



Extracción Automática de Documentos de Definición de Procesos en proyectos RPA a través de la Minería de Tareas y Decisiones

Escuela Técnica Superior de Ingeniería Informática

Grado en Ingeniería de la Salud

Trabajo Fin de Grado

Realizado por D. Alejandro Inglés Martínez

Supervisado por D. Antonio Martínez Rojas y Dr. D. José González Enríquez

Departamento Lenguajes y Sistemas Informáticos

Curso 2023/2024 - Sevilla, junio de 2024

Índice de contenidos

1	Introducción	10
1.1	Objetivo	10
1.2	Motivación	11
2	Planificación	12
2.1	Temporal	12
2.1.1	Estimación temporal	13
2.1.2	Desviación temporal	15
2.2	Costes	17
2.2.1	Estimación de costes	17
2.2.1.1	Costes directos	17
2.2.1.2	Costes indirectos	20
2.2.1.3	Presupuesto final	21
2.2.2	Desviación de costes	21
3	Contexto de la propuesta dentro del framework ScreenRPA	24
3.1	Introducción al Enriquecimiento del UI Log con ScreenRPA	24
3.2	Transformación del UI Log en un Dataset por cada punto de decisión	24
3.2.1	Creación del Dataset Aplanado Etiquetado	25
3.3	Generación y Explicación de Decisiones mediante Árboles de Decisión	26
3.3.1	Proceso de Entrenamiento del Árbol de Decisión	26
4	Propuesta: Mejora en la minería de decisiones aplicada en ScreenRPA mediante la identificación de reglas no deterministas combinando árboles de decisión	27
4.1	Proceso de mejora propuesto	27
4.1.1	Entrenamiento del modelo sobre el UI log completo:	27
4.1.2	Algoritmo para la Extracción de Reglas Superpuestas	27
4.2	Ejemplo de aplicación del proceso propuesto	29
5	Conceptos y materias aplicadas	34
5.1	Investigación relacionada	34
5.2	Metodologías	35
5.2.1	Metodologías ágiles	36
5.2.1.1	Scrum	37
5.2.1.2	Design Sprint y NDT	40
6	Desarrollo del sistema	41
6.1	Requisitos	41
6.1.1	Objetivos del sistema	42
6.1.1.1	Diagrama de objetivos	42
6.1.1.2	Definición de objetivos	42
6.1.2	Catálogo de requisitos	44
6.1.2.1	Requisitos de almacenamiento de información	44
6.1.2.2	Definición de actores	56
6.1.2.3	Requisitos funcionales	61

6.1.2.4	Requisitos no funcionales	65
6.2	Análisis	67
6.2.1	Modelo de clases de contenido	67
6.2.1.1	Diagrama de clases de contenido	68
6.2.1.2	Definición de clases de contenido	68
6.2.2	Modelos de clases de procesos	83
6.2.2.1	Diagrama de clases de procesos	83
6.2.2.2	Definición de clases de procesos	83
6.3	Diseño	85
6.3.1	Modelo de la base de datos	86
6.3.2	Arquitectura	87
6.3.2.1	MVC	87
6.3.2.2	MTV	87
6.3.2.3	Patrón Modelo-Vista-Controlador MVC	89
6.3.3	Análisis de alternativas tecnológicas	91
6.3.3.1	Control de versiones	91
6.3.3.2	Frontend	91
6.3.3.3	Backend	93
6.3.3.4	Base de datos	94
6.3.4	Prototipos	95
6.4	Implementación	102
6.4.1	Entorno de desarrollo	102
6.4.1.1	Visual Code	102
6.4.1.2	Python	103
6.4.1.3	Django 4.1.2	103
6.4.1.4	Django Template	103
6.4.1.5	Gestor de paquetes Anaconda	103
6.4.1.6	PgAdmin	104
6.4.2	Elección de librerías	104
6.4.2.1	Scikit-learn	104
6.4.2.2	Python-docx	104
6.4.2.3	Graphviz	104
6.4.2.4	Pydotplus	104
6.4.3	Pruebas	105
7	Conclusiones	113
7.1	Lecciones aprendidas	113
7.2	Consecución de los objetivos del proyecto	113
7.3	Trabajos futuros	113
8	Apéndices	113
8.1	Glosario de términos	113
8.2	Bibliografía	114
ANEXO	118

A.	Manual de instalación.....	118
a.	Despliegue	118
B.	Manual de usuario.....	119

Índice de figuras

Figura 1: Resultados del informe CHAOS [11].....	12
Figura 2. División de planificación temporal en bloques.....	14
Figura 3. Diagrama de Gantt estimado.....	15
Figura 4. Diagrama de Gantt real.....	16
Figura 5. Porcentaje de costes acumulados.....	23
Figura 6: Flujo de información RPA para la extracción de las reglas que deciden las decisiones del proceso.	25
Figura 7. Proceso de registro de venta de cemento.	29
Figura 8. Notificación por correo de oferta aceptada y pago realizado.	30
Figura 9. Pago realizado registrado en la plataforma de facturación.....	30
Figura 10. Dataset etiquetado utilizado para el ejemplo de experimentación.....	31
Figura 11. Representación gráfica del primer árbol de clasificación.....	31
Figura 12. Resumen del conocimiento general extraído durante el proceso.	33
Figura 13. Valores predichos vs valores reales.....	33
Figura 14. Matriz de confusión de resultados del clasificador.....	34
Figura 15. Ciclos de vida tradicionales vs ágiles.....	37
Figura 16. Origen del término scrum [31]	38
Figura 17. What is a standard scrum process.....	39
Figura 18. Fases de Design sprint.....	40
Figura 19: Diagrama de objetivos.....	42
Figura 20: Diagrama de requisitos de almacenamiento.....	45
Figura 21: Diagrama de actores.....	56
Figura 22: Diagrama de casos de uso.....	61
Figura 23 : RF-01.Ver detalle de una ejecución.....	62
Figura 24: RF-02.Crear reporte de una ejecución.....	64
Figura 25: Diagrama de requisitos no funcionales.....	65
Figura 26: Diagrama de clases de contenido.....	68
Figura 27 : Diagrama de clases de procesos.....	83
Figura 28: rf-01. ver detalle de una ejecución.....	84
Figura 29: rf-02.crear reporte de una ejecución.....	85
Figura 30. Modelo de la base de datos.....	86
Figura 31: Arquitectura de MVC.....	90
Figura 32. Git.....	91
Figura 33. GitHub.....	91
Figura 34. HTML5, CSS3 y Javascript.	92
Figura 35. JQuery.....	92

Figura 36. Bootstrap.....	93
Figura 37. Django.....	93
Figura 38. PostgreSQL.	94
Figura 39. Prototipo de vista en detalle de una ejecución.....	95
Figura 40. Prototipo de vista de en detalle de configuración de la fase “Log Processing”	96
Figura 41. Prototipo de vista en detalle de configuración de la fase “Extract Training Dataset”	97
Figura 42. Prototipo de vista en detalle de resultados de la fase “Log Processing” . ”	98
Figura 43. Prototipo de vista en detalle de resultados de la fase “Process Discovery”.	99
Figura 44. Prototipo de vista en detalle de resultados de la fase “Extract Training Dataset”.	100
Figura 45. Prototipo de vista en detalle de configuración para la generación de reporte. ”	101
Figura 46. Prototipo de vista en detalle de configuración de reporte. ”	102
Figura 47: PS-40.Crear reporte de una ejecución	106
Figura 48: PS-41.Crear reporte de una ejecución	107
Figura 49: PS-42.Crear reporte de una ejecución	109
Figura 50. PS-44.Ver detalle de una ejecución	111
Figura 51: PS-45.Ver detalle de una ejecución	112
Figura 52: Página de inicio	119
Figura 53: Formulario de registro de usuario relleno	120
Figura 54: Listado de estudio de casos vacío	120
Figura 55: Formulario de creación de un caso de estudio	121
Figura 56: Acceso a las configuraciones de las fases de un caso de estudio	122
Figura 57: Formulario de creación de configuración de la fase “UI Element Detection”	123
Figura 58: Activación de configuración de una fase	123
Figura 59: Formulario de creación de configuración de la fase “Feature Extraction”	124
Figura 60: Formulario de creación de configuración de la fase “Process Discovery”	125
Figura 61: Formulario de creación de configuración de la fase “Extract Training Dataset”	126
Figura 62: Formulario de creación de configuración de la fase “Decision Tree Training”	127
Figura 63: Ejecución de un caso de estudio	127
Figura 64: Vista en detalle de una ejecución	128
Figura 65: Fases realizadas durante la ejecución	129
Figura 66: Vista en detalle de configuración de la fase “UI Element Detection”	129
Figura 67: Vista de resultado de la fase “Feature Extraction”	130
Figura 68: Vista de resultado de la fase “Process Discovery”	131
Figura 69: Vista de una variante específica del proceso descubierto	131
Figura 70: Vista de resultado de la fase “Extract Training Dataset”	132
Figura 71: Vista de resultado de la fase “Decision Tree Training”	132
Figura 72: Zoom de las reglas descubiertas por el árbol de decisión	133
Figura 73: Zoom del árbol de decisión entrenado en la fase de “Decisión Tree Training”	133
Figura 74: Formulario de creación de reporte	134
Figura 75: Listado de reportes generado a partir de una ejecución	135
Figura 76: Detalle de configuración de un reporte	136
Figura 77: Comprimido con la descarga de un reporte	136

Figura 78: Introducción del reporte	137
Figura 79: Aplicaciones utilizadas por el usuario durante el proceso	138
Figura 80: Descripción gráfica del proceso descubierto	138
Figura 81: Flujo de actividades descrito en una variante	139
Figura 82: Descripción de actividad de click	140
Figura 83: Descripción de actividad de click fuera de resolución de pantalla	141
Figura 84: Descripción de actividad de escritura por teclado donde no se posee el texto escrito	141
Figura 85: Descripción de reglas descubiertas en un punto de decisión	142
Figura 86: Descripción gráfica de reglas deterministas	143
Figura 87: Descripción gráfica de reglas solapadas entre ramas	143
Figura 88: Datos iniciales de entrada en la plataforma	144

Índice de tablas

Tabla 1. Planificación temporal.....	13
Tabla 2. Desviación temporal.....	16
Tabla 3. Tabla salarial según convenio [12]	18
Tabla 4. Bases de cotización mínima de contrato de trabajo a tiempo parcial [13]	18
Tabla 5. Tipos de cotización: contingencias comunes, desempleo, FOGASA y FP [13]	18
Tabla 6. Coste de personal por hora	19
Tabla 7. Coste de personal total	19
Tabla 8. Costes de infraestructuras	20
Tabla 9. Estimación de coste total	21
Tabla 10. Desviación del coste de personal	22
Tabla 11. Desviación del coste total.....	22
Tabla 12 . OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA	42
Tabla 13: OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis	43
Tabla 14: OBJ-03.Aglutinar toda la información extraída en forma de reporte.....	43
Tabla 15 : OBJ-04.Aplicar nuevas técnicas de minería de decisiones para análisis más complejos	43
Tabla 16: OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma	43
Tabla 17. Nomenclatura de requisitos	44
Tabla 18 : «RA» RA-01.Decision Tree Training	46
Tabla 19 : «RA» RA-02.Feature Extraction Technique	47
Tabla 20: «RA» RA-03.Extract Training Dataset	48
Tabla 21: «RA» RA-04.UI Elements Classification	49
Tabla 22: «RA» RA-05.Monitoring.....	50
Tabla 23: «RA» RA-06.Prefilters	51
Tabla 24: «RA» RA-07.Postfilters	52
Tabla 25 : «RA» RA-08.Process Discovery.....	52
Tabla 26: «RA» RA-09.Reporting PDD	53
Tabla 27: «RA» RA-10.Case Study	54

Tabla 28: «RA» RA-12.UI Elemets Detection.....	54
Tabla 29 : «RA» RA-13.Execution	55
Tabla 30: «RA» RA-14.User	56
Tabla 31: «AC» AC-01.Usuario	60
Tabla 32: «AC» AC-02.Administrador.....	61
Tabla 33: «RF» RF-01.Ver detalle de una ejecución.....	63
Tabla 34: «RF» RF-21.Crear reporte de una ejecución	65
Tabla 35: «RNF» RNF-01.Disponibilidad.....	66
Tabla 36: «RNF» RNF-02.Escalabilidad.....	66
Tabla 37: «RNF» RNF-03.Usabilidad.....	67
Tabla 38: «RNF» RNF-04.Atomicidad	67
Tabla 39: «RNF» RNF-07.Reutilización	67
Tabla 41: «CL» CL-02.Feature Extraction Technique	71
Tabla 42: «CL» CL-03.Extract Training Dataset	72
Tabla 43: «CL» CL-04.UI Elements Classification.....	73
Tabla 44: «CL» CL-05.Monitoring.....	74
Tabla 45 : «CL» CL-06.Prefilters	75
Tabla 46 : «CL» CL-07.Postfilters.....	76
Tabla 47: «CL» CL-08.Process Discovery	77
Tabla 48: «CL» CL-09.Reporting PDD	78
Tabla 49: «CL» CL-10.Case Study	79
Tabla 50: «CL» CL-12.UI Elemets Detection.....	80
Tabla 51: «CL» CL-13.Execution	81
Tabla 52: «CL» CL-14.User	82
Tabla 53: PS-39.Crear reporte de una ejecución	105
Tabla 54: PS-40.Crear reporte de una ejecución	106
Tabla 55: PS-41.Crear reporte de una ejecución	108
Tabla 56: PS-42.Crear reporte de una ejecución	110
Tabla 57: PS-43.Ver detalle de una ejecución.....	110
Tabla 58: PS-44.Ver detalle de una ejecución.....	111
Tabla 59: PS-45.Ver detalle de una ejecución.....	112

Resumen

La Automatización Robótica de Procesos [1] (RPA) es una disciplina que crece cada vez más de la mano de la Inteligencia Artificial (IA) y el Aprendizaje automático (ML) que permiten hacer realidad la llamada automatización cognitiva.

En ese contexto, un desafío común entre las diferentes plataformas existentes es la legibilidad y accesibilidad al conocimiento generado por los procesos analizados. Con frecuencia, los resultados de estos procesos no permiten una interpretación directa ni un análisis transparente de las decisiones implementadas.

Para abordar estos desafíos, el proyecto propone un conjunto de soluciones orientadas a mejorar la experiencia del usuario en proyectos de RPA. En primer lugar, se presenta una propuesta de una técnica de minería de decisiones para identificar las reglas no deterministas que dictan las decisiones tomadas. En segundo lugar, se desarrollan mejoras en la plataforma para optimizar la accesibilidad y la interpretabilidad de los resultados obtenidos. Finalmente, se desarrolla un método para la generación automática de informes que sinteticen el conocimiento adquirido durante el proceso RPA.

Este enfoque integral busca no solo superar las barreras en la legibilidad y accesibilidad del conocimiento, sino también aumentar la eficiencia y efectividad de la Automatización Robótica de Procesos, reforzando así el impacto transformador de la tecnología en el ámbito empresarial.

Agradecimientos

A mis padres, por hacer esto posible, y a mis mentores, por ser mis referentes y darme las herramientas necesarias para crecer e iniciarme en el apasionante mundo de la investigación y el desarrollo.

1 Introducción

En esta memoria, se detalla el desarrollo de una propuesta que busca mejorar la interpretabilidad y accesibilidad del conocimiento generado por una plataforma existente de Automatización Robótica de Procesos (RPA) desarrollada por el grupo de ES3 (Engineering and Science for Software Systems) [2]. Antes de explicar cómo, primero es fundamental comprender qué es el RPA, los robots software y qué papel juegan estos en la eficiencia y automatización de procesos.

La Automatización Robótica de Procesos, o RPA, es una tecnología que permite automatizar tareas digitales a través de "robots software". Los robots software forman el core de RPA, son programas diseñados para simular la interacción humana con sistemas de información, descomponiendo cualquier proceso en bloques manejables que imitan acciones humanas para resolver tareas específicas.

A pesar de la importancia de la información extraída en procesos RPA, se ha identificado un desafío común entre las plataformas existentes: la dificultad en la accesibilidad e interpretación del conocimiento generado. Esta limitación obstaculiza la comprensión y el aprovechamiento del potencial completo de estas tecnologías. Un ejemplo es que en la actualidad tareas como la redacción de informes para los clientes suelen ser realizadas manualmente por analistas, lo que supone una clara oportunidad de mejora.

Con el objetivo de superar estos retos se han planificado varias mejoras específicas en la plataforma. La primera mejora es la investigación y aplicación de una nueva técnica de Decision Mining. Esta metodología se centra en extraer reglas que reflejen la toma de decisiones en las distintas variantes de los procesos. A diferencia de métodos más rígidos, esta técnica permite identificar reglas que no son deterministas y presentan solapamientos, lo cual refleja de manera más fiel la naturaleza de los escenarios del mundo real.

La segunda mejora consiste en el desarrollo de una parte de la plataforma que permita visualizar y gestionar la ejecución de los procesos, facilitando la observación detallada de los resultados obtenidos en cada fase. Esto, no solo mejora la transparencia, sino que también incrementa la capacidad de seguimiento y ajuste de los procesos en tiempo real.

La tercera y última mejora de la propuesta de este trabajo, es la generación automática de reportes que sean completamente comprensibles y fáciles de interpretar, incluso para profesionales que no sean expertos en RPA. Estos reportes documentarán exhaustivamente el conocimiento extraído durante la ejecución de las fases software del proceso RPA, cubriendo los resultados y decisiones claves de los procesos analizados.

Mediante estas mejoras, se espera no solo optimizar la funcionalidad de la plataforma, sino también transformar la manera en que los usuarios interactúan y se benefician de la tecnología RPA, potenciando su capacidad para tomar decisiones informadas basadas en datos procesados de forma más efectiva y accesible.

1.1 Objetivo

Tras un exhaustivo estudio del área se ha identificado que en el ámbito del RPA una característica con mucho potencial de mejora y que es de fundamental importancia para el correcto funcionamiento de

las plataformas que lo aplican, es la interpretabilidad y accesibilidad del conocimiento generado durante el proceso. Debido a esto, con este proyecto se desea abordar esta necesidad en una plataforma existente de Automatización Robótica de Procesos (RPA), optimizando así la funcionalidad de la plataforma y transformando la interacción de los usuarios con la tecnología RPA.

Para llevar a cabo esto, el trabajo fin de grado expondrá:

1. En primer lugar, la investigación e implementación de una nueva técnica de Decision Mining que permita extraer reglas que reflejen la toma de decisiones en las distintas variantes de los procesos RPA. Esta técnica deberá ser capaz de identificar reglas no deterministas y con solapamientos, reflejando así la complejidad de los escenarios reales y mejorando la toma de decisiones basada en procesos automatizados.
2. En segundo lugar, desarrollar una sección dentro de la plataforma que facilite la visualización y gestión de las ejecuciones de procesos y los resultados de sus fases.
3. Por último, crear un sistema para la generación de reportes automáticos y fácilmente interpretables, incluso para aquellos profesionales que no sean expertos en RPA. Estos reportes deberán documentar de manera exhaustiva el conocimiento extraído durante la ejecución de las fases del proceso RPA, asegurando una documentación detallada y accesible de los resultados y decisiones clave.

