, 2003, vol. 3, no. 2, pp. 93-116.
- [35] S. Martins, “Modelo de proceso para proyectos de explotación de información”, tesis doctoral, Facultad de Informática Universidad Nacional de la Plata, 2020.

- [36] P. Britos, O. Dieste y R. García-Martínez. “Requirements Elicitation in Data Mining for Business Intelligence Projects”, In *Advances in Information Systems Research, Education and Practice*. IFIP WCC TC8 2008. IFIP – The International Federation for Information Processing, Springer, Boston, MA., 2008, vol. 274. doi.org/10.1007/978-0-387-09682-7-9_12
- [37] P. Pytel, A. Hossian, P. Britos, R. García-Martínez, “Feasibility and Effort Estimation Models for Medium and Small Size Information Mining Projects”. *Information Systems Journal*, Elsevier, 2015, vol. 47, pp. 01-14 <https://doi.org/10.1016/j.is.2014.06.004>.
- [38] P. Pytel, “Viabilidad y Estimación de proyectos de explotación de información, tesis doctoral. Universidad Nacional de La Plata, 2014. <https://doi.org/10.35537/10915/43572>
- [39] P. Pytel, P. Britos, R. García Martínez, “Propuesta de un Modelo para Evaluar la Viabilidad de Proyectos de Explotación de Información”, En XVIII Congreso Argentino de Ciencias de la Computación, 2012.
- [40] P. Britos, M. Abasolo, R. García-Martínez y F. Perales, “Identification of MPEG- 4 Patterns in Human Faces Using Data Mining Techniques”, *Proceedings 13th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*, 2005, pp. 9-10. [En línea]. Disponible en:

- http://wscg.zcu.cz/WSCG2005/Papers_2005/Poster/!WSCG2005_Poster_Proceedings_Final.pdf
- [41] M. Cogliati, P. Britos y R. García-Martínez, “Patterns in Temporal Series of Meteorological Variables Using SOM & TDIDT”. In: Bramer M. (eds) Artificial Intelligence in Theory and Practice. IFIP AI 2006. IFIP International Federation for Information Processing, vol 217. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-34747-9_32
- [42] A. Naik, y L. Samant. “Correlation review of classification algorithm using data mining tool: WEKA, Rapidminer, Tanagra, Orange and Knime”, Procedia Computer Science, 2016, vol. 85, pp. 662-668. doi.org/10.1016/j.procs.2016.05.251
- [43] A.K. Sharma, A. Kumar, S. Saxenay, M. Beniwal, “Evaluating WEKA over the Open Source Web Data Mining Tools”, International Journal of Engineering Research and Technology, 2015, vol. 8, no.1, pp.128-132.
- [44] Y. Ramamohan, K. Vasantharao, C. K. Chakravarti, y A. S. K. Ratnam, “A study of data mining tools in knowledge discovery process”, International Journal of Soft Computing and Engineering (IJSCE), 2012, vol. 2, no 3, pp. 191-194.
- [45] F. Pollo-Cattaneo, “Modelo de Proceso para Elicitación de Requerimientos en Proyectos de Explotación de Información”, tesis doctoral, Editorial de la Universidad Nacional de La Plata (EDULP), 2020.

- [46] J. Gallardo, “Metodología para la Definición de Requisitos en Proyectos de Data Mining (ER-DM)”, tesis Doctoral. Facultad de Informática (UPM), Madrid, España, 2009.
- [47] P. Britos, H. Merlino, E. Fernández, M. A. Ochoa, E. Diez, Y R. García-Martínez, “Tool Selection Methodology in Data Mining”, In JIISIC, 2006, pp. 85-90.
- [48] M. Hofmann, “The development of a generic data mining life cycle (DMLC)”, dissertation, School of Computing, Dublin Institute of Technology, 2003.
- [49] P. Chapman, J. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Shearer y R. Wirth, “CRISP-DM 1.0”, Step-by-step data mining guide, 2000.
- [50] D. Rodríguez, M. Pollo Cattaneo, P. Britos, R. García Martínez, “Estimación Empírica de Carga de Trabajo en Proyectos de Explotación de Información. In XVI Congreso Argentino de Ciencias de la Computación, 2010.
- [51] D. Basso, D. Rodríguez, R. García Martínez, “Comportamiento de métricas para proyectos de explotación de información en PyMEs”, XXI Congreso Argentino de Ciencias de la Computación, 2015.
- [52] Project Management Institute, Inc., “A guide to the project management body of knowledge (PMBOK® guide)” Fifth edition, 2013.
- [53] A. Velito y S. Tejada., “La historia clínica como instrumento de calidad”, Curso Anual de Auditoría

Médica del Hospital Alemán, Argentina, Buenos Aires, 2010.

- [54] S. Martins, D. Rodríguez, R. García-Martínez, “Deriving Processes of Information Mining Based on Semantic Nets and Frames. Lecture Notes on Artificial Intelligence”, Springer, Cham, 2014, vol. 8482, pp. 150-159, https://doi.org/10.1007/978-3-319-07467-2_16.
- [55] S. Umadevi, K. S. J. Marseline, "A survey on data mining classification algorithms," 2017 International Conference on Signal Processing and Communication (ICSPEC), 2017, pp. 264-268, doi: 10.1109/CSPEC.2017.8305851.
- [56] L. J. M. Camaré, “Aprendizaje automático a partir de conjuntos de datos no balanceados y su aplicación en el diagnóstico y pronóstico médico”, tesis doctoral, Instituto Nacional de Astrofísica, Óptica y Electrónica, 2008.
- [57] M. Ramasamy, S. Selvaraj, M. Mayilvaganan, “An empirical analysis of decision tree algorithms: Modeling hepatitis data”, In 2015 IEEE International Conference on Engineering and Technology (ICETECH), pp. 1-4, March 2015.