

ÍNDICE DE CONTENIDOS

INTRODUCCIÓN.....	1
CAPÍTULO 1: INTRODUCCIÓN.....	2
1.1. Lugar de aplicación	3
1.1.1. Modelo de negocio	3
1.1.2. Visión	4
1.1.3. Misión	4
1.1.4. Valores	4
1.1.5. Organigrama	4
1.1.6. Características del servicio	5
1.2. Propuesta de mejora	6
1.3. Objetivo general	7
1.4. Objetivos específicos	7
1.5. Resultados tangibles esperados	7
CAPÍTULO 2: MARCO TEÓRICO Y METODOLOGÍA.....	9
2.1. Marco teórico	10
2.1.1. Big data	10
2.1.2. Ciencia de datos	10
2.1.3. Minería de datos	11
2.1.4. Clasificación supervisada	11
2.1.5. Fuga de trabajadores	15
2.2. Metodología de solución	17
2.2.1. CRISP-DM	18
2.2.2. Design Thinking	19
2.2.3. Propuesta metodológica	21
2.3. Secuenciación de actividades	22
CAPÍTULO 3: ANÁLISIS Y DIAGNÓSTICO.....	23
3.1. Obtención de los datos	24
3.1.1. Servicio web	24
3.1.2. Endpoints	26
3.2. Calidad de los datos	28
3.2.1. Estado de los datos	31

CAPÍTULO 4: ESTRUCTURA DE MEJORA.....	36
4.1. Herramientas	37
4.1.1. Open Refine	37
4.1.2. KNIME	38
4.1.3. Power BI	41
CAPÍTULO 5: APLICACIÓN DE LA METODOLOGIA.....	43
5.1. Primera iteración	44
5.1.1. Preparación de datos	44
5.1.2. Construcción y evaluación del modelo	50
5.1.3. Evaluación	51
5.2. Segunda iteración	51
5.3. Tercera iteración	53
5.4. Cuarta iteración	55
5.5. Quinta iteración	57
5.6. Segmentación	61
5.6.1. Estratégico	62
5.6.2. Táctico	64
5.6.3. Operativo	65
5.7. Dashboard de validación	67
5.8. Propuesta de implementación	68
5.9. Predicción en fecha	70
CAPÍTULO 6: EVALUACION DE IMPACTO.....	74
6.1. Impacto económico	75
6.2. Impacto social	78
CONCLUSIONES.....	81
BIBLIOGRAFÍA.....	84
ANEXOS.....	86

ÍNDICE DE FIGURAS

Ilustración 1: Organigrama de la empresa.....	4
Ilustración 2: Grafico de K vecinos más cercanos	12
Ilustración 3: Representación del árbol de decisión	12

Ilustración 4: Representación del árbol de decisión	13
Ilustración 5: Representación del árbol de decisión	13
Ilustración 6: Representación de una red neuronal.....	14
Ilustración 7: Encuesta de metodologías usadas en minería de datos	18
Ilustración 8: Fases de CRISP-DM	18
Ilustración 9: Fases de design thinking.....	20
Ilustración 10: Diagrama BPMN de <i>Web Service</i>	25
Ilustración 11: Diagrama BPMN de la aplicación del proyecto	25
Ilustración 12: Categorización de variables en fuga de trabajadores	29
Ilustración 13: Variables a utilizar	29
Ilustración 14: Variables representadas	30
Ilustración 15: Grafico de causa de fuga	31
Ilustración 16: Grafico de datos faltantes	32
Ilustración 17: Grafico de sueldo base vs causa de fuga	33
Ilustración 18: Diagrama de cajas y bigotes del sueldo base	33
Ilustración 19: <i>Dashboard</i> de los datos	34
Ilustración 20: <i>Dashboard</i> de los datos filtrados.....	35
Ilustración 21: Interfaz de Open Refine	37
Ilustración 22: Interfaz de KNIME.....	38
Ilustración 23: Ciclo de minería de datos en KNIME	39
Ilustración 24: Cuadrante mágico de Gartner en <i>softwares</i> de minería de datos	41
Ilustración 25: Interfaz de Power BI.....	41
Ilustración 26: Cuadrante mágico de Gartner en <i>softwares</i> de inteligencia de negocio.....	42
Ilustración 27: Muestra de los valores de “Cargo” post limpieza	45
Ilustración 28: Valores del atributo "Cargo"	45
Ilustración 29: Nodo " <i>Column Filter</i> "	46
Ilustración 30: Nodo " <i>Missing value</i> "	47
Ilustración 31: Nodos " <i>Math Formula</i> "	48
Ilustración 32: Valores del atributo “Motivo término”	48
Ilustración 33: Nodo “ <i>String Replace (Dictionary)</i> ”.....	48
Ilustración 34: Grafico de fuga.....	49
Ilustración 35: Nodos que tratan con fechas.....	49
Ilustración 36: Nodos de clase y filtro	50
Ilustración 37: Nodo de partición	50