1.2 Motivación

El término Automatización Robótica de Procesos [1] (RPA) se refiere a un paradigma de software donde los robots son programas que imitan el comportamiento de los trabajadores humanos que interactúan con los sistemas de información (SI) [3]-[4], es decir, conjuntos de componentes que realizan acciones que resuelven una tarea particular de RPA. Tal paradigma se ha vuelto cada vez más popular debido a que el RPA despierta mucho interés para las organizaciones. En ese contexto, las soluciones basadas en la Inteligencia Artificial (IA) denominadas soluciones cognitivas de RPA [5] están recibiendo cada vez más atención, ya que la combinación de ambas disciplinas ofrece diversas ventajas. Por un lado, los métodos de IA mejoran las soluciones de RPA al proporcionar nuevas capacidades y, por el otro, las soluciones RPA producen datos relativos a la propia ejecución de los procesos que sirven para surtir a los modelos de conocimiento de datos que permiten mejorar mediante el aprendizaje, el rendimiento y la precisión de las propuestas basadas en IA. Es decir, permiten un entrenamiento continuo de los modelos de IA. Por lo tanto, las principales plataformas RPA [6] (por ejemplo, BluePrism [7], UiPath [8] y Automation Anywhere [9]) ya tienen una batería de técnicas basadas en IA.

En el contexto de un proyecto RPA, el cliente del RPA [10] –las personas que contratan los servicios y necesitan de la plataforma para automatizar sus procesos– debe tomar decisiones basadas en el análisis de los datos extraídos. Para ello, es crucial que el cliente comprenda claramente los resultados de la plataforma, ya que uno de los factores clave que conducen al fracaso de los proyectos de RPA es la falta de comprensión por parte de los clientes de la información extraída, y por lo tanto no pueden expresar de forma satisfactoria esta. Esto es especialmente importante en el contexto de las soluciones de RPA basadas en IA, ya que dificultan de una manera considerable la comprensibilidad de los resultados debido a la complejidad inherente a estas técnicas y que el cliente no tiene por qué conocer. Por lo que hay que “dárselo ya digerido”.

Por otro lado, desde el punto de vista de los proveedores RPA –(empresa que desarrolla y ofrece los servicios de RPA), la tarea de documentar y digerir esta información es altamente compleja, y requiere

de expertos analistas que comprendan a fondo tanto el problema del proceso a automatizar como el funcionamiento de la plataforma RPA.

Para proporcionar apoyo tanto a los clientes como a los analistas de RPA en el contexto de un proyecto de IA-RPA, este trabajo propone una serie de mejoras para aumentar la comprensibilidad de los resultados y conocimiento extraído.

2 Planificación

En esta sección se abordará tanto la planificación de tiempos como de presupuesto, elementos esenciales para el avance del proyecto y claves en su éxito final. La previsión inicial es un desafío notorio debido a que las estimaciones de tiempo y coste frecuentemente experimentan desviaciones significativas, tal como se evidencia en el informe CHAOS [11].

Desde el año 2015, The Standish Group ha identificado tres aspectos críticos para el éxito de un proyecto: el respeto a los plazos establecidos, la adherencia al presupuesto y la consecución de resultados que satisfagan las expectativas. A diferencia de ediciones previas del Informe CHAOS, donde se valoraban los plazos, el presupuesto y el cumplimiento del alcance del proyecto, la nueva perspectiva permite cierta flexibilidad en cuanto al alcance completo del proyecto, aunque aún se enfatiza la importancia de los plazos y el presupuesto. La etapa de planificación de tiempos y costes, por su parte, representa una fuente considerable de dificultades, dado que el 60% de los proyectos analizados en este informe [11] registraron desviaciones en los tiempos y el 56% en los costes, elementos estrechamente vinculados al éxito del proyecto.

TRADITIONAL RESOLUTION FOR ALL PROJECTS					
	2011	2012	2013	2014	2015
SUCCESSFUL	39%	37%	41%	36%	36%
CHALLENGED	39%	46%	40%	47%	45%
FAILED	22%	17%	19%	17%	19%

The Traditional resolution of all software projects from FY2011–2015 within the new CHAOS database.

Figura 1. Resultados del informe CHAOS [11]

2.1 Temporal

La organización temporal del proyecto incluye la fase en la que se identifican las tareas necesarias, se asignan los tiempos y recursos a estas actividades, y se establece un orden de ejecución para minimizar la duración total del proyecto. Posteriormente, se define la ruta crítica, que es simplemente el cálculo del tiempo mínimo requerido para completar el proyecto de la manera más eficaz. Además, se mantiene un monitoreo constante sobre el avance del proyecto para identificar y gestionar posibles demoras.

En el contexto específico de este proyecto, la planificación quedará notablemente más simplificada ya que solo involucra a una persona en el equipo de trabajo. Esto implica que no se pueden llevar a cabo tareas simultáneamente ni hay posibilidad de especializar a los miembros del equipo en distintas disciplinas. Por lo tanto, todas las tareas se planificarán de manera secuencial y serán asignadas a la misma persona.

Además, se debe considerar que el proyecto incorporará tecnologías que son nuevas para el autor y que, dada la esencia del proyecto, se abordarán tareas relacionadas con las etapas de investigación e innovación. Esto introduce una complejidad adicional en la estimación del tiempo y esfuerzo requeridos para completar el proyecto.

2.1.1 Estimación temporal

A continuación, se desarrolla la planificación temporal inicial del proyecto, siguiendo la metodología Scrum desarrollada en la sección 5.2.1.1 Scrum.

Tarea	Fecha de inicio	Fecha de fin	Duración (horas)
Investigación			
Estudio del ámbito de Process Mining y Decision Mining	21/03/2024	05/04/2024	20
Desarrollo e implementación de técnica de Decision Mining sobre UI logs	25/03/2024	05/04/2024	35
Desarrollo y memoria			
Planificación temporal y de costes	26/02/2024	28/02/2024	15
Estudio del estado del arte	28/02/2024	29/02/2024	5
Estudio de las metodologías, tecnologías y herramientas a utilizar	29/02/2024	08/03/2024	30
Elicitación de requisitos	08/03/2024	15/03/2024	25
Diseño del sistema	15/03/2024	21/03/2024	20
Desarrollo de la aplicación	05/04/2024	29/04/2024	80
Pruebas	29/04/2024	03/05/2024	20
Instalación e implantación	03/05/2024	07/05/2024	10
Manual de usuario	07/05/2024	09/05/2024	10
Análisis y cierre del proyecto	09/05/2024	14/05/2024	15
Presentación del proyecto			
Preparación para la defensa del proyecto	05/06/2024	07/06/2024	15
Total de horas			300 horas

Tabla 1. Planificación temporal

Esta división se ha realizado teniendo en cuenta los grandes bloques dentro del proyecto, que son aquellas tareas relacionadas con el trabajado de investigación en el ámbito de Decision Mining (55 horas), el desarrollo del proyecto y la memoria (230 horas), y la presentación del proyecto (15 horas).

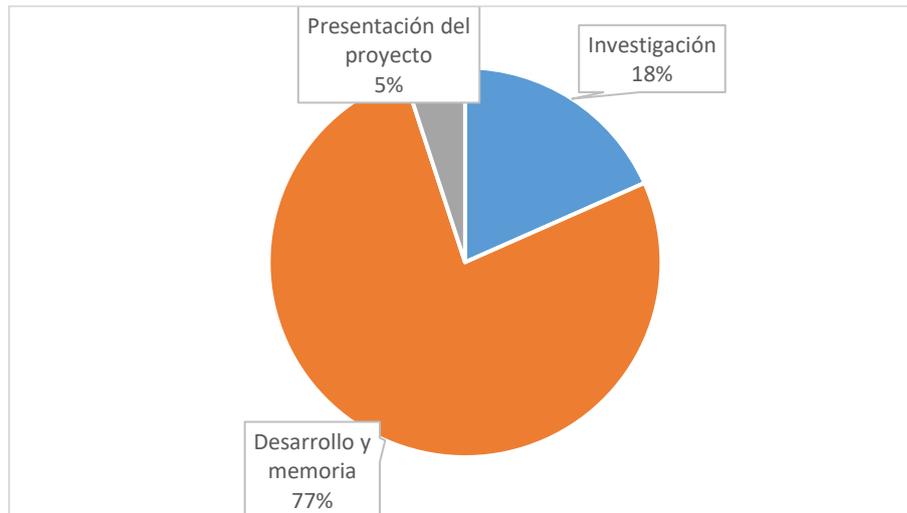


Figura 2. División de planificación temporal en bloques

Dentro del bloque de investigación se agrupan las tareas de estudio en el ámbito de Process Mining y Decision Mining a partir de la lectura de artículos científicos relacionados, además del desarrollo e implementación de una técnica de Decision Mining sobre UI logs para encontrar reglas solapantes. El bloque de desarrollo y memoria, como su nombre indica, implica el propio desarrollo del sistema de información, además de la redacción de la presente memoria. Este bloque es donde se agrupan entre otras tareas, la planificación, que incluye lo tratado en este apartado, la elicitación de requisitos, el diseño del sistema, la redacción del plan de pruebas, los resultados de su ejecución o el cierre del proyecto. Se finaliza con la preparación para la presentación del proyecto.

A continuación, se muestra el diagrama de Gantt asociado a la planificación general de este proyecto, para clarificar la **distribución** temporal de sus tareas:

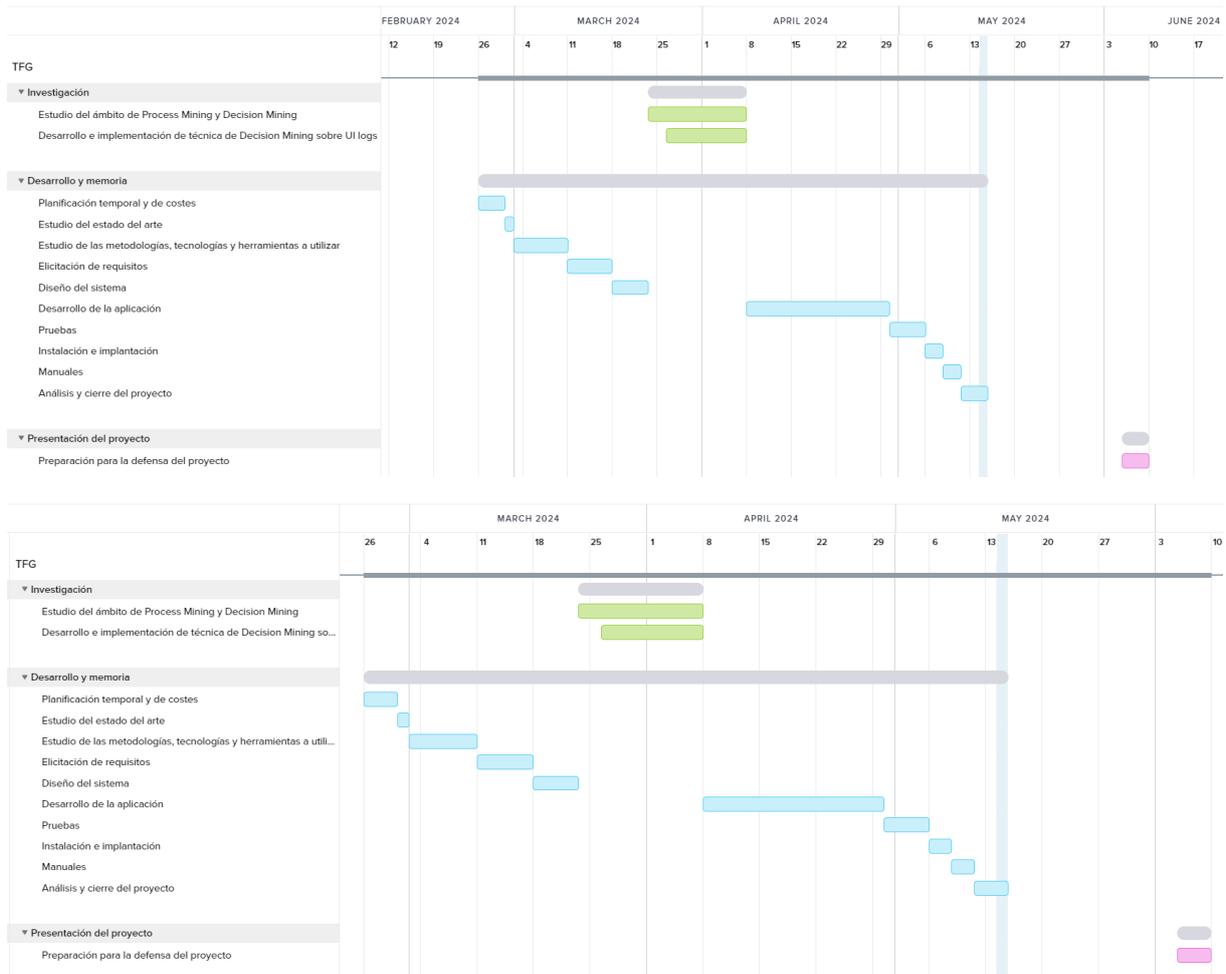


Figura 3. Diagrama de Gantt estimado

2.1.2 Desviación temporal

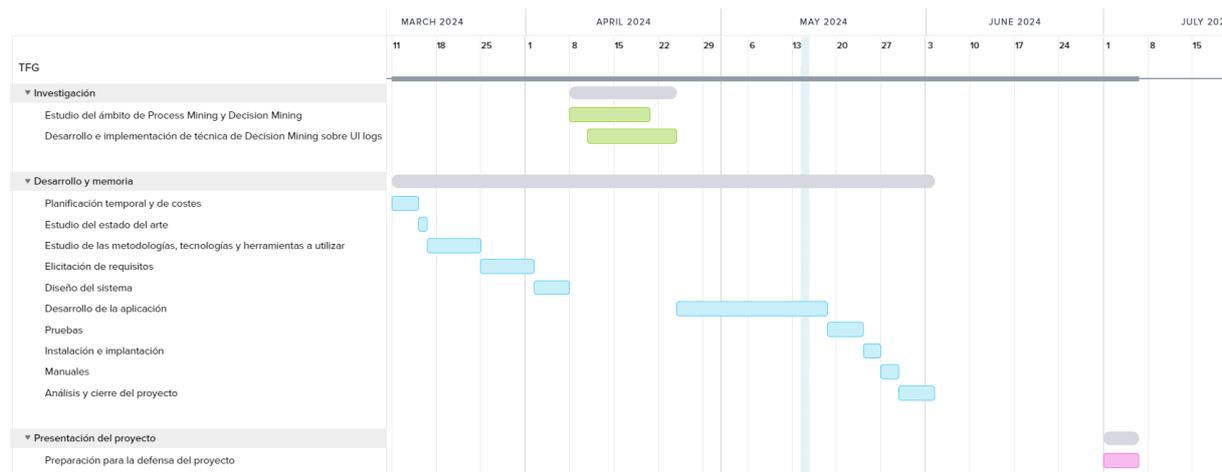
Este apartado muestra el tiempo real que se ha empleado en el desarrollo de las tareas durante el desarrollo del proyecto, frente a la estimación inicial.

Tarea	Fecha de inicio	Fecha de fin	Duración (horas)
Investigación			
Estudio del ámbito de Process Mining y Decision Mining	05/04/2024	18/04/2024	20
Desarrollo e implementación de técnica de Decision Mining sobre UI logs	09/03/2024	23/04/2024	35 + 5
Desarrollo y memoria			
Planificación temporal y de costes	11/03/2024	13/02/2024	15
Estudio del estado del arte	13/03/2024	14/03/2024	5
Estudio de las metodologías, tecnologías y herramientas a utilizar	14/03/2024	22/03/2024	30
Elicitación de requisitos	22/03/2024	01/04/2024	25 + 5
Diseño del sistema	01/04/2024	05/04/2024	20
Desarrollo de la aplicación	23/04/2024	16/05/2024	80 + 5
Pruebas	16/05/2024	22/05/2024	20
Instalación e implantación	22/05/2024	24/05/2024	10
Manual de usuario	24/05/2024	28/05/2024	10
Análisis y cierre del proyecto	28/05/2024	03/06/2024	15 + 5
Presentación del proyecto			
Preparación para la defensa del proyecto	01/07/2024	04/07/2024	15 + 5
Total de horas			300 + 25 horas

Tabla 2. Desviación temporal

Se puede observar que la desviación con respecto a la estimación inicial de la duración del proyecto fue del 7.7%, lo cual incrementó su tiempo de ejecución en 25 horas. Además, la fecha de inicio del proyecto se retrasó del 26 de febrero al 11 de marzo, lo que desplazó la fecha de finalización del 14 de mayo al 3 de junio.

Por consiguiente, el diagrama de Gantt real sufriría cambios en las fechas de inicio y finalización de la mayoría de sus tareas debido a que existen retrasos acumulados en tareas iniciales del proyecto, con lo que resultaría de esta manera.



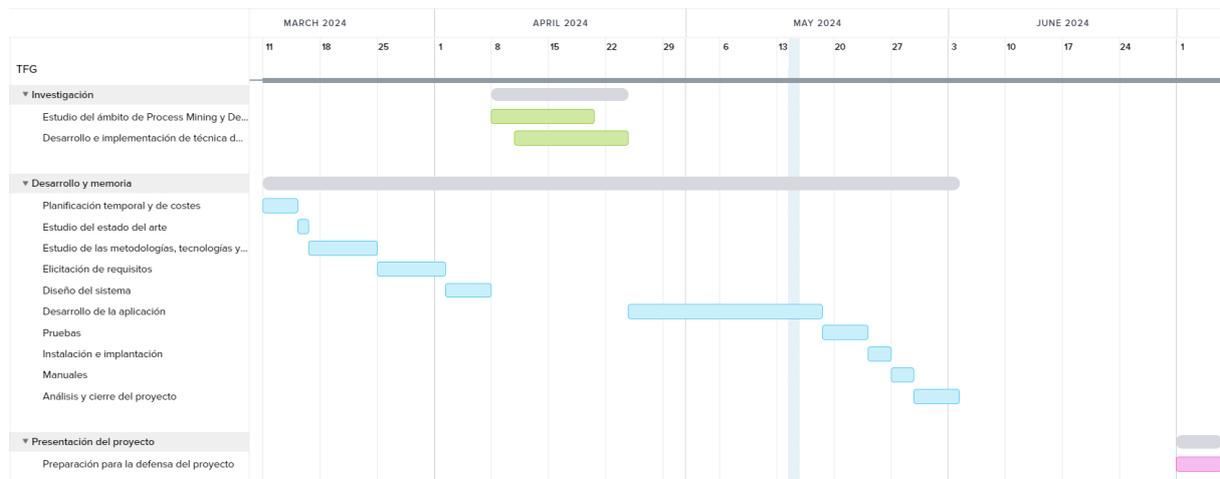


Figura 4. Diagrama de Gantt real

Debido a estos retrasos, tanto en el inicio como en la duración del proyecto, este se presentó para su evaluación en la segunda convocatoria. Se llegó con holgura a esta fecha, beneficiándose de un margen de varias semanas previsto hasta la fecha límite de entrega. Este margen se planificó especialmente, considerando la naturaleza de investigación e innovación de algunas etapas del proyecto. Estas actividades, al ser más complejas de prever y debido a la experiencia limitada del autor en estos temas, tienden a subestimarse en tiempo, resultando en una previsión inicial demasiado optimista en comparación con el tiempo real requerido.

2.2 Costes

La planificación de costes constituye una fase esencial y crítica en cualquier proyecto de ingeniería, puesto que el objetivo final de cualquier proyecto es obtener un retorno, ya sea económico en proyectos de consultoría, o en forma de un producto innovador en proyectos de I+D+i. En todos estos casos, es crucial contar con una idea precisa del costo total del proyecto, a pesar de que se base en estimaciones, ya que esto es fundamental para tomar decisiones respecto al mismo.

2.2.1 Estimación de costes

2.2.1.1 Costes directos

Se expone a continuación una estimación minuciosa de los costes directos vinculados a este proyecto, iniciando con una especificación de los **costes de personal**. A pesar de que el proyecto está coordinado por un único responsable, su magnitud y complejidad demandan un análisis como si fuese ejecutado por un equipo multidisciplinar. A continuación, se presenta una tabla que detalla los roles involucrados en el proyecto y la remuneración asignada a cada uno, conforme a lo establecido en el convenio colectivo estatal.

Rol	Salario anual	Salario bruto por hora
Jefe de proyecto	28.850,23 €	16,02 €
Analista	27.985,64 €	15,55 €
Programador senior	26.534,12 €	14,73 €

Tabla 3. Tabla salarial según convenio [12]

Estos datos, suministrados por el convenio colectivo estatal de empresas de consultoría [12], sirven como referencia para calcular el coste bruto por hora de cada rol. Sin embargo, estos no representan el coste real para el proyecto, puesto que al salario por hora se debe añadir el coste estimado asociado a la seguridad social, el cual es responsabilidad de la empresa. Estos costes adicionales se calcularán usando los porcentajes proporcionados por el Ministerio de Inclusión, Seguridad Social y Migraciones en el régimen general de la seguridad social del año 2024 [13].

Grupo de Cotización	Categorías Profesionales	Bases mínimas/hora
1	Ingenieros y Licenciados. Personal de alta dirección no incluido en el artículo 1.3.c) del Estatuto de los Trabajadores	11,13

Tabla 4. Bases de cotización mínima de contrato de trabajo a tiempo parcial [13]

El cálculo de la base de cotización se determina sumando todos los conceptos salariales de la nómina del trabajador. Es importante señalar que, para todos los roles implicados, la base supera el mínimo estipulado en la Tabla 4 para el Grupo de Cotización 1, correspondiente a ingenieros y licenciados. Los costes sociales a cargo de la empresa se calcularán utilizando los porcentajes especificados a continuación aplicables a contratos de duración determinada a tiempo parcial y a la categoría profesional indicada.

	Empresa	Trabajadores	Total
Tipo de cotización: Contingencias comunes	23,60	4,70	28,30
Desempleo: Contrato duración determinada tiempo parcial	6,70	1,60	8,30
FOGASA	0,20	-	0,20
Formación profesional	0,60	0,10	0,70
Total acumulado	31,10 %	6,40 %	

Tabla 5. Tipos de cotización: contingencias comunes, desempleo, FOGASA y FP [13]

Estas cotizaciones incluyen coberturas para contingencias comunes (como enfermedades comunes, maternidad o accidentes no laborales), desempleo, formación profesional (para la certificación y el reciclaje profesional), y FOGASA (Fondo de Garantía Salarial, que asegura los pagos a los trabajadores ante la insolvencia del empleador).

Además, se añadirá un porcentaje adicional del 1% para Incapacidad, Muerte y Supervivencia (IMS) y 0.65% para Incapacidad Temporal (IT) derivados del trabajo, sumando un total de 1.65%. De este modo, los costes sociales de la empresa equivalen al 31.10% + 1.65% del salario bruto percibido por el trabajador.

El cálculo del coste total por rol sería:

Coste por hora = Salario bruto €/hora + (Salario bruto €/hora * 32,75/100)

Esto se aplica a cada uno de los roles involucrados, reflejando los costes por rol en la Tabla 6

Rol	Salario bruto por hora	Costes sociales de la empresa	Coste por rol
Jefe de proyecto	16,02 €	5,25	21,26
Analista	15,55 €	5,09	20,64
Programador senior	14,73 €	4,82	19,55

Tabla 6. Coste de personal por hora

Partiendo de estos datos y considerando que el proyecto tiene una duración de 9 semanas, que abarcará actividades tanto de investigación y documentación como de desarrollo y pruebas, se procederá a estimar el coste total de personal utilizando la Planificación Temporal para determinar el número de tareas y su duración, que se multiplicará por el coste por hora del rol correspondiente.

Tarea	Rol	Duración (horas)	Coste (euros)
Investigación			
Estudio del ámbito de Process Mining y Decision Mining	Analista	20	412.80
Desarrollo e implementación de técnica de Decision Mining sobre UI logs	Analista	35	722.40
Desarrollo y memoria			
Planificación temporal y de costes	Jefe de proyecto	15	318,90
Estudio del estado del arte	Analista	5	103,20
Estudio de las metodologías, tecnologías y herramientas a utilizar	Jefe de proyecto	15	318,90
	Analista	15	309.60
Elicitación de requisitos	Jefe de proyecto	5	106.30
	Analista	20	412.80
Diseño del sistema	Analista	20	412.80
Desarrollo de la aplicación	Programador senior	80	1.564.00
Pruebas	Analista	10	206,40
	Programador senior	10	195.50
Instalación e implantación	Programador senior	10	195.50
Manual de usuario	Analista	10	206,40
Análisis y cierre del proyecto	Jefe de proyecto	15	318,90
Presentación del proyecto			
Preparación para la defensa del proyecto	Jefe de proyecto	15	318,90
Total		300	6.123,30

Tabla 7. Coste de personal total

Finalmente se determina que para la ejecución del proyecto serán necesarias 65 horas del rol de Jefe de proyecto a un costo de 21,26 € por hora, 145 horas del rol de Analista a 20,64 € por hora, y 100 horas de Programador senior a 19,55 € por hora. Esto resulta en un costo respectivo de 1.381,90 € para el Jefe de proyecto, 2.992,80 € para el Analista y 1.955,00 € para el Programador senior, sumando un costo total estimado de personal de **6.123,30 €**.

Tras establecer los costes de personal, y para concluir con los costes directos del proyecto, a continuación se detallan los **costes de material** asociados al proyecto:

Hardware

Para el desempeño de las funciones de los diversos roles, se ha seleccionado un único modelo de ordenador portátil con especificaciones estándar, requiriendo un total de tres unidades, una por cada rol implicado. Por lo tanto, los costes de hardware corresponden a tres unidades del modelo MSI Modern 15 B13M-281XES, equipados con procesador Intel i7-1355U, 16 GB de memoria RAM, 1 TB de disco duro SSD y una pantalla de 15.6 pulgadas, cada uno valorado en 729 € [14]. Esto resulta en un coste total de hardware de $3 \times 729 = 2,187$ €. Este gasto se considerará dentro de los costes indirectos y, por lo tanto, no se sumará al total de costes de material, ya que los equipos informáticos son considerados bienes amortizables según la Agencia Tributaria [15].

Software

En cuanto al software utilizado, las tecnologías implementadas en el proyecto, que se especificarán más adelante, incluyen el sistema operativo de los ordenadores empleados en este proyecto, que es Windows 11 Pro. El coste por licencia es de 289 € [16], resultando en un total de $3 \times 289 = 867$ € para las tres licencias.

2.2.1.2 Costes indirectos

Los costes indirectos son aquellos que no pueden ser atribuidos directamente a un objeto de costo específico, como un proyecto, instalación, función o producto, afectando más bien al proceso productivo de manera general. Estos costes pueden ser tanto fijos como variables, e incluyen gastos de administración, personal y seguridad, entre otros. Mientras que algunos gastos generales se pueden atribuir directamente a un proyecto y son considerados gastos directos [17], otros afectan a la operación en su conjunto.

En este contexto, se realizará una estimación de estos costes, clasificados en amortizaciones e infraestructuras, agrupando en cada categoría las amortizaciones de bienes, alquileres, gastos de energía, seguros y otros costes similares.

Infraestructuras

Los costes de infraestructuras asociados al proyecto comprenderán los gastos de mantenimiento del establecimiento donde este se desarrolle, de forma que tanto el alquiler como los gastos derivados del propio inmueble, como luz, agua o comunidad, entran dentro de esta categoría. En esta tabla se detalla la estimación de ellos para este proyecto.

Descripción	Duración	Coste	Total
Alquiler de oficina. Calle. Luis Montoto, 84.	3 meses	300 €/mes [18]	900 €
Conexión a internet. Lowi Fibra 300 Mbps	3 meses	25,95 €/mes + 30 € instalación	80,28 €
Luz	3 meses y 3 personas	0,0892 €/kWh [19]	78,34 €
Agua	3 meses y 3 personas	2,25 €/m ³ [20]	13,50 €
Total			1.072,12 €

Tabla 8. Costes de infraestructuras

Amortizaciones

Como se ha indicado previamente, los equipos informáticos se clasifican como bienes amortizables. De acuerdo con la tabla de coeficientes de amortización lineal, dentro de la sección "Equipos electrónicos e informáticos", se especifica la categoría "Equipos para procesos de información". Según establece la Agencia Tributaria [15], estos equipos son susceptibles de una amortización de hasta el 25% anual durante un período máximo de 8 años. Así, cada uno de los portátiles, que tiene un valor de 729 €, puede ser amortizado en una cantidad de 182,25 € durante los primeros cuatro años de uso. Esto implica que la amortización mensual, que utilizaremos solo para calcular el valor del coste durante la duración del proyecto, sería de 15,19 €. Por lo tanto, resultaría en una amortización de 45,57 € por cada equipo durante los tres meses que dura el proyecto, sumando un total de amortización de 136,71 €.

Con lo que tendríamos unos costes indirectos estimados de **1.208,83 €**.

2.2.2 Presupuesto final

En resumen del análisis efectuado en las secciones anteriores, el total estimado de los costes del proyecto sería de **10.522,84 €**. Este resultado surge de la acumulación de los costes detallados, los cuales se reflejarán en la Tabla 9.

Tipo de coste	Total
Coste de personal	6.123,30 €
Coste de material	3.054,00 €
Coste de infraestructura	1.208,83 €
Coste de amortización	136,71 €
	10.522,84€

Tabla 9. Estimación de coste total

2.2.3 Desviación de costes

La Tabla 11 presenta la desviación observada respecto a la estimación inicial de costes. El coste total estimado al comienzo del proyecto era de 10.522,84 €, mientras que el coste real al finalizar fue de **10.916,03 €**. Esto representa un 103,73% del coste inicial, con una diferencia de 393,19 €.

Este aumento en el coste se debe a desviaciones en la planificación temporal de las tareas, lo que a su vez ocasionó desviaciones en los costes de personal debido al tiempo adicional necesario para completar estas.

Tarea	Rol	Duración (horas)	Coste (euros)
Investigación			
Estudio del ámbito de Process Mining y Decision Mining	Analista	20	412,80
Desarrollo e implementación de técnica de Decision Mining sobre UI logs	Analista	35 + 5	825,60
Desarrollo y memoria			
Planificación temporal y de costes	Jefe de proyecto	15	318,90
Estudio del estado del arte	Analista	5	103,20
Estudio de las metodologías, tecnologías y herramientas a utilizar	Jefe de proyecto	15	318,90
	Analista	15	309,60
Elicitación de requisitos	Jefe de proyecto	5 + 2	148,82
	Analista	20 + 3	474,72
Diseño del sistema	Analista	20	412,80
Desarrollo de la aplicación	Programador senior	80 + 5	1.661,75
Pruebas	Analista	10	206,40
	Programador senior	10	195,50
Instalación e implantación	Programador senior	10	195,50
Manual de usuario	Analista	10	206,40
Análisis y cierre del proyecto	Jefe de proyecto	15 + 5	425,20
Presentación del proyecto			
Preparación para la defensa del proyecto	Jefe de proyecto	15 + 5	425,20
Total		325	6.516,49

Tabla 10. Desviación del coste de personal

Comparado con un valor inicial de 6.123,30 euros, la desviación total en los costes se relaciona directamente con el aumento en los gastos de personal, resultado de un incremento de 25 horas de trabajo. No se registraron modificaciones en los costes de materiales o indirectos, dado que estos corresponden a tarifas anuales o mensuales. En resumen, los costes solo se desviaron un 4,22% con respecto al valor inicial, lo cual representa una desviación mínima para un proyecto de este tamaño. Los resultados detallados se muestran en la Tabla 11.

Tipo de coste	Total
Coste de personal	6.516,49 €
Coste de material	3.054,00 €
Coste de infraestructura	1.208,83 €
Coste de amortización	136,71 €
	10.916,03 €

Tabla 11. Desviación del coste total

Con lo que se concluye con una distribución del coste real del proyecto tal y como se muestra en la figura 5.

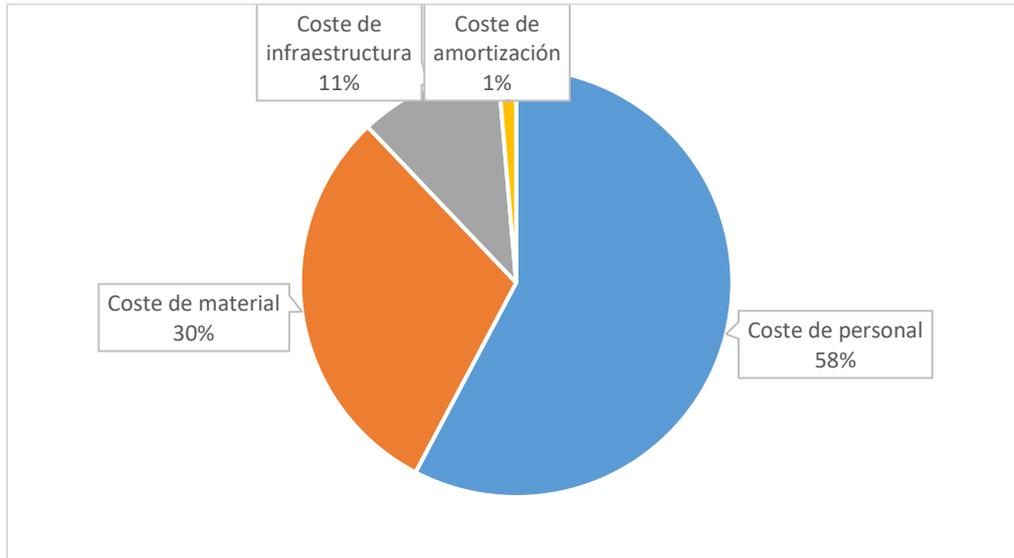


Figura 5. Porcentaje de costes acumulados

3 Contexto de la propuesta dentro del framework ScreenRPA

3.1 Introducción al Enriquecimiento del UI Log con ScreenRPA

El primer paso en nuestro enfoque es partir de un registro de interfaz de usuario (UI log) que ha sido enriquecido por el framework ScreenRPA. ScreenRPA ofrece herramientas y técnicas avanzadas para el procesamiento de UI logs, permitiendo un enriquecimiento significativo de los eventos capturados a partir de la información obtenida de las capturas de pantalla asociadas con cada evento registrado. Estos logs son esenciales, ya que registran todas las interacciones del usuario con la interfaz, capturando detalles cruciales que pueden ser analizados posteriormente para obtener una comprensión más profunda y detallada del comportamiento del usuario y las decisiones tomadas durante el proceso.

El enriquecimiento de los UI logs proporciona un contexto visual crucial para analizar el comportamiento del usuario. Las capturas de pantalla incorporadas ayudan a comprender las acciones del usuario, permitiendo a los analistas observar con precisión qué elementos de la interfaz fueron interactuados y en qué condiciones. Este nivel de detalle es fundamental para comprender no solo las acciones específicas que los usuarios realizan, sino también el flujo de trabajo y las decisiones que influyen en el curso del proceso, facilitando así mejoras significativas en la interfaz y la eficiencia del proceso de negocio.

3.2 Transformación del UI Log en un Dataset por cada punto de decisión

El proceso de transformación del UI log enriquecido en un dataset estructurado se realiza meticulosamente para cada punto de decisión dentro del proceso, generando así tantos “flatned dataset” como puntos de decisión tenga el proceso, ya que cada punto de decisión requiere un análisis detallado y específico en la plataforma ScreenRPA. La plataforma utiliza un procedimiento estandarizado para implementar esta transformación de manera eficiente y efectiva.

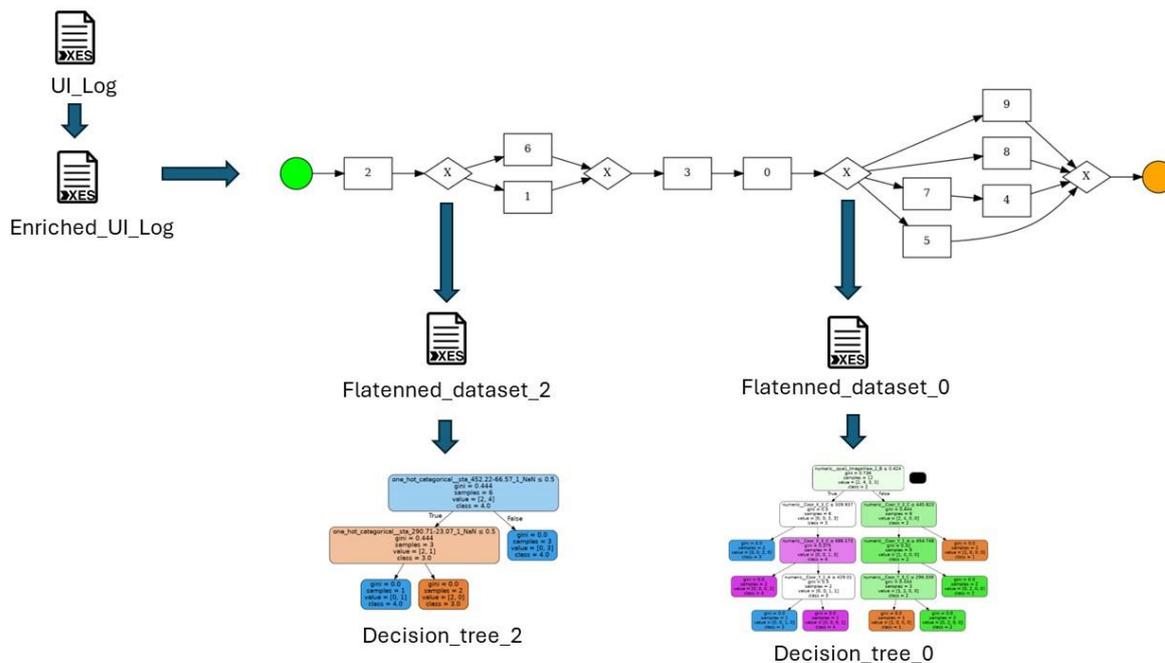


Figura 6: Flujo de información RPA para la extracción de las reglas que deciden las decisiones del proceso.