Ilustración 38: Nodo de entrenamiento	50
Ilustración 39: Nodo de predicción	50
Ilustración 40: Nodo constructor de la matriz de confusión.....	51
Ilustración 41: Nodo de filtrado de “Despidos”	52
Ilustración 42: Curva ROC para los modelos.....	52
Ilustración 43: Nodo de correlación lineal	53
Ilustración 44: Valores de tipo de contrato.....	55
Ilustración 45: Grafico de correlaciones.....	57
Ilustración 46: Nodos de validación cruzada.....	58
Ilustración 47: Árbol de decisión general resultante (muestra referencial).....	59
Ilustración 48: Diagrama de cajas y bigotes del sueldo base post preprocesamiento	61
Ilustración 49: Árbol de decisión del conjunto estratégico (muestra referencial).....	62
Ilustración 50: Diagrama de cajas y bigotes del sueldo base estratégico	63
Ilustración 51: Árbol de decisión del conjunto táctico (muestra referencial).....	64
Ilustración 52: Diagrama de cajas y bigotes del sueldo base operativo	65
Ilustración 53: Árbol de decisión del conjunto operativo (muestra referencial)	66
Ilustración 54: Diagrama de cajas y bigotes del sueldo base táctico.....	67
Ilustración 55: <i>Dashboard</i> de <i>testing</i>	68
Ilustración 56: Icono para configurar el espacio de trabajo.....	68
Ilustración 57: <i>KNIME Explorer</i>	69
Ilustración 58: Mostrar código API.....	69
Ilustración 59: Mostrar código	70
Ilustración 60: Código POST	70
Ilustración 61: Preparación de datos para predecir fecha de renuncia	71
Ilustración 62: Nodos de AutoML.....	71
Ilustración 63: Nodos de <i>Deep Learning Keras</i>	72
Ilustración 64: Optimización de capas ocultas	72

ÍNDICE DE TABLAS

Tabla 1: Revisión de literatura.....	16
Tabla 2: Resultados de precisión de la iteración 1	51
Tabla 3: Resultados de precisión de la iteración 2	52
Tabla 4: Resultados de correlación de las variables entre si	54

Tabla 5: Resultados de precisión de la iteración 3	54
Tabla 6: Resultados de precisión de la iteración 4	56
Tabla 7: Resultados de precisión de la iteración 5	57
Tabla 8: Resultados de la validación cruzada.....	58
Tabla 9: Ahorro en entrenamiento de cargos operativos.....	76
Tabla 10: Ahorro en entrenamiento de cargos tácticos	77
Tabla 11: Ahorro en entrenamiento de cargos operativos.....	77

ÍNDICE DE ANEXOS

Anexo 1: <i>Endpoint</i> de "Persona"	86
Anexo 2: Objeto "Detalle" proveniente de "Persona".....	86
Anexo 3: Objeto "Permiso" proveniente de "Persona"	86
Anexo 4: <i>Endpoint</i> de "Contratos"	86
Anexo 5: <i>Endpoint</i> de "Contratos v2"	88
Anexo 6: <i>Endpoint</i> de "Contratos resumido"	89
Anexo 7: <i>Endpoint</i> de "Cargos"	90
Anexo 8: <i>Endpoint</i> de "Vacaciones"	90
Anexo 9: <i>Endpoint</i> de "Vacaciones resumido"	90
Anexo 10: <i>Endpoint</i> de "Ausentismo".....	90
Anexo 11: <i>Endpoint</i> de "Ausentismo resumido".....	91
Anexo 12: <i>Endpoint</i> de "Prorrateo por centro de costo"	91
Anexo 13: <i>Endpoint</i> de "Centralización contable"	91
Anexo 14: <i>Endpoint</i> de "Asignación de ítems de pago"	92
Anexo 15: <i>Endpoint</i> de "Descarga de documentos de personas"	92
Anexo 16: <i>Endpoint</i> de "Creación de documentos de personas"	92
Anexo 17: <i>Endpoint</i> de "Días administrativos".....	92
Anexo 18: <i>Endpoint</i> de "Días administrativos resumido".....	92
Anexo 19: <i>Endpoint</i> de "Enrolamiento de firma digital"	93
Anexo 20: <i>Endpoint</i> de "Solicitud de firma digital"	93
Anexo 21: <i>Endpoint</i> de "Solicitud de firma digital" asociado al registro	93
Anexo 22: <i>Endpoint</i> de "Asignación de personas a turnos".....	93
Anexo 23: <i>Endpoint</i> de "Inyección y visualización de marcas".....	93
Anexo 24: <i>Endpoint</i> de "Días trabajados por contrato"	93

Anexo 25: *Endpoint* de “Turnos y horarios asignados por trabajador” 94
Anexo 26: Objeto “Asignación” de “Turnos y horarios asignados por trabajador” 94
Anexo 27: Actividades 94