3.2.1 Creación del Dataset Aplanado Etiquetado

Para convertir el UI log en un dataset funcional para la aplicación de algoritmos de aprendizaje automático, es crucial seguir un conjunto de pasos sistemáticos, tal como se describe en el pseudocódigo que describe el Algoritmo 1. Este algoritmo está diseñado para procesar cada caso registrado en el log, identificando y extrayendo los eventos que ocurren antes de cada punto de decisión específico.

```

Input  : UILog log, DecisionPoint decision
Output : Dataset dataset
1 dataset ← ∅
2 instances ← separateLogByCases(log)
3 for instance in instances do
4     row ← ∅
5     for event in instance.eventsBefore(decision) do
6         for attr in event.attributes() do
7             row ← (column=event.activityID + '_' + attr.name, value= attr.value)
8         end
9     end
10    row ← (column='LABEL', value=decision.branch(instance))
11    dataset ← row
12 end
13 return dataset
    
```

Algoritmo 1: Algoritmo para transformar un UI log a dataset para un punto de decisión

Mediante este método, los eventos se aplanan en una sola fila por cada observación. Se incluye cada atributo de los eventos como una nueva columna en la fila correspondiente, y se asegura que los nombres de las columnas sean únicos para prevenir cualquier posibilidad de colisión entre ellos.

El dataset resultante está estructurado de tal manera que cada fila representa un conjunto de eventos que conducen a un punto de decisión. La etiqueta ("label") de cada fila indica la decisión que se tomó en ese momento. Este formato estructurado asegura que el dataset sea especialmente adecuado para el entrenamiento de algoritmos de clasificación, permitiendo que la plataforma aprenda de los datos registrados en el UI log pudiendo explicar las decisiones basadas en el contexto y los datos específicos de los eventos registrados. Al dataset resultante de este proceso, se le denomina dataset aplanado o en inglés, "flattened dataset".

3.3 Generación y Explicación de Decisiones mediante Árboles de Decisión

Una vez generado el dataset etiquetado, el siguiente paso dentro de la plataforma es entrenar un modelo de clasificación para explicar y analizar las decisiones registradas en el UI log. Hasta ahora, la metodología adoptada por la plataforma se basa en la propuesta de Rozinat et al. [21], utilizando árboles de decisión para llevar a cabo el análisis de los puntos de decisión (decision mining). La principal razón para elegir árboles de decisión sobre otros modelos de clasificación radica en su alta explicabilidad. Los árboles de decisión facilitan la interpretación y comprensión claras por parte de humanos gracias a la visibilidad directa de las reglas descubiertas que determinan cada decisión.

3.3.1 Proceso de Entrenamiento del Árbol de Decisión

El proceso de entrenamiento del modelo en la plataforma para analizar un punto de decisión consiste en ajustar un árbol de decisión a los datos del dataset etiquetado. Los atributos (columnas) del dataset se emplean para predecir las etiquetas (decisiones). El árbol de decisión resultante ofrece una representación visual clara de las reglas que explican las decisiones tomadas; cada nodo del árbol representa una regla crucial del proceso, mientras que cada hoja del árbol corresponde a una etiqueta (decisión).

Sin embargo, es crucial ser consciente del riesgo de sobreajuste (overfitting) durante el entrenamiento del modelo. Un árbol de decisión que muestra una precisión del 100% en los datos de entrenamiento podría estar capturando ruido y valores atípicos, adaptándose de más a las particularidades de los datos de entrenamiento y disminuyendo su capacidad de generalización y adaptación a nuevos conjuntos de datos. Para mitigar este problema, se lleva a cabo una evaluación exhaustiva del modelo utilizando métricas adicionales como la precisión, el recall y la puntuación F1. Estas métricas ayudan a asegurar que el modelo no solo es preciso, sino también robusto y fiable para aplicaciones prácticas, proporcionando así una base sólida para decisiones informadas y basadas en datos.

4 Propuesta: Mejora en la minería de decisiones aplicada en ScreenRPA mediante la identificación de reglas no deterministas combinando árboles de decisión

La propuesta central de este trabajo fin de grado es mejorar el enfoque actual en la minería de decisiones llevado a cabo por la plataforma de ScreenRPA. El nuevo enfoque se basa en realizar un análisis más profundo y detallado de las reglas descubiertas en el proceso, consiguiendo identificar reglas superpuestas (overlapping rules).

Las reglas superpuestas son aquellas que no son deterministas en las decisiones tomadas durante un proceso. Es decir, en casos diferentes, con las mismas circunstancias a priori (características identificadas), se toman decisiones diferentes. Esto generalmente ocurre debido a la falta de caracterización de ciertos factores que influyen en las decisiones, aunque en algunos casos, también puede deberse a la propia naturaleza entrópica y azarosa de ciertos procesos. Buscando identificar estas reglas, la nueva perspectiva persigue proporcionar una mayor claridad y comprensión en situaciones de incertidumbre durante los análisis de los procesos.

Esta mejora se inspira en la metodología de descubrimiento de reglas superpuestas en logs de eventos tradicionales, desarrollada por Mannhardt et al [22]. Sin embargo, en lugar de aplicar esta metodología directamente a los logs de eventos tradicionales, la propuesta se adapta al contexto de los datasets obtenidos de logs de interacción de usuarios (UI logs) en el ámbito de la Automatización de Procesos Robóticos (RPA). Al adaptar esta metodología a los UI logs, se anticipa que se obtendrá un análisis más detallado y preciso de las interacciones del usuario, permitiendo identificar patrones y relaciones que podrían ser menos evidentes con los métodos tradicionales.

4.1 Proceso de mejora propuesto

4.1.1 Entrenamiento del modelo sobre el UI log completo:

Para lograr esta mejora, el primer paso es entrenar un modelo sobre el UI log completo. Utilizamos el dataset etiquetado generado previamente como base para este entrenamiento. Continuaremos utilizando árboles de decisión debido a su fácil interpretación tanto por humanos como por máquinas. Este tipo de modelo visualiza las reglas de clasificación en forma de un árbol estructurado, donde los nodos del árbol representan los diferentes atributos del UI log (columnas del dataset), y las ramas del árbol indican condiciones claras y no superpuestas evaluadas sobre estos nodos.

4.1.2 Algoritmo para la Extracción de Reglas Superpuestas

Sobre este modelo se desarrolla un algoritmo para analizar la precisión del árbol de decisión y extraer las reglas superpuestas si las hubiese. Para ello se siguen los pasos descritos en el pseudocódigo del Algoritmo 2.

```
Input   : Flattened dataset, columna_objetivo, min_samples_split, merge_ratio_e
Output  : Diccionario paths_dict
1  X, y ← dividirDatos(df, columna_objetivo)
2  X_train, X_test, y_train, y_test ← dividirTrainTest(X, y, 0.2)
3  tree ← entrenarArbolDecision(X_train, y_train)
4  y_pred, accuracy ← predecirYCalcularPrecision(tree, X_test, y_test)
5
6  if accuracy == 1.0 then
7      guardarArbolCompleto(tree, X)
8  else
9      guardarArbolErroneo(tree, X)
10     misclassified_data, misclassified_labels, misclassified_preds ←
identificarErrores(X_test, y_test, y_pred)
11
12     for label in etiquetasUnicas(misclassified_labels) do
13         subset, subset_target, subset_preds ← extraerSubconjunto(misclassified_data,
misclassified_labels, misclassified_preds, label)
14         t_prima, freq_t_prima ← etiquetaMasFrecuente(subset_preds)
15
16         if cumpleCriterios(subset, min_samples_split, freq_t_prima, merge_ratio_e)
then
17             subset_paths ← extraerRutasDecision(tree, subset)
18             guardarSubconjunto(subset, subset_target, subset_preds, subset_paths,
label)
19             actualizarPathsDict(paths_dict, subset, t_prima)
20
21             sub_tree ← entrenarNuevoArbol(subset, subset_target)
22             guardarSubArbol(sub_tree, subset, i)
23
24             paths_dict ← actualizarDiccionario(paths_dict, subset, t_prima)
25         end
26     end
27 end
28 return paths_dict
```

Algoritmo 2: Extracción de reglas no deterministas en puntos de decisión.

La propuesta se centra en el algoritmo para la extracción de reglas no deterministas en puntos de decisión (Algoritmo 2). Este algoritmo toma como entrada el dataset aplanado etiquetado, el nombre de la variable objetivo y los criterios para el entrenamiento de los árboles de clasificación. Como salida, se espera un diccionario con las reglas descubiertas para cada clase del punto de decisión.

En primer lugar, el conjunto de datos de entrada se divide en un conjunto para entrenamiento y otro para pruebas. Con estos datos se entrena y evalúa el primer árbol de clasificación (líneas 1-6). Si la precisión del modelo es del 100%, se exporta y guarda la estructura completa del árbol. En caso contrario, se exporta y guarda la estructura del árbol con las clasificaciones erróneas (líneas 6-9).

En la línea 10, se identifican las muestras mal clasificadas y se extraen sus etiquetas verdaderas y predichas. Para cada etiqueta mal clasificada, se extrae el subconjunto de datos correspondiente y se determina la etiqueta predicha más frecuente dentro de este subconjunto (líneas 12-14). Si el subconjunto cumple con ciertos criterios, se extraen las rutas de decisión y se almacenan junto con los datos de verdad y predicciones (líneas 16-18).

En la línea 19, se actualiza el diccionario de rutas con las reglas superpuestas y se elimina la regla asociada de la clase de predicción. A continuación, se entrena un nuevo árbol de decisión con el subconjunto de muestras mal clasificadas para refinar las reglas de clasificación (líneas 21-22). Finalmente, en la línea 24, se actualiza el diccionario de rutas con metadatos sobre las muestras mal clasificadas y sus relaciones, proporcionando una visión comprensiva de los patrones de clasificación errónea y las reglas de decisión asociadas.

Este enfoque avanzado permite mejorar la interpretabilidad y la calidad del modelo de decisión, proporcionando información adicional en casos de incertidumbre y mejorando la comprensión general del proceso de toma de decisiones. De esta manera, este trabajo presenta una mejora significativa en el análisis y explicación de decisiones basadas en logs de interfaz de usuario, integrando técnicas avanzadas de aprendizaje automático y extracción de reglas superpuestas. La combinación del enfoque tradicional de árboles de decisión con la identificación de reglas superpuestas proporciona un marco robusto y transparente para entender y optimizar los procesos de negocio capturados por ScreenRPA.

4.2 Ejemplo de aplicación del proceso propuesto

Para llevar a cabo la experimentación del nuevo enfoque de minería de decisiones, es fundamental explicar el contexto del proceso en estudio. Nos situamos en el caso de un usuario comercial de una empresa proveedora de cemento. Una empresa constructora solicita sus servicios y negocian la compra de una cantidad específica de toneladas de material.

Para procesar el pago, dar la venta por finalizada y proceder con el envío, el comercial necesita confirmar que la oferta ha sido aceptada y que el pago se ha realizado correctamente. Para ello, el comercial se basa en tres características "virtuales" medidas durante el proceso: "attachment_A1" (si se adjunta el contrato aceptando el pago), "attachment_A2" (si se adjunta el justificante de pago) y "status_label_B1" (si en la plataforma de facturación de la empresa aparece el pago como realizado). Estas propiedades se recopilan durante dos actividades previas a la toma de decisión (Figura 6). La actividad A es la lectura del correo electrónico por parte del comercial, donde la constructora detalla si la oferta ha sido aceptada y si el pago se ha realizado (Figura 7). La actividad B, que se realiza tras verificar el correo electrónico, consiste en acceder a la plataforma de facturación de la cementera y verificar si el pago ha sido registrado correctamente (Figura 8).

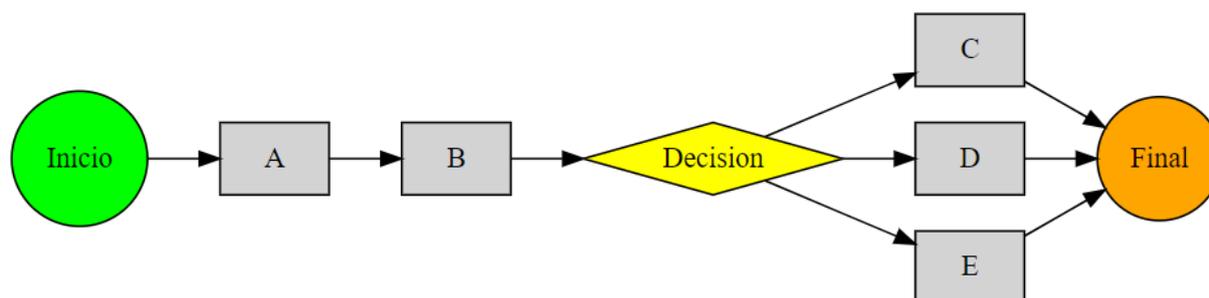


Figura 7. Proceso de registro de venta de cemento.

En la Figura 7 se muestra un correo de notificación de aceptación de la oferta y confirmación del pago (actividad A), donde las características "attachment_A1" y "attachment_A2" son ambos de valor 1. En la Figura 8 se muestra una captura de pantalla del programa de facturación de la empresa (actividad B) donde la característica "status_label_B1" (señalada en la imagen) tiene valor 1.

Aceptación de oferta y realización de pago. Recibidos x

a Alejandro Inglés Martínez
para mí ▾

Traductor español ▾ español ▾ Traducir correo ←

→ Reenviar correo t

10:42 (hace 0 minutos) ☆ 😊

Buenas tardes Alejandro,
hemos comprobado formalmente la oferta y hemos decidido aceptarla. Debido a la naturaleza de la urgencia de la compra, ya hemos tramitado el pago.
Te adjunto el justificante bancario.
Muchas gracias.

2 archivos adjuntos • Analizado por Gmail ⓘ

Justificante-pag... Contrato aceptac...

Figura 8. Notificación por correo de oferta aceptada y pago realizado.

Ventas Cómo gestionar tus gastos

Gastos Gestionar gastos periódicos Validar gastos periódicos

Todos los tipos ▾ Todos los estados ▾ Todos los métodos ▾

2022 ▾ Trimestre 4 ▾

Buscar por ▾ Mínimo 3 caracteres 🔍

🔍 Ocultar el resumen Estado de pago Evolución

IVA deducible	Base	Pagado	Pendientes	Vencido
16.830,95 €	515.523,38 €	486.540,11 €	14.975,93 €	5523,07 €
Retención	Total	<div style="display: flex; justify-content: space-between;"> 9% 9% 5% </div>		
-28.266,52 €	504.087,81 €			

Mostrando 1 - 20 de 548 resultados

ESTADO	NÚMERO	CONTACTO	EMISIÓN	VENCIMIENTO	PAGO	CONCEPTO	BASE (€)	TOTAL (€)
PAGADO	12249	B. Sabadell	28/12/2022	28/12/2022	28/12/2022	Comisiones Sendinblue - US\$	65,74	65,74
PAGADO	12750 / 1785186	Programa A	28/12/2022	28/12/2022	28/12/2022	Empresa X	28	28
VENCIDO	12751 / YI-MAN-00122	Digital	28/12/2022	28/12/2022		Servicios	840	1016,4
PENDIENTE	12754 / 2022-12	Joe Doe	28/12/2022	4/1/2023		December	5608,33	5944,83
PENDIENTE	12720 / B22120208	GYM SERVICES ONLINE S.L.	25/12/2022	24/1/2023		Gym	108,33	131,08

Figura 9. Pago realizado registrado en la plataforma de facturación

En este contexto, se recopilan numerosos casos y, tras la aplicación del pipeline de ScreenRPA, se obtiene finalmente el dataset aplanado y etiquetado del que partimos para la experimentación. Este dataset tiene una dimensionalidad de 42 filas y cuatro atributos, donde cada fila corresponde a una instancia distinta de ejecución del proceso. Como se mencionó anteriormente, de estos atributos, tres son características del proceso registradas antes de tomar la decisión, mientras que la cuarta variable es la variable objetivo o etiqueta a predecir, denominada "variant", que representa el curso de acción que decide tomar el usuario tras su decisión.

	attachment_A1	attachment_A2	status_label_B1	variant
0	1	1	1	1
1	1	0	0	1
2	0	0	0	2
3	0	0	1	2
4	0	1	0	3

Figura 10. Dataset etiquetado utilizado para el ejemplo de experimentación

Después de un mínimo preprocesamiento, los datos se dividen en un conjunto de entrenamiento, que comprende el 80% del total, y un conjunto de prueba, que constituye el 20% restante. A continuación, se procede a aplicar un primer modelo de árbol de clasificación C45 a estos datos.

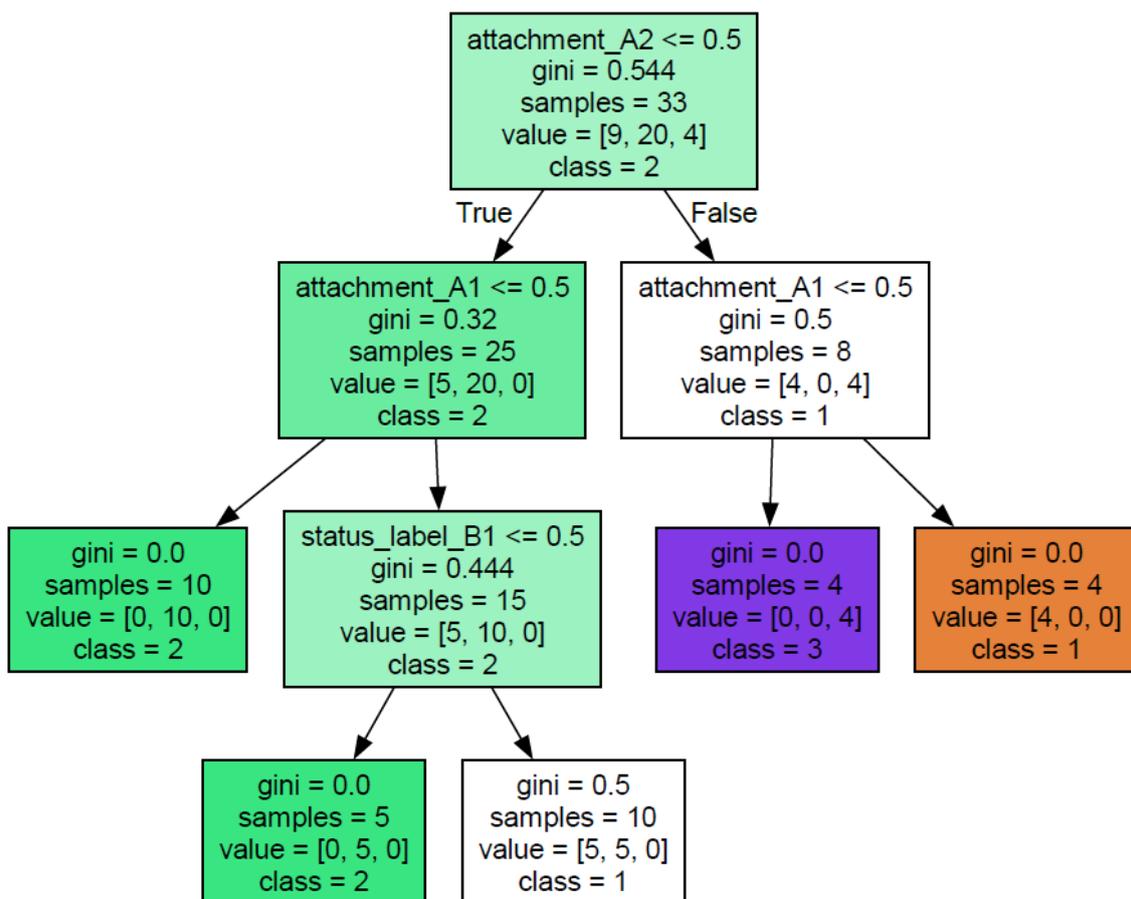


Figura 11. Representación gráfica del primer árbol de clasificación

Una vez que el modelo ha sido entrenado y evaluado, se recorren las ramas del árbol para extraer las reglas descubiertas, las cuales se almacenan en un diccionario cuyas claves son las etiquetas correspondientes a las hojas de las ramas. Si el rendimiento del modelo no alcanza el 100%, para cada etiqueta con muestras mal clasificadas (clases 1 y 2), las instancias mal clasificadas se someten a un nuevo ciclo de clasificación utilizando un segundo árbol de decisión, esta vez entrenado exclusivamente con los datos que fueron erróneamente clasificados por el primer árbol. Al igual que en el primer modelo, se revisa el árbol y se extraen todas las reglas identificadas.

La figura 12 presenta un resumen del conocimiento extraído tras la aplicación de todo este proceso. El primer modelo de árbol de decisión consigue una precisión de 88.89%. Aunque es un resultado relativamente alto, implica que un 11.11% de las muestras son clasificadas erróneamente, motivando la necesidad de un análisis más detallado y un segundo proceso de clasificación para gestionar estos errores.

En este resumen (figura 12) se observa que de las 3 vías posibles tras la decisión, la única clase en la que se descubren reglas 100% deterministas es la clase 3, mientras que en las dos clases restantes, obtenemos una regla superpuestas entre ambas. Esta regla provoca una confusión en el modelo y por lo tanto también en los criterios de clasificación, lo que se manifiesta en la forma en que algunas muestras de la clase 2 son incorrectamente clasificadas como pertenecientes a la clase 1, y viceversa (Figura 13).

Para entender el significado de las reglas descubiertas por el algoritmo, se describe el contexto de estas a continuación:

En la clase 3, el comercial cuando verifica el correo del cliente tiene adjunto el justificante bancario (por lo que el pago se ha realizado con certeza) pero no se encuentra adjunto el contrato de la oferta firmado. Por lo que el usuario decide ponerse en contacto con su empresa para verificar que el contrato firmado no se tuviera con anterioridad, antes de ponerse en contacto con el cliente para pedirle este.

En la clase 1, el comercial decide tramitar el envío ya que el cliente ha realizado el pago correctamente; tiene el justificante adjunto y además en la plataforma de facturación se marca el pago como realizado.

En la clase 2, el comercial decide ponerse en contacto con la constructora por diversos motivos: porque no se ha aceptado la oferta y, por lo tanto, no tiene adjunto ni el justificante de pago ni el contrato de aceptación; o bien, porque aunque se haya aceptado la oferta, el cliente no le envía el justificante de pago adjunto y tampoco aparece el pago como realizado en la plataforma de facturación.

Sin embargo, existen causas por las que el usuario decide tomar diferentes caminos aun teniendo las mismas características registradas. Se ha aceptado la oferta (contrato de aceptación adjunto en el correo), la plataforma de facturación marca el pago como realizado, pero, sin embargo, no está el justificante de pago adjunto en el correo. ¿Por qué el usuario decide tomar diferentes caminos en estas circunstancias? Sin aplicar este nuevo enfoque, no obtendríamos ninguna información y estos casos resultarían totalmente opacos y escaparían a nuestro razonamiento y capacidad predictora. Sin embargo, gracias a la propuesta, conseguimos arrojar luz sobre estas situaciones. Aunque no somos capaces de predecir de forma segura la decisión a tomar, sí tenemos algo de conocimiento donde antes solo había incertidumbre.

En nuestro caso, este solapamiento se puede justificar debido a que el usuario decide ponerse en contacto con el cliente (caso 2) para pedirle el justificante del pago, ya que es la prueba inequívoca de que el pago se ha realizado correctamente y debido a que el usuario en las mismas circunstancias decide tramitar el envío (clase 1), ya que no tiene adjunto el justificante de forma electrónica pero sí lo posee de forma física. Lo que es suficiente al comercial para poder cerrar el proceso y decidir esa acción.

En el caso de este proceso, el hecho de que exista una regla superpuesta entre clases es la falta de caracterización de la información de si el usuario recibe el justificante de forma física o no.

```
=====
Tree Accuracy: 88.89%

=====
Variants with 100% certainty:
=====
Classes whose guards have a 100% certainty: {3}
Class 3:
- (certainty 100%) attachment_A2 > 0.50 and attachment_A1 <= 0.50

=====
Variants with uncertainty:
=====
Overlapping rules: the class 2 (ground_truth) is missclassified as the class 1 (prediction)

Class 2:
- (precision 100%) attachment_A2 <= 0.50 and attachment_A1 <= 0.50
- (precision 100%) attachment_A2 <= 0.50 and attachment_A1 > 0.50 and status_label_B1 <= 0.50
- (overlapped) attachment_A2 <= 0.50 and attachment_A1 > 0.50 and status_label_B1 > 0.50
Class 1:
- (precision 100%) attachment_A2 > 0.50 and attachment_A1 > 0.50
- (overlapped) attachment_A2 <= 0.50 and attachment_A1 > 0.50 and status_label_B1 > 0.50
```

Figura 12. Resumen del conocimiento general extraído durante el proceso.

Finalmente, para visualizar los resultados del análisis de forma gráfica se presentan en forma de matriz de confusión, detallando el desempeño del modelo.

```
y_test
31  2
41  2
4   3
15  1
28  1
19  2
16  2
7   1
18  3
Name: variant, dtype: int64
y_pred
[2 1 3 1 1 2 2 1 3]
```

Figura 13. Valores predichos vs valores reales

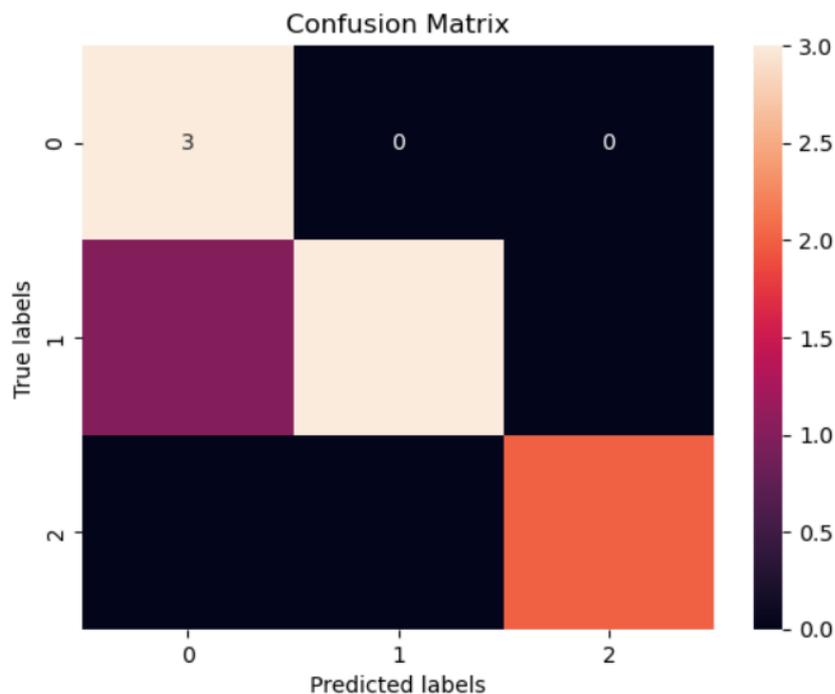


Figura 14. Matriz de confusión de resultados del clasificador

5 Conceptos y materias aplicadas

En esta sección se detallarán los conceptos teóricos que se ponen en práctica a la hora de definir el ciclo de vida del software y el desarrollo del proyecto, además de los trabajos relacionados que se han encontrado al respecto de este.

5.1 Investigación relacionada

Se ha identificado que el estudio con título “Robotic process automation using process mining — A systematic literature review” [23], es la revisión más reciente con respecto al área de la minería de procesos aplicada al ámbito del RPA, que es el contexto de este proyecto. Por tanto, analizando esta revisión se han conseguido extraer los trabajos más destacados en el ámbito según la SLR. Fijándonos en la selección de “papers” que se puede ver en la tabla 2, se puede apreciar que los trabajos más representativos, aparte de los correspondientes a los representantes del grupo ES3 [26], y, por tanto, lo relacionado con el framework sobre el que se construirá la propuesta de este TFG, son los frameworks de Agostinelly [24] y Leno [25]. Se han seleccionado y examinado los estudios que describen más extensamente los frameworks de Leno y Agostinelli, comparando estos con el del grupo ES3.

Comparativa entre Robidium y ScreenRPA

El enfoque de Robidium se diferencia del de ScreenRPA en diferentes puntos. Desde el punto de vista de la extracción de datos, Robidium exclusivamente extrae información de registros de interfaz de usuario provenientes de rutinas de transferencia de datos, es decir, rutinas que consisten en

operaciones de copiar, cortar, pegar y editar, sin incluir la capacidad de capturar información a partir de capturas de pantallas.

En cuanto al análisis del proceso, ScreenRPA incorpora la minería de decisiones, análisis al que no llega Robidium. Además, Leno implementa la automatización mediante la generación automática de scripts RPA ejecutables necesarios para automatizar la implementación de rutinas por parte de los robots software en un sistema informático de destino, lo que optimiza las rutinas repetitivas sin necesidad de intervención manual en cada paso.

Comparativa entre SmartRPA y ScreenRPA

SmartRPA se distingue por su enfoque en la extracción de datos de ScreenRPA en dos puntos. En primer lugar, extrae la información directamente de la interacción del usuario con el sistema, limitándose a eventos de teclado y clics, sin incorporar la extracción de características de capturas de pantalla. En segundo lugar, esta información la obtiene a partir de datos segmentados, característica diferencial con las otras dos plataformas.

En cuanto al análisis del proceso, la propuesta de Agostinelli no aborda la minería de decisiones, se centra exclusivamente en la minería de procesos. Además, al igual que Robidium, SmartRPA facilita la automatización a través de la generación automática de scripts RPA ejecutables, necesarios para automatizar rutinas en sistemas informáticos de destino.

Conclusión

En resumen, el enfoque de ScreenRPA, desarrollado por el grupo ES3, incorpora la capacidad adicional de extraer información de capturas de pantalla, una funcionalidad que no poseen los otros dos frameworks. Esto es especialmente útil en entornos virtualizados donde no es posible acceder a la estructura del sistema operativo y solo se pueden utilizar imágenes, fortaleciendo el enfoque de la plataforma, ya que los demás frameworks se limitan solo a las características extraídas a partir de los datos de eventos de interacciones del usuario con el sistema (UI logs).

La otra diferencia principal con respecto a los otros frameworks, es la aplicación de minería de decisiones, con lo que consigue no solo descubrir el proceso de negocio que modela el flujo de tareas sino también identificar las decisiones que se toman y decidir entre las distintas variantes del proceso.

En general, ScreenRPA desarrolla un análisis más profundo de los datos y procesos, llegando a aplicar técnicas de minería de decisiones, pero, sin embargo, no llega a implementar la generación automática de scripts ni a poner en marcha robots software capaces de realizar tareas automatizadas en los sistemas. Lo que supone que en este momento ScreenRPA es capaz de analizar flujos de información y extraer conocimientos en profundidad de ellos, pero no logra todavía la automatización real de las tareas.

Para ampliar esta investigación y llegar más allá, el planteamiento del grupo ES3, y sobre lo que se centra el enfoque de este apartado del trabajo, es valorar si se posee información suficiente para identificar las razones detrás de cada decisión, y no solo la decisión en sí.

5.2 Metodologías

En el desarrollo de un producto software, numerosas etapas, actividades y roles son cruciales para la finalización exitosa del proyecto. Desde la planificación inicial hasta la implementación del software, diversos factores influyen en el éxito o fracaso de este. En años recientes, se han definido diversas

metodologías para abordar proyectos, siendo la elección de la metodología adecuada clave para su culminación exitosa.

Una metodología es esencialmente un conjunto integrado de técnicas y métodos que se aplican de manera coherente a lo largo de las actividades del ciclo de vida de un proyecto de desarrollo, ofreciendo un proceso detallado y completo para el desarrollo de software.

Existen principalmente dos categorías de metodologías. Las primeras definidas son las metodologías tradicionales, que enfatizan la importancia de la documentación a lo largo del ciclo de vida del proyecto. En contraste, las metodologías ágiles buscan mejorar el rendimiento y reducir el énfasis en la documentación en favor de una comunicación continua y efectiva entre el equipo de desarrollo y el cliente, así como la adaptabilidad rápida a los cambios y la detección temprana de errores.

En este proyecto, nos enfocaremos en las metodologías ágiles, que serán las empleadas a lo largo del desarrollo de este. A continuación, detallaremos las características de estas metodologías, para concluir con las razones por las que resultan ideales para el desarrollo de este proyecto y cómo se definirá su ciclo de vida.

5.2.1 Metodologías ágiles

Ante la posibilidad de cambios significativos en las necesidades del cliente desde la planificación hasta la entrega final de un proyecto, se hace imprescindible adoptar metodologías que permitan una rápida adaptación a estos cambios imprevistos. De esta necesidad emergen las metodologías ágiles, conocidas por su capacidad de ofrecer respuestas rápidas y efectivas ante modificaciones durante el desarrollo del proyecto.

En febrero de 2001, con la expansión del uso de computadores personales desde principios de los noventa y el aumento de las expectativas de los usuarios, diecisiete líderes de la industria del software se reunieron en Utah, Estados Unidos. Esta reunión, conocida como "The Agile Alliance", fue convocada para discutir mejoras sobre las metodologías existentes, que en ese momento eran predominantemente tradicionales y rígidas. En este encuentro, se acuñó el término "ágil" para describir un enfoque de desarrollo de software que enfatiza la calidad, la flexibilidad y la adaptabilidad, posponiendo decisiones hasta que sean absolutamente necesarias y adoptando una planificación que pueda adaptarse a cambios rápidos.

Como fruto de esta reunión, se formuló el "Manifiesto para el Desarrollo Ágil de Software" [29], que establece un marco basado en valores y principios orientados a mejorar la eficiencia en el desarrollo de software. Los cuatro valores fundamentales reflejados en este manifiesto son:

1. **Priorizar a los individuos y las interacciones** sobre los procesos y las herramientas.
2. **Valorar el software funcional** más que la documentación exhaustiva.
3. **Fomentar la colaboración con el cliente** más que la negociación contractual.
4. **Responder al cambio** por encima de seguir un plan estricto.

Estos valores enfatizan la importancia de los elementos del lado derecho sobre los tradicionales del lado izquierdo. Los principios subyacentes [27], [28] incluyen:

- Satisfacer al cliente mediante entregas tempranas y continuas de software valioso.
- Ser receptivos a los cambios en los requisitos, incluso en etapas avanzadas del desarrollo.
- Entregar software funcional frecuentemente, con ciclos que varían de unas pocas semanas a unos pocos meses.

- Promover la colaboración diaria entre los empresarios y desarrolladores durante el proyecto.
- Construir proyectos alrededor de personas motivadas, proporcionándoles el entorno y apoyo necesario y confiando en ellos para completar la tarea.
- Admitir que la conversación cara a cara es el método más eficaz de comunicación dentro del equipo de desarrollo.
- Mantener un ritmo sostenible de trabajo para todos los involucrados en el proyecto.
- Fomentar la atención continua a la excelencia técnica y al buen diseño.

Las metodologías ágiles permiten un desarrollo iterativo e incremental, donde cada iteración abarca todas las fases del ciclo de vida del software, permitiendo que los requisitos y soluciones evolucionen progresivamente según las necesidades emergentes del proyecto.

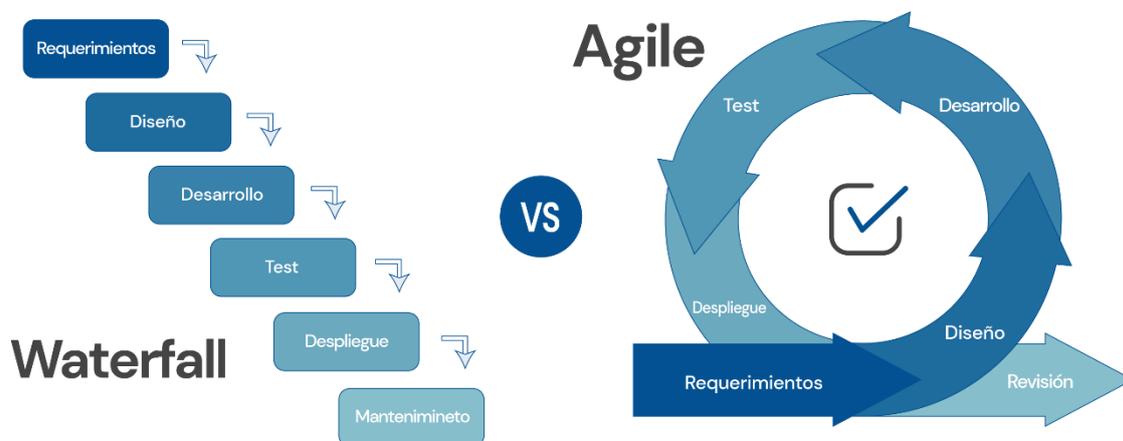


Figura 15. Ciclos de vida tradicionales vs ágiles

Este enfoque ágil, ilustrado por la comparación de ciclos de vida tradicionales frente a ágiles en la Figura 15, sienta las bases para metodologías como Scrum, Extreme Programming (XP), Dynamic Systems Development Method (DSDM), Crystal Clear y Agile Unified Process (AUP). A continuación, se explorará en detalle la metodología Scrum, seleccionada para este proyecto, basándonos en los principios establecidos en los artículos [28], [30].

5.2.1.1 Scrum

En el ámbito del desarrollo de software, las exigencias empresariales son dinámicas, lo que implica que los requisitos de un proyecto pueden cambiar durante su ciclo de vida. Esta realidad presenta desafíos adicionales para los equipos de desarrollo. Ante esto, la metodología ágil de Scrum emerge como una solución efectiva para optimizar los procesos de desarrollo de software. A continuación, exploraremos qué es Scrum y cómo surgió.

Origen del término Scrum

El término "Scrum" se originó en el rugby, donde describe una formación de ocho jugadores que trabajan en conjunto para avanzar con la pelota. Esta estrategia de equipo ha sido adaptada al desarrollo

de software para ilustrar cómo un equipo cohesionado puede trabajar hacia un objetivo común de manera efectiva y organizada, como se muestra en la Figura 16.



Figura 16. Origen del término “scrum” [31]

Definición

Scrum es un enfoque de desarrollo ágil que opta por un proceso incremental de desarrollo de software en lugar de una planificación y ejecución total. Prioriza la experiencia del equipo sobre la calidad de los procesos utilizados y favorece la ejecución paralela de fases en vez de secuencial. Según Coplien y Brooks, Scrum es particularmente eficaz para equipos reducidos, demostrando que grupos pequeños e independientes son más eficientes [32], [33]. Cada proyecto inicia con una fase de planificación donde se define la arquitectura y se designa un arquitecto de software. A lo largo del desarrollo, el equipo debe estar dispuesto a adaptar esta arquitectura, siguiendo un plan inicial establecido. El arquitecto software establece y mantiene la visión del proyecto durante todas las fases de desarrollo, respaldado por el patrón de Coplien "The Architect Controls Product" [32]. Tras la fase inicial, el desarrollo se divide en sprints, ciclos cortos que agregan incrementos al producto, típicamente durando de una a cuatro semanas, y enlazan con la metodología Design Sprint [34]. El proyecto culmina con una fase de cierre que finaliza el desarrollo del producto.

En este marco, las tareas identificadas se registran en una lista denominada "product backlog", que se actualiza después de cada sprint para reorganizar y priorizar tareas. Al inicio de cada sprint, las tareas asignadas forman el "sprint backlog" y se ejecutan durante el sprint. Durante este periodo, se prohíben cambios externos, permitiendo al equipo definir el alcance de trabajo de su sprint backlog. Cada sprint busca introducir nuevas funcionalidades valiosas para el usuario.

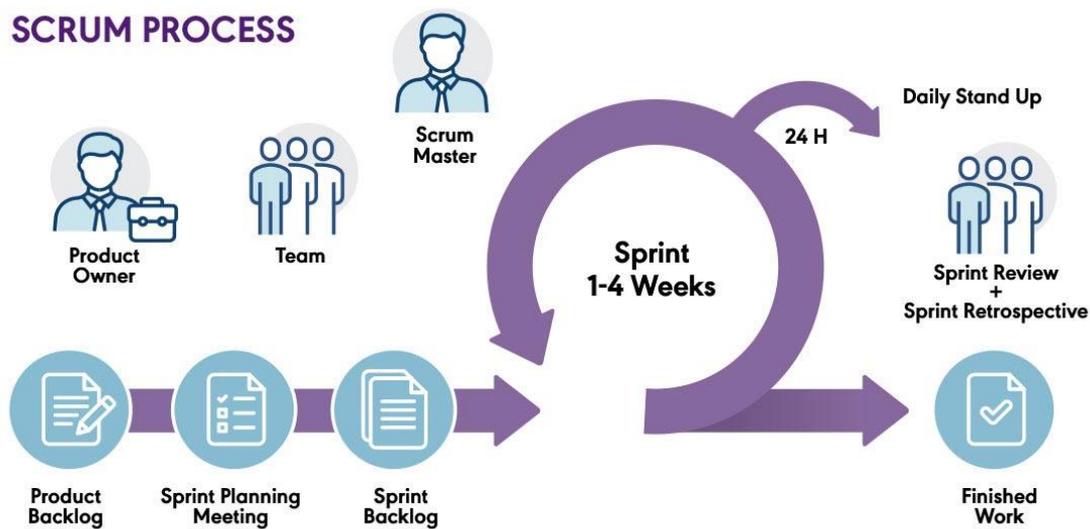


Figura 17. What is a standard scrum process?

Un sprint resulta en un incremento visible, utilizable y entregable del producto, incorporando nuevas interacciones del usuario con el sistema. Los sprints se desarrollan en tiempo real, manteniendo una fecha de finalización fija, aunque la funcionalidad entregada pueda ajustarse. Las reuniones diarias, conocidas como "Daily Scrum", facilitan la comunicación y resolución rápida de problemas para evitar retrasos prolongados.

Daily Scrum

Estas reuniones diarias fortalecen la cohesión del equipo, incluso integrando a colaboradores remotos, haciéndolos sentir parte activa del grupo. En ellas se revisa el trabajo realizado y por realizar, y se tratan los obstáculos encontrados, buscando ser breves y efectivas. Su duración se recomienda entre 15 y 30 minutos, concentrando los esfuerzos en abordar obstáculos y mantener al equipo informado y alineado con los objetivos.

Roles

En Scrum, identificamos al Scrum Master en cada reunión diaria, quien guía el proceso y mide el progreso del equipo hacia los objetivos del sprint, además de gestionar riesgos y cambios. Junto a él, el "Product Owner" maneja y prioriza el "product backlog", mientras que desarrolladores, usuarios y stakeholders interactúan para asegurar la relevancia y calidad del producto en desarrollo.

Sprint review y retrospective

Tras cada sprint, se realiza una revisión con todos los interesados para evaluar el progreso y planificar futuros sprints. Seguido por la retrospectiva del sprint, que busca mejorar el rendimiento del equipo mediante la reflexión sobre las prácticas actuales y compromisos futuros.

Sprint planning

Esta fase planifica el próximo sprint, determinando los incrementos necesarios y asignando tareas del "product backlog" según las prioridades y capacidades del equipo, aprovechando la experiencia acumulada en sprints anteriores para mejorar la estimación y ejecución de tareas.

5.2.1.2 Design Sprint y NDT

La optimización de la plataforma ha sido llevada a cabo con la colaboración del grupo de investigación ES3 [2], lo que implica la integración de sus metodologías de trabajo en el proyecto. El enfoque metodológico seleccionado para este desarrollo es el de Scrum, el cual es favorable frente a las metodologías más tradicionales debido a su énfasis en software funcional y la interacción continua con el cliente, minimizando el peso de las reglas rígidas y la documentación extensiva. Dentro de este contexto, se adoptarán otras metodologías compatibles con el marco ágil de Scrum.

En particular, se implementará la metodología Design Sprint para fomentar un comportamiento de equipo más dinámico y eficaz durante los sprints. Esta técnica, utilizada por el grupo de investigación ES3 para el prototipado rápido y la validación de ideas con usuarios finales, permite acelerar el desarrollo con iteraciones más cortas y rápidos lanzamientos de incrementos del producto. El Design Sprint fue desarrollado por Google Ventures [34] y está diseñado para abordar y solucionar preguntas críticas de negocio en solo cinco días, distribuidos de la siguiente manera:

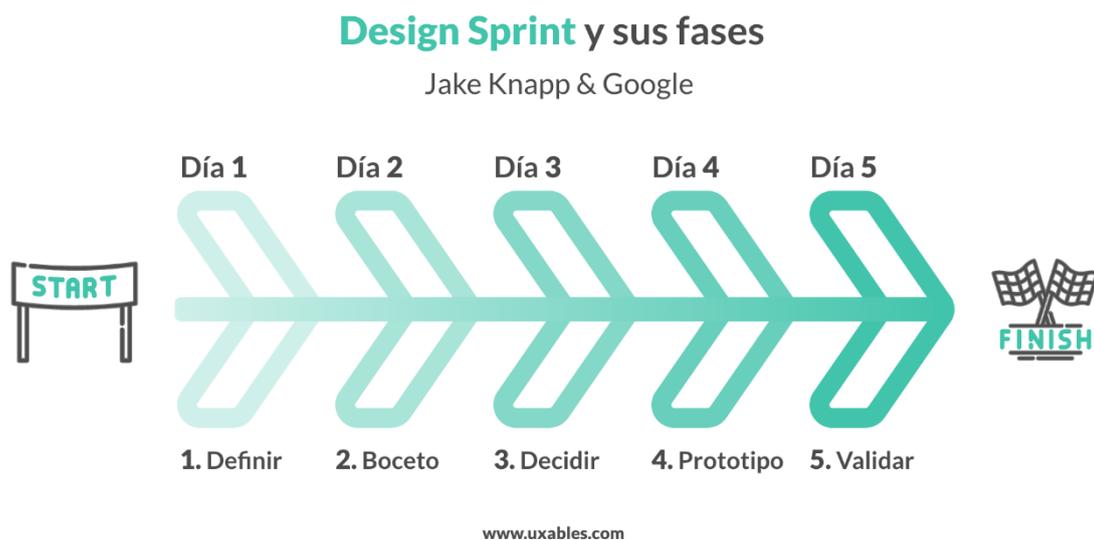


Figura 18. Fases de Design sprint

1. Lunes, **Comprender**: El equipo intercambia información relevante, definiendo objetivos, alcance y necesidades del usuario.
2. Martes, **Boceto**: Cada miembro del equipo desarrolla de forma independiente soluciones creativas para el problema identificado.
3. Miércoles, **Decidir**: Se evalúan todas las soluciones propuestas y se selecciona la más adecuada.
4. Jueves, **Prototipar**: Se construye un prototipo funcional con la mínima inversión de tiempo y recursos.
5. Viernes, **Validar**: Se valida el prototipo con usuarios reales para determinar su viabilidad y valor. Dependiendo del resultado, se continúa su desarrollo o se aprende de los errores, limitando el costo temporal al de una semana.

Esta metodología agiliza el proceso de desarrollo, reduciendo la duración de las iteraciones y comprimiendo los debates prolongados en un ciclo de trabajo intensivo y concentrado.

Adicionalmente, en la mejora de la plataforma RPA, se utilizará la NDT-Suite para el análisis, requisitos y pruebas del sistema. NDT (Navigational Development Techniques) se enmarca en el paradigma de la ingeniería guiada por modelos en la Ingeniería Web. Su elección se justifica por el alto nivel de automatización que ofrece y la eficiencia de las herramientas que compone la NDT-Suite. Además, se incluye el NDTQ-Framework, que proporciona una definición más estructurada y completa de los procesos de software [36], [37], reforzando así la metodología de desarrollo elegida para este proyecto.

6 Desarrollo del sistema

Siguiendo la metodología de NDT para el desarrollo del sistema, se respetará cada una de las fases de su ciclo de vida, enumerándose a continuación en orden, correspondiéndose con requisitos, análisis, diseño, implementación y pruebas.

6.1 Requisitos

En este apartado veremos tanto los objetivos principales que se persiguen con la mejora de la plataforma RPA como sus correspondientes requisitos.

6.1.1 Objetivos del sistema

6.1.1.1 Diagrama de objetivos

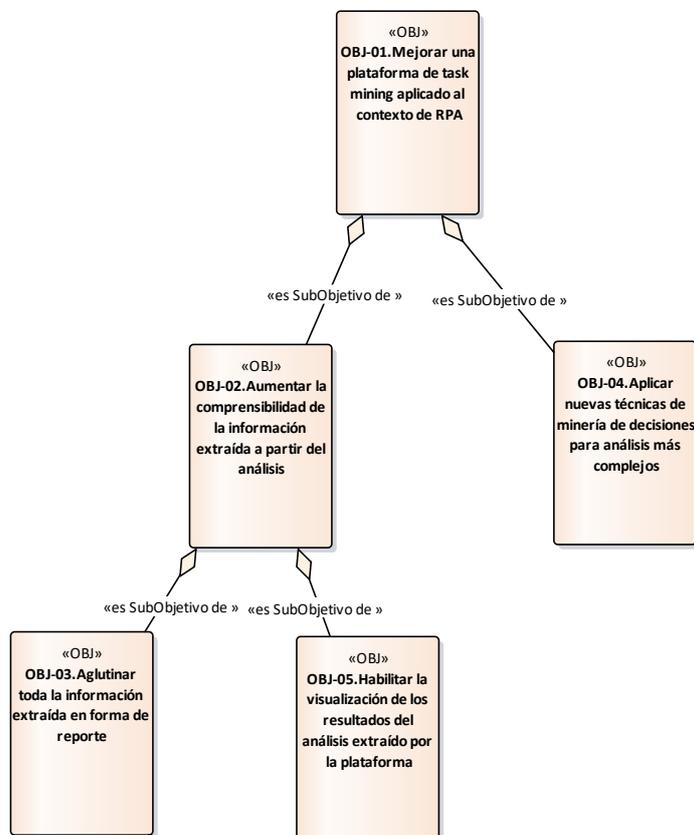


Figura 19: Diagrama de objetivos

6.1.1.2 Definición de objetivos

OBJ-01. Mejorar una plataforma de task mining aplicado al contexto de RPA		
Versión	1.0	25/03/2024 10:51:19
Autor	User	
Descripción	Mejorar una plataforma de servicios tecnológicos avanzados para la automatización de tareas.	
Subobjetivos	OBJ-02. Aumentar la comprensibilidad de la información extraída a partir del análisis OBJ-04. Aplicar nuevas técnicas de minería de decisiones para análisis más complejos	
Estado	Proposed	

Tabla 12.12. OBJ-01. Mejorar una plataforma de task mining aplicado al contexto de RPA

OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis		
Versión	1.0	25/03/2024 11:03:41
Autor	User	
Descripción	Aumentar la comprensibilidad de la información extraída a partir del análisis a través de nuevas funcionalidades que permiten la gestión y visualización de reportes y resultados por parte del usuario a través de la interfaz de la plataforma	
Subobjetivos	OBJ-03.Aglutinar toda la información extraída en forma de reporte OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma	
Estado	Proposed	

Tabla 1313: OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis

OBJ-03.Aglutinar toda la información extraída en forma de reporte		
Versión	1.0	25/03/2024 11:06:46
Autor	User	
Descripción	Aglutinar toda la información extraída en forma de reporte perfectamente legible, facilitando así a las personas la interpretación del conocimiento adquirido.	
Subobjetivos		
Estado	Proposed	

Tabla 1414: OBJ-03.Aglutinar toda la información extraída en forma de reporte

OBJ-04.Aplicar nuevas técnicas de minería de decisiones para análisis más complejos		
Versión	1.0	25/03/2024 11:11:44
Autor	User	
Descripción	Aplicar nuevas técnicas de minería de decisiones para análisis más complejos de los procesos a mejorar.	
Subobjetivos		
Estado	Proposed	

Tabla 15 15: OBJ-04.Aplicar nuevas técnicas de minería de decisiones para análisis más complejos

OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma		
Versión	1.0	26/03/2024 13:01:52
Autor	User	
Descripción	Habilitar la visualización de los resultados del análisis extraído por la plataforma pudiendo observar el conocimiento extraído en cada una de las fases del pipeline.	
Subobjetivos		
Estado	Proposed	

Tabla 1616: OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma

6.1.2 Catálogo de requisitos

A continuación se describen los distintos artefactos definidos por NDT para los requisitos, de manera que cada artefacto conllevará un código identificativo específico. Los códigos de cada artefacto se muestran a continuación.

Requisito	Nombre
Objetivo	OBJ-XX.Nombre
Requisito de almacenamiento	RA-XX.Nombre
Actor	AC-XX.Nombre
Requisito funcional	RF-XX.Nombre
Requisito no funcional	RNF-XX.Nombre

Tabla 1717. Nomenclatura de requisitos

6.1.2.1 Requisitos de almacenamiento de información

6.1.2.1.1 Diagrama de requisitos de almacenamiento

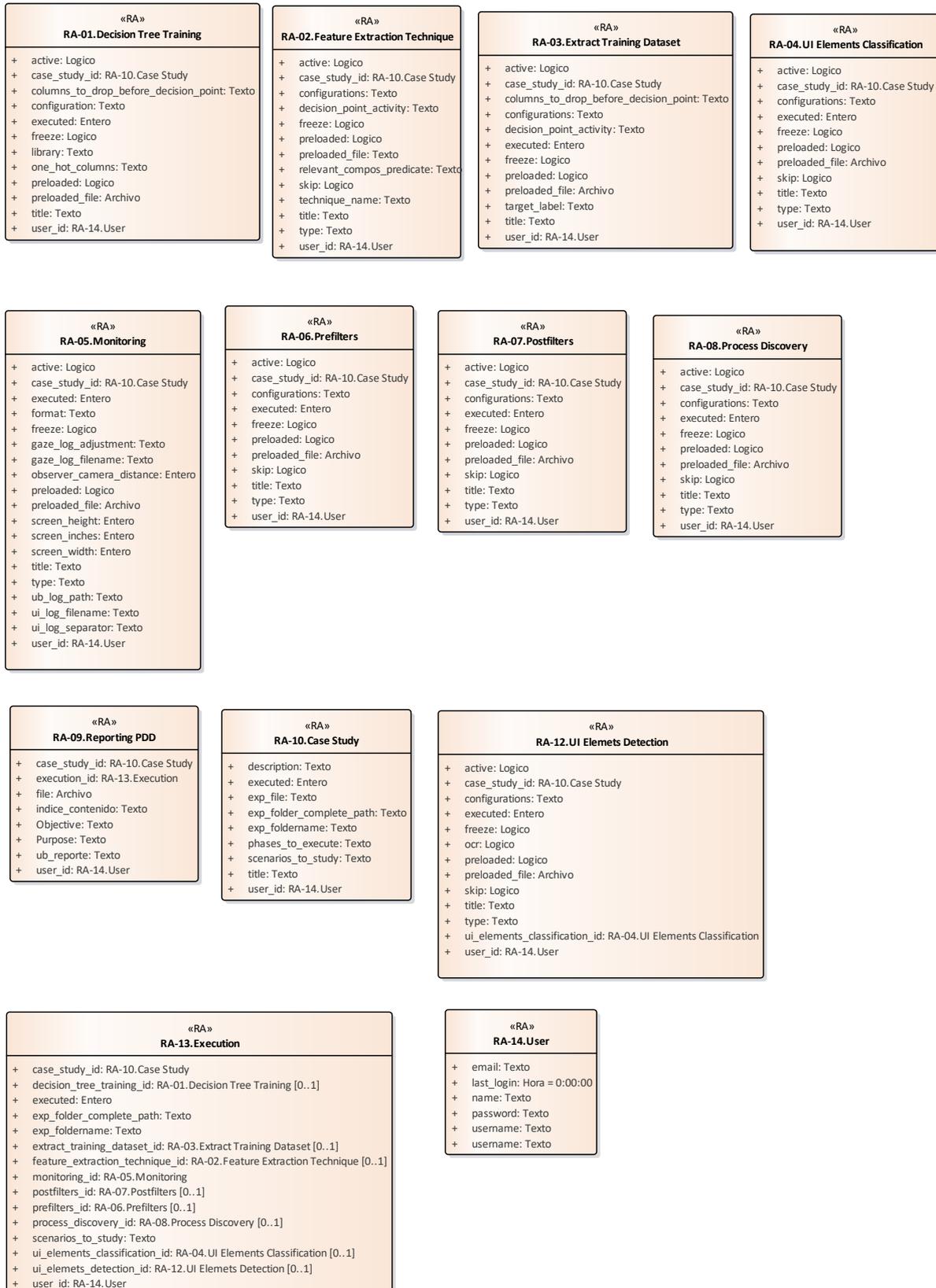


Figura 20: Diagrama de requisitos de almacenamiento

6.1.2.1.2 Definición de requisitos de almacenamiento

«RA» RA-01.Decision Tree Training	
Versión	1.0 05/04/2024 9:36:32
Autor	User
Objetivos asociados	RA-01.Decision Tree Training OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA OBJ-04.Aplicar nuevas técnicas de minería de decisiones para análisis más complejos OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma
Descripción	Configuración para el entrenamiento de un arbol de decisión en base a un log enriquecido de eventos
Atributos	<p>Nombre, Tipo y Descripción</p> <p>active: Logico <i>Descripción:</i> Si la configuración es la configuración a ejecutar al ejecutar el caso de estudio asociado</p> <hr/> <p>case_study_id: RA-10.Case Study <i>Descripción:</i> Caso de estudio al que pertenece la configuración</p> <hr/> <p>columns_to_drop_before_decision_point: Texto <i>Descripción:</i> Columnas a eliminar antes del punto de decisión</p> <hr/> <p>configuration: Texto <i>Descripción:</i> JSON para configuracion adicional</p> <hr/> <p>executed: Entero <i>Descripción:</i> Número de ejecuciones con la configuración</p> <hr/> <p>freeze: Logico <i>Descripción:</i> Si la configuración es de solo lectura</p> <hr/> <p>library: Texto <i>Descripción:</i> Librería a la que pertenecen los algoritmos utilizados en la configuración</p> <hr/> <p>one_hot_columns: Texto <i>Descripción:</i> Número de columnas tras transformar variables categóricas a numéricas</p> <hr/> <p>preloaded: Logico <i>Descripción:</i> Si los resultados de la fase están precargados</p> <hr/> <p>preloaded_file: Archivo <i>Descripción:</i> Resultados precargados de la fase</p> <hr/> <p>title: Texto <i>Descripción:</i> Título de la configuración de la fase</p> <hr/> <p>user_id: RA-14.User <i>Descripción:</i> Usuario que creó la configuración</p>
Estado	Proposed

Tabla 18 18: «RA» RA-01.Decision Tree Training

«RA» RA-02.Feature Extraction Technique	
Versión	1.0 05/04/2024 9:38:03
Autor	User
Objetivos asociados	RA-02.Feature Extraction Technique OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA
Descripción	Configuración de la técnica para extraer valor a un log de eventos
Atributos	<p>Nombre, Tipo y Descripción</p> <p>active: Logico <i>Descripción:</i> Si la configuración es la configuración a ejecutar al ejecutar el caso de estudio asociado</p> <p>case_study_id: RA-10.Case Study <i>Descripción:</i> Caso de estudio al que pertenece la configuración</p> <p>configurations: Texto <i>Descripción:</i> JSON para configuracion adicional</p> <p>decision_point_activity: Texto <i>Descripción:</i> Punto de decisión a estudiar</p> <p>freeze: Logico <i>Descripción:</i> Si la configuración es de solo lectura</p> <p>preloaded: Logico <i>Descripción:</i> Si los resultados de la fase están precargados</p> <p>preloaded_file: Texto <i>Descripción:</i> Resultados precargados de la fase</p> <p>relevant_compos_predicate: Texto <i>Descripción:</i> Condición para que un componente UI sea relevante</p> <p>skip: Logico <i>Descripción:</i> Si la fase se salta (no se ejecuta) al realizar la ejecución del caso de estudio</p> <p>technique_name: Texto <i>Descripción:</i> Nombre de la técnica configurada</p> <p>title: Texto <i>Descripción:</i> Título de la configuración de la fase</p> <p>type: Texto <i>Descripción:</i> Técnica para extraer conocimiento de un log de eventos</p> <p>user_id: RA-14.User <i>Descripción:</i> Usuario que creó la configuración</p>
Estado	Proposed

Tabla 19 19: «RA» RA-02.Feature Extraction Technique

«RA» RA-03.Extract Training Dataset		
Versión	1.0	05/04/2024 9:39:02
Autor	User	
Objetivos asociados	RA-03.Extract Training Dataset OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis	
Descripción	Configuración para la extracción de datos para entrenar el arbol de decision	
Atributos	Nombre, Tipo y Descripción	
	active: Logico	<i>Descripción:</i> Si la configuración es la configuración a ejecutar al ejecutar el caso de estudio asociado
	case_study_id: RA-10.Case Study	<i>Descripción:</i> Caso de estudio al que pertenece la configuración
	columns_to_drop_before_decision_point: Texto	<i>Descripción:</i> Columnas a eliminar antes del punto de decisión
	configurations: Texto	<i>Descripción:</i> JSON para configuracion adicional
	decision_point_activity: Texto	<i>Descripción:</i> Punto de decisión a estudiar
	executed: Entero	<i>Descripción:</i> Número de ejecuciones con la configuración
	freeze: Logico	<i>Descripción:</i> Si la configuración es de solo lectura
	preloaded: Logico	<i>Descripción:</i> Si los resultados de la fase están precargados
	preloaded_file: Archivo	<i>Descripción:</i> Resultados precargados de la fase
	target_label: Texto	<i>Descripción:</i> Variable objetivo a predecir
	title: Texto	<i>Descripción:</i> Título de la configuración de la fase
	user_id: RA-14.User	<i>Descripción:</i> Usuario que creó la configuración
	Estado	Proposed

Tabla 2020: «RA» RA-03.Extract Training Dataset

«RA» RA-04.UI Elements Classification		
Versión	1.0	05/04/2024 9:39:31
Autor	User	
Objetivos asociados	RA-04.UI Elements Classification OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA	

Descripción	Configuración para la técnica de clasificación de elementos UI
Atributos	<p>Nombre, Tipo y Descripción</p> <p>active: Logico <i>Descripción:</i> Si la configuración es la configuración a ejecutar al ejecutar el caso de estudio asociado</p> <hr/> <p>case_study_id: RA-10.Case Study <i>Descripción:</i> Caso de estudio al que pertenece la configuración</p> <hr/> <p>configurations: Texto <i>Descripción:</i> JSON para configuracion adicional</p> <hr/> <p>executed: Entero <i>Descripción:</i> Número de ejecuciones con la configuración</p> <hr/> <p>freeze: Logico <i>Descripción:</i> Si la configuración es de solo lectura</p> <hr/> <p>preloaded: Logico <i>Descripción:</i> Si los resultados de la fase están precargados</p> <hr/> <p>preloaded_file: Archivo <i>Descripción:</i> Resultados precargados de la fase</p> <hr/> <p>skip: Logico <i>Descripción:</i> Si la fase se salta (no se ejecuta) al realizar la ejecución del caso de estudio</p> <hr/> <p>title: Texto <i>Descripción:</i> Título de la configuración de la fase</p> <hr/> <p>type: Texto <i>Descripción:</i> Técnica usada para la clasificación de elementos UI</p> <hr/> <p>user_id: RA-14.User <i>Descripción:</i> Usuario que creó la configuración</p>

Tabla 2121: «RA» RA-04.UI Elements Classification

«RA» RA-05.Monitoring	
Versión	1.0 05/04/2024 9:39:57
Autor	User
Objetivos asociados	<p>RA-05.Monitoring</p> <p>OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA</p> <p>OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma</p> <p>OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis</p>
Descripción	Configuración sobre la información del log de eventos con las acciones de un usuario (keystrokes, clicks, etc.)
Atributos	<p>Nombre, Tipo y Descripción</p> <p>active: Logico <i>Descripción:</i> Configuración activa a ejecutar al ejecutar un caso de estudio</p> <hr/> <p>case_study_id: RA-10.Case Study <i>Descripción:</i> Caso de estudio a la que pertenece la configuración</p> <hr/> <p>executed: Entero <i>Descripción:</i> Número ejecuciones con esta configuración de la fase</p> <hr/> <p>format: Texto <i>Descripción:</i> Si esta en formato .mht u otro. Depende de la herramienta con la que se haya grabado el log</p> <hr/> <p>freeze: Logico <i>Descripción:</i> Si la configuración está en modo de solo lectura</p>

	gaze_log_adjustment: Texto <i>Descripción:</i> Ajuste al gaze log
	gaze_log_filename: Texto <i>Descripción:</i> Nombre del log de eyetracking
	observer_camera_distance: Entero <i>Descripción:</i> Distancia desde la cámara al usuario monitorizado
	preloaded: Logico <i>Descripción:</i> Si los resultados de la fase están precargados
	preloaded_file: Archivo <i>Descripción:</i> Resultados precargados de la fase
	screen_height: Entero <i>Descripción:</i> Altura de la pantalla monitorizada
	screen_inches: Entero <i>Descripción:</i> Pulgadas de la pantalla monitorizada
	screen_width: Entero <i>Descripción:</i> Ancho de la pantalla monitorizada
	title: Texto <i>Descripción:</i> Título de la configuración de la fase
	type: Texto <i>Descripción:</i> Herramienta usada para monitorear una actividad
	ub_log_path: Texto <i>Descripción:</i> Ubicación en la jerarquía de carpeta donde se guarda el log
	ui_log_filename: Texto <i>Descripción:</i> Nombre del archivo (log) contenedor de las acciones del usuario
	ui_log_separator: Texto <i>Descripción:</i> Separador usado en el log (;,)
	user_id: RA-14.User <i>Descripción:</i> Usuario que creó la configuración
Estado	Proposed

Tabla 2222: «RA» RA-05.Monitoring

«RA» RA-06.Prefilters	
Versión	1.0 05/04/2024 9:40:09
Autor	User
Objetivos asociados	RA-06.Prefilters OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA
Descripción	Configuración para el filtrado de información del log de eventos
Atributos	Nombre, Tipo y Descripción
	active: Logico <i>Descripción:</i> Si la configuración es la configuración a ejecutar al ejecutar el caso de estudio asociado
	case_study_id: RA-10.Case Study <i>Descripción:</i> Caso de estudio al que pertenece la configuración
	configurations: Texto <i>Descripción:</i> JSON para configuración adicional
	executed: Entero <i>Descripción:</i> Número de ejecuciones con la configuración

	freeze: Logico <i>Descripción:</i> Si la configuración es de solo lectura
	preloaded: Logico <i>Descripción:</i> Si los resultados de la fase están precargados
	preloaded_file: Archivo <i>Descripción:</i> Resultados precargados de la fase
	skip: Logico <i>Descripción:</i> Si la fase se salta (no se ejecuta) al realizar la ejecución del caso de estudio
	title: Texto <i>Descripción:</i> Título de la configuración de la fase
	type: Texto <i>Descripción:</i> Herramienta usada para filtrar el contenido de un log
	user_id: RA-14.User <i>Descripción:</i> Usuario que creó la configuración
Estado	Proposed

Tabla 2323: «RA» RA-06.Prefilters

«RA» RA-07.Postfilters	
Versión	1.0 05/04/2024 9:40:23
Autor	User
Objetivos asociados	RA-07.Postfilters OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis
Descripción	Configuración para el filtrado de información del log de eventos enriquecido
Atributos	<p>Nombre, Tipo y Descripción</p> <p>active: Logico <i>Descripción:</i> Si la configuración es la configuración a ejecutar al ejecutar el caso de estudio asociado</p> <p>case_study_id: RA-10.Case Study <i>Descripción:</i> Caso de estudio al que pertenece la configuración</p> <p>configurations: Texto <i>Descripción:</i> JSON para configuracion adicional</p> <p>executed: Entero <i>Descripción:</i> Número de ejecuciones con la configuración</p> <p>freeze: Logico <i>Descripción:</i> Si la configuración es de solo lectura</p> <p>preloaded: Logico <i>Descripción:</i> Si los resultados de la fase están precargados</p> <p>preloaded_file: Archivo <i>Descripción:</i> Resultados precargados de la fase</p> <p>skip: Logico <i>Descripción:</i> Si la fase se salta (no se ejecuta) al realizar la ejecución del caso de estudio</p> <p>title: Texto <i>Descripción:</i> Título de la configuración de la fase</p> <p>type: Texto <i>Descripción:</i> Técnica para filtrar el contenido de un log enriquecido</p> <p>user_id: RA-14.User <i>Descripción:</i> Usuario que creó la configuración</p>

Estado	Proposed
--------	----------

Tabla 2424: «RA» RA-07.Postfilters

«RA» RA-08.Process Discovery	
Versión	1.0 05/04/2024 9:40:44
Autor	User
Objetivos asociados	RA-08.Process Discovery OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis
Descripción	Configuración para el descubrimiento de procesos BPM
Atributos	<p>Nombre, Tipo y Descripción</p> <p>active: Logico <i>Descripción:</i> Si la configuración es la configuración a ejecutar al ejecutar el caso de estudio asociado</p> <p>case_study_id: RA-10.Case Study <i>Descripción:</i> Caso de estudio al que pertenece la configuración</p> <p>configurations: Texto <i>Descripción:</i> JSON para configuración adicional</p> <p>executed: Entero <i>Descripción:</i> Número de ejecuciones con la configuración</p> <p>freeze: Logico <i>Descripción:</i> Si la configuración es de solo lectura</p> <p>preloaded: Logico <i>Descripción:</i> Si los resultados de la fase están precargados</p> <p>preloaded_file: Archivo <i>Descripción:</i> Resultados precargados de la fase</p> <p>skip: Logico <i>Descripción:</i> Si la fase se salta (no se ejecuta) al realizar la ejecución del caso de estudio</p> <p>title: Texto <i>Descripción:</i> Título de la configuración de la fase</p> <p>type: Texto <i>Descripción:</i> Técnica para descubrir el proceso de negocio</p> <p>user_id: RA-14.User <i>Descripción:</i> Usuario que creó la configuración</p>
Estado	Proposed

Tabla 25 25: «RA» RA-08.Process Discovery

«RA» RA-09.Reporting PDD	
Versión	1.0 05/04/2024 9:41:01
Autor	User
Objetivos asociados	RA-09.Reporting PDD OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis OBJ-03.Aglutinar toda la información extraída en forma de reporte
Descripción	Reporte de una ejecución de un caso de estudio

Atributos	Nombre, Tipo y Descripción
	case_study_id: RA-10.Case Study <i>Descripción:</i> Caso de estudio asociado a la ejecución que se reporta
	execution_id: RA-13.Execution <i>Descripción:</i> Ejecución cuyos resultados reporta el documento
	file: Archivo <i>Descripción:</i> Documento de reporte
	indice_contenido: Texto <i>Descripción:</i> Índice de contenido configurado por el usuario a mostrar en el reporte
	Objective: Texto <i>Descripción:</i> Objetivo del estudio
	Purpose: Texto <i>Descripción:</i> Propósito del estudio
	ub_reporte: Texto <i>Descripción:</i> Ubicación en la jerarquía de carpetas donde se almacena el reporte
	user_id: RA-14.User <i>Descripción:</i> Usuario que genera el reporte
Estado	Proposed

Tabla 2626: «RA» RA-09.Reporting PDD

«RA» RA-10.Case Study	
Versión	1.0 05/04/2024 9:41:31
Autor	User
Objetivos asociados	RA-10.Case Study OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis
Descripción	Caso de estudio
Atributos	Nombre, Tipo y Descripción
	description: Texto <i>Descripción:</i> Descripción del caso de estudio
	executed: Entero <i>Descripción:</i> Número de veces que ha sido ejecutado el caso de estudio
	exp_file: Texto <i>Descripción:</i> Archivo que contiene toda la información del caso de estudio
	exp_folder_complete_path: Texto <i>Descripción:</i> Ruta completa donde se almacena el contenido relacionado con el caso de estudio
	exp_foldername: Texto <i>Descripción:</i> Nombre del directorio donde se almacena el contenido relacionado con el caso de estudio
	phases_to_execute: Texto <i>Descripción:</i> Fases a ejecutar cuando se ejecute el caso de estudio
	scenarios_to_study: Texto <i>Descripción:</i> Escenarios del caso de estudio a ejecutar cuando se ejecute el caso de estudio
	title: Texto <i>Descripción:</i> Título del caso de estudio
	user_id: RA-14.User <i>Descripción:</i> Usuario que generó el caso del estudio
	Estado

Tabla 2727: «RA» RA-10.Case Study

«RA» RA-12.UI Elemets Detection		
Versión	1.0	05/04/2024 9:42:36
Autor	User	
Objetivos asociados	RA-12.UI Elements Detection OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma	
Descripción	Configuración para la técnica de detección de elementos UI	
Atributos	Nombre, Tipo y Descripción	
	active: Logico	<i>Descripción:</i> Si la configuración es la configuración a ejecutar al ejecutar el caso de estudio asociado
	case_study_id: RA-10.Case Study	<i>Descripción:</i> Caso de estudio al que pertenece la configuración
	configurations: Texto	<i>Descripción:</i> JSON para configuracion adicional
	executed: Entero	<i>Descripción:</i> Número de ejecuciones con la configuración
	freeze: Logico	<i>Descripción:</i> Si la configuración es de solo lectura
	ocr: Logico	<i>Descripción:</i> Si se aplica OCR
	preloaded: Logico	<i>Descripción:</i> Si los resultados de la fase están precargados
	preloaded_file: Archivo	<i>Descripción:</i> Resultados precargados de la fase
	skip: Logico	<i>Descripción:</i> Si la fase se salta (no se ejecuta) al realizar la ejecución del caso de estudio
	title: Texto	<i>Descripción:</i> Título de la configuración de la fase
	type: Texto	<i>Descripción:</i> Técnica usada para detectar elementos UI
	ui_elements_classification_id: RA-04.UI Elements Classification	<i>Descripción:</i> Técnica de clasificación de elementos UI asociada
	user_id: RA-14.User	<i>Descripción:</i> Usuario que creó la configuración
Estado	Proposed	

Tabla 2828: «RA» RA-12.UI Elements Detection

«RA» RA-13.Execution		
Versión	1.0	05/04/2024 9:43:05
Autor	User	
Objetivos asociados	RA-13.Execution OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis	

Descripción	Ejecucion de un caso de estudio
Atributos	<p>Nombre, Tipo y Descripción</p> <p>case_study_id: RA-10.Case Study <i>Descripción:</i> Caso de estudio ejecutado</p> <p>decision_tree_training_id: RA-01.Decision Tree Training <i>Descripción:</i> Configuración de la fase Decision Tree Training utilizada en la ejecución</p> <p>executed: Entero <i>Descripción:</i> Porcentaje que representa si la ejecucion se ha completado satisfactoriamente o en que porcentaje se ha quedado</p> <p>exp_folder_complete_path: Texto <i>Descripción:</i> Ruta completa donde se almacena el contenido relacionado con la ejecución</p> <p>exp_foldername: Texto <i>Descripción:</i> Nombre del directorio donde se almacena el contenido relacionado con la ejecución</p> <p>extract_training_dataset_id: RA-03.Extract Training Dataset <i>Descripción:</i> Configuración de la fase Extract Training Dataset utilizada en la ejecución</p> <p>feature_extraction_technique_id: RA-02.Feature Extraction Technique <i>Descripción:</i> Configuración de la fase Feature Extraction utilizada en la ejecución</p> <p>monitoring_id: RA-05.Monitoring <i>Descripción:</i> Configuración de la fase Log Processing utilizada en la ejecución</p> <p>postfilters_id: RA-07.Postfilters <i>Descripción:</i> Configuración de la fase Post-filtering utilizada en la ejecución</p> <p>prefilters_id: RA-06.Prefilters <i>Descripción:</i> Configuración de la fase Pre-filtering utilizada en la ejecución</p> <p>process_discovery_id: RA-08.Process Discovery <i>Descripción:</i> Configuración de la fase Process Discovery utilizada en la ejecución</p> <p>scenarios_to_study: Texto <i>Descripción:</i> Escenarios ejecutados del caso de estudio</p> <p>ui_elements_classification_id: RA-04.UI Elements Classification <i>Descripción:</i> Configuración de la etapa de clasificación de la fase UI Element Configuration utilizada en la ejecución</p> <p>ui_elemets_detection_id: RA-12.UI Elemets Detection <i>Descripción:</i> Configuración de la fase UI Element Detection utilizada en la ejecución</p> <p>user_id: RA-14.User <i>Descripción:</i> Uusuario que ha realizado la ejecución</p>
Estado	Proposed

Tabla 29 29: «RA» RA-13.Execution

«RA» RA-14.User	
Versión	1.0 05/04/2024 9:48:57
Autor	User
Objetivos asociados	<p>RA-14.User</p> <p>OBJ-03.Aglutinar toda la información extraída en forma de reporte</p> <p>OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis</p> <p>OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA</p> <p>OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma</p>
Descripción	Usuario del sistema
Atributos	Nombre, Tipo y Descripción

	email: Texto <i>Descripción:</i> Email del usuario
	last_login: Hora <i>Descripción:</i> Hora del último inicio de sesión del usuario
	name: Texto <i>Descripción:</i> Apodo del usuario
	password: Texto <i>Descripción:</i> Contraseña del usuario
	username: Texto <i>Descripción:</i> Apellidos del usuario
	username: Texto <i>Descripción:</i> Nombre del usuario
Estado	Proposed

Tabla 3030: «RA» RA-14.User

6.1.2.2 Definición de actores

6.1.2.2.1 Diagrama de actores



Figura 21: Diagrama de actores

6.1.2.2.2 Definición de actores del sistema

«AC» AC-01.Usuario			
Versión	1.0	25/03/2024 11:21:53	
Autor	User		
Conexiones	. Fuente	Estereotipo	Destino
	AC-01.Usuario		PS-36.Eliminar reporte de una ejecución

AC-01.Usuario		PS-12.Visualizar contenido de Modelos de procesos
AC-01.Usuario		PS-39.Crear reporte de una ejecución
AC-01.Usuario		RF-01.Ver detalle de una ejecución
AC-01.Usuario		PS-29.Descargar reporte
AC-01.Usuario		RF-09.Descargar resultado de la fase de una ejecución
AC-01.Usuario		PS-43.Ver detalle de una ejecución
AC-01.Usuario		PS-32.Ver configuración de reporte de una ejecución
AC-01.Usuario		PS-27.Descargar reporte
AC-01.Usuario		PS-09.Visualizar contenido de Modelos de decisión
AC-01.Usuario		PS-02.Visualizar contenido de Modelos de decisión
AC-01.Usuario		PS-44.Ver detalle de una ejecución
AC-01.Usuario		PS-06.Visualizar contenido de Modelos de decisión
AC-01.Usuario		PS-40.Crear reporte de una ejecución
AC-01.Usuario		PS-05.Visualizar contenido de Modelos de decisión
AC-01.Usuario		RF-08.Ver resultado de las fases de una ejecución
AC-01.Usuario		PS-21.Ver resultado de las fases de una ejecución
AC-01.Usuario		PS-11.Visualizar contenido de Modelos de procesos
AC-01.Usuario		PS-10.Visualizar contenido de Modelos de procesos

AC-01.Usuario		PS-04.Visualizar contenido de Modelos de decisión
AC-01.Usuario		PS-07.Visualizar contenido de Modelos de decisión
AC-01.Usuario		PS-33.Listar reportes de una ejecución
AC-01.Usuario		PS-26.Ver configuraciones de las fases de una ejecución
AC-01.Usuario		PS-25.Ver configuraciones de las fases de una ejecución
AC-01.Usuario		PS-18.Descargar resultado de la fase de una ejecución
AC-01.Usuario		PS-41.Crear reporte de una ejecución
AC-01.Usuario		RF-11.Visualizar contenido de Modelos de procesos
AC-01.Usuario		RF-06.Descargar reporte
AC-01.Usuario		PS-34.Listar reportes de una ejecución
AC-01.Usuario		PS-38.Eliminar reporte de una ejecución
AC-01.Usuario		PS-28.Descargar reporte
AC-01.Usuario		PS-45.Ver detalle de una ejecución
AC-01.Usuario	«Relationship»	AE-01.Usuario
AC-01.Usuario		RF-02.Crear reporte de una ejecución
AC-01.Usuario		PS-08.Visualizar contenido de Modelos de decisión
AC-01.Usuario		RF-10.Visualizar contenido de UI logs
AC-01.Usuario		PS-31.Ver configuración de reporte de una ejecución

AC-01.Usuario		PS-23.Ver resultado de las fases de una ejecución
AC-01.Usuario		PS-37.Eliminar reporte de una ejecución
AC-01.Usuario		PS-24.Ver configuraciones de las fases de una ejecución
AC-01.Usuario		RF-05.Ver configuración de reporte de una ejecución
AC-01.Usuario		PS-42.Crear reporte de una ejecución
AC-01.Usuario		PS-13.Visualizar contenido de UI logs
AC-01.Usuario		PS-22.Ver resultado de las fases de una ejecución
AC-01.Usuario		PS-17.Descargar resultado de la fase de una ejecución
AC-01.Usuario		PS-03.Visualizar contenido de Modelos de decisión
AC-01.Usuario		RF-07.Ver configuraciones de las fases de una ejecución
AC-01.Usuario		RF-12.Visualizar contenido de Modelos de decisión
AC-01.Usuario		RF-04.Listar reportes de una ejecución
AC-01.Usuario		PS-35.Listar reportes de una ejecución
AC-01.Usuario		PS-01.Visualizar contenido de Modelos de decisión
AC-01.Usuario		PS-30.Ver configuración de reporte de una ejecución
AC-01.Usuario		PS-20.Ver resultado de las fases de una ejecución
AC-01.Usuario		PS-19.Ver resultado de las fases de una ejecución

	AC-01.Usuario		PS-16.Descargar resultado de la fase de una ejecución
	AC-01.Usuario		PS-14.Visualizar contenido de UI logs
	AC-01.Usuario		PS-15.Visualizar contenido de UI logs
	AC-01.Usuario		RF-03.Eliminar reporte de una ejecución
	OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA	«Relationship»	AC-01.Usuario
	OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis	«Relationship»	AC-01.Usuario
	OBJ-03.Aglutinar toda la información extraída en forma de reporte	«Relationship»	AC-01.Usuario
	OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma	«Relationship»	AC-01.Usuario
Descripción	El usuario es un actor con capacidad para realizar ejecuciones de casos de estudios y gestionar sus reportes		
Estado	Proposed		

Tabla 3131: «AC» AC-01.Usuario

«AC» AC-02.Administrador			
Versión	1.0	25/03/2024 11:22:32	
Autor	User		
Conexiones	. Fuente	Estereotipo	Destino
	AC-02.Administrador	«Relationship»	AE-02.Administrador
	OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA	«Relationship»	AC-02.Administrador
	OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma	«Relationship»	AC-02.Administrador
	OBJ-03.Aglutinar toda la información extraída en forma de reporte	«Relationship»	AC-02.Administrador

	OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis	«Relationship»	AC-02.Administrador
Descripción	El administrador es un actor que consta de los permisos del usuario y además tiene la capacidad de configurar los parámetros y entidades del sistema.		
Estado	Proposed		

Tabla 3232: «AC» AC-02.Administrador

6.1.2.3 Requisitos funcionales

6.1.2.3.1 Diagramas de casos de uso

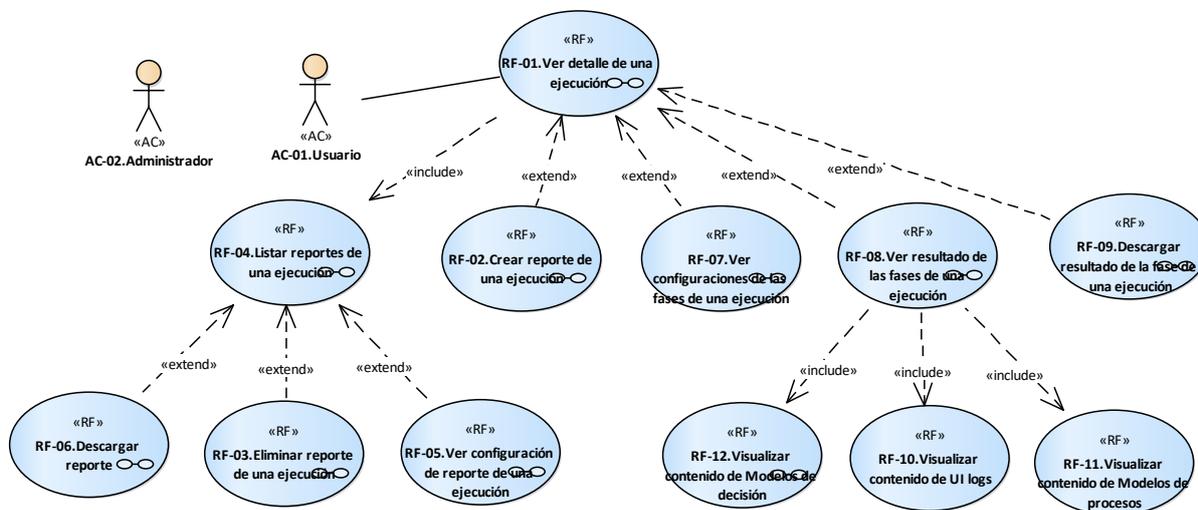


Figura 22: Diagrama de casos de uso

6.1.2.3.2 Definición de casos de uso del sistema

A continuación, se muestra la estructura general que seguirán los requisitos funcionales definidos para este sistema, para poder visualizar el detalle tanto de los diagramas como de la definición de todos los casos de uso, ir

a Memoria técnica - Anexo Requisitos Funcionales.docx incluido en el mismo directorio que el presente documento.

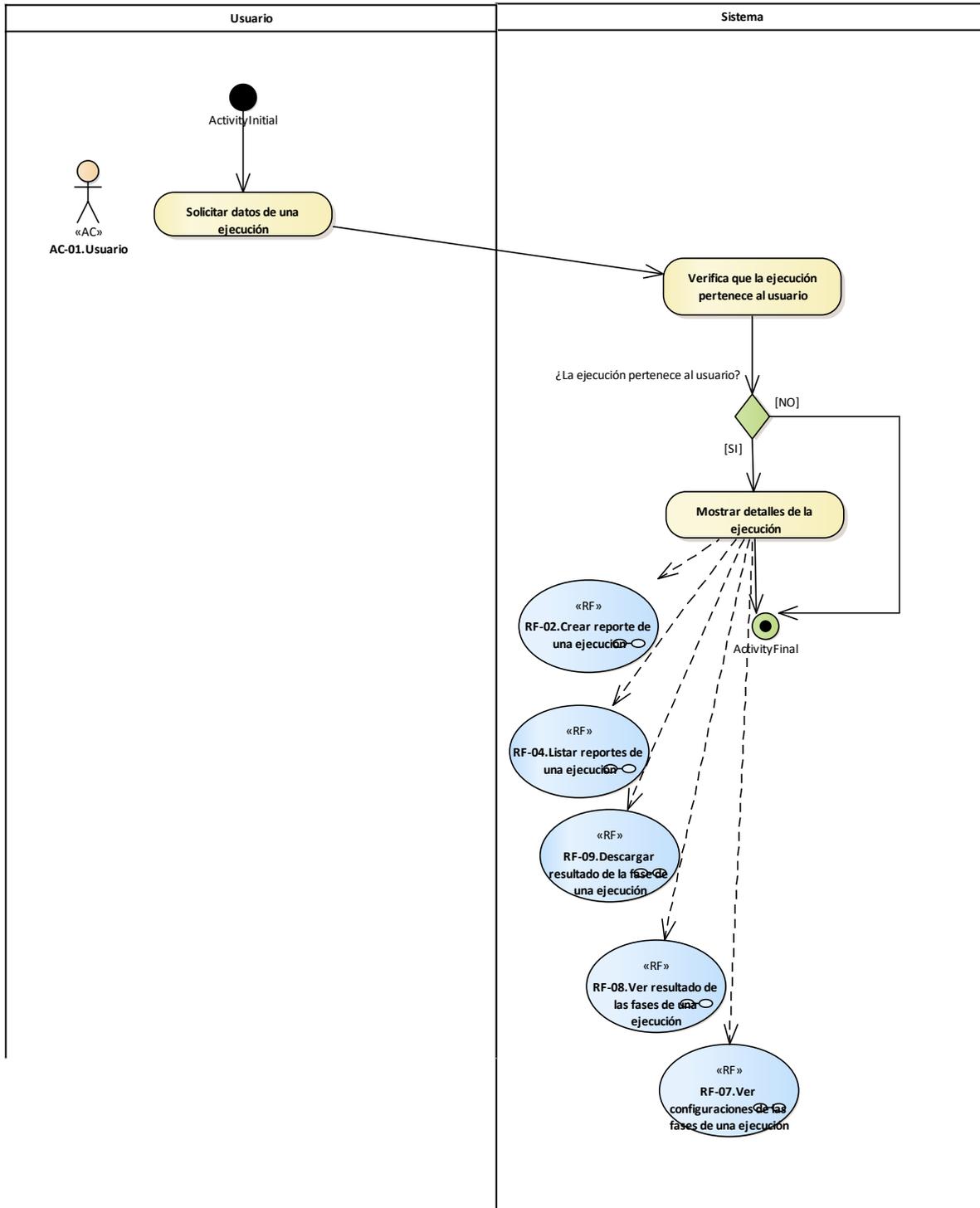


Figura 23 : RF-01.Ver detalle de una ejecución

«RF» RF-01.Ver detalle de una ejecución			
Versión	1.0	25/03/2024 11:26:59	
Autor	User		
Conexiones	Fuente	Estereotipo	Destino
	RF-01.Ver detalle de una ejecución	«Relationship»	PS-44.Ver detalle de una ejecución
	RF-01.Ver detalle de una ejecución	«Relationship»	CP-01.Subsistema general
	RF-01.Ver detalle de una ejecución	«Relationship»	PS-45.Ver detalle de una ejecución
	RF-01.Ver detalle de una ejecución	«Relationship»	PS-43.Ver detalle de una ejecución
	RF-01.Ver detalle de una ejecución	«include»	RF-04.Listar reportes de una ejecución
	RF-02.Crear reporte de una ejecución	«extend»	RF-01.Ver detalle de una ejecución
	AC-01.Usuario		RF-01.Ver detalle de una ejecución
	RF-08.Ver resultado de las fases de una ejecución	«extend»	RF-01.Ver detalle de una ejecución
	RF-09.Descargar resultado de la fase de una ejecución	«extend»	RF-01.Ver detalle de una ejecución
	OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA	«Relationship»	RF-01.Ver detalle de una ejecución
	RF-07.Ver configuraciones de las fases de una ejecución	«extend»	RF-01.Ver detalle de una ejecución
Descripción	El usuario debe tener acceso a una vista de los detalles de una ejecución, vista central desde donde se centralizará la manipulación de los reportes y resultados asociados a esa ejecución.		
Pre-condición y Post-condición	Post-condición: <i>Detalles</i> . Se redirige al usuario a la vista de detalles del caso de estudio Pre-condición: <i>Casos de Estudio</i> . El usuario debe tener al menos un caso de estudio Pre-condición: <i>Sesión Iniciada</i> . El usuario debe de haber iniciado sesión		
Estado	Proposed		

Tabla 3333: «RF» RF-01.Ver detalle de una ejecución

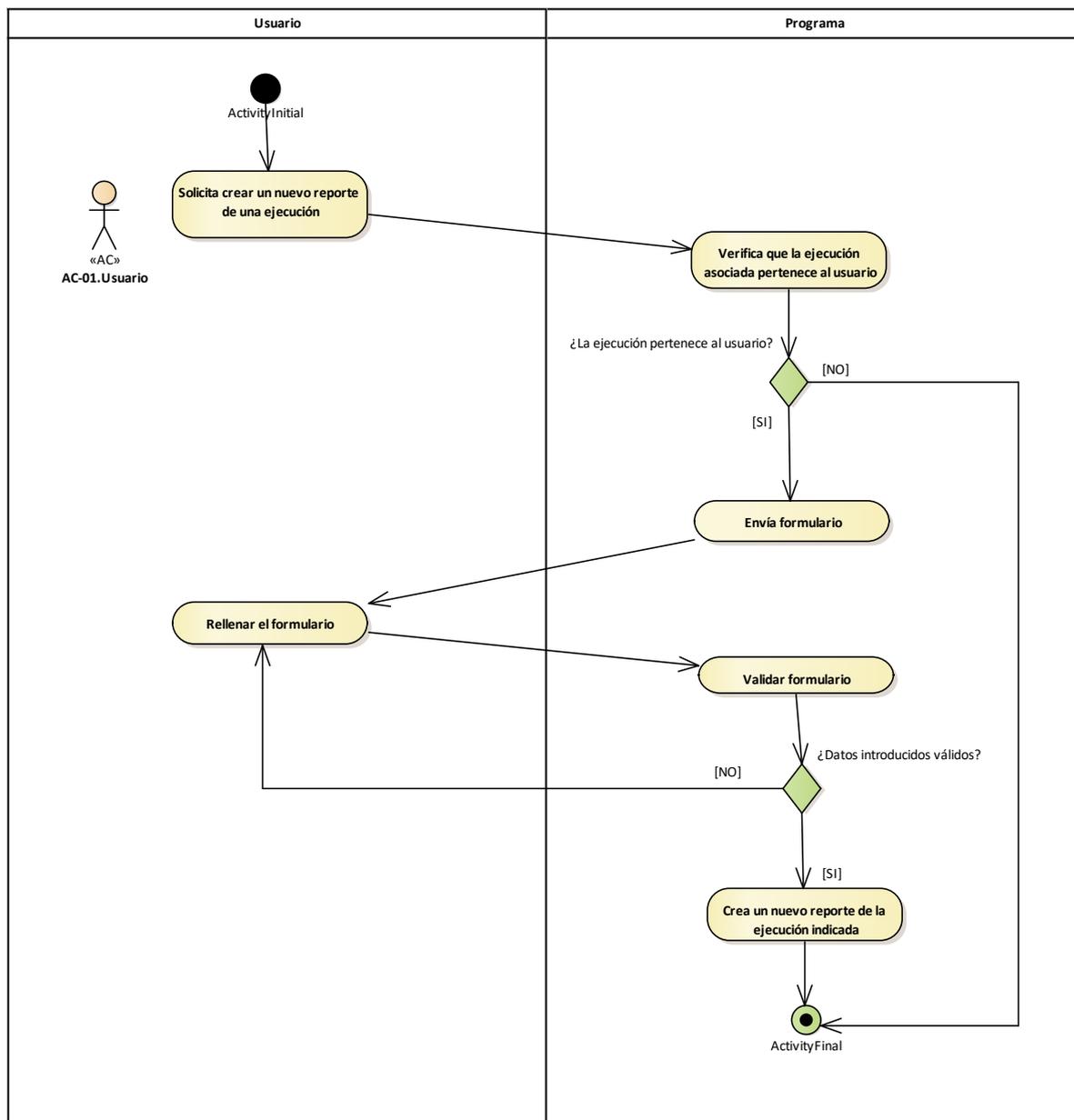


Figura 24: RF-02.Crear reporte de una ejecución

«RF» RF-02.Crear reporte de una ejecución			
Versión	1.0	25/03/2024 11:36:17	
Autor	User		
Conexiones	Fuente	Estereotipo	Destino
	RF-02.Crear reporte de una ejecución	«Relationship»	PS-41.Crear reporte de una ejecución
	RF-02.Crear reporte de una ejecución	«Relationship»	PS-42.Crear reporte de una ejecución
	RF-02.Crear reporte de una ejecución	«extend»	RF-01.Ver detalle de una ejecución

	RF-02.Crear reporte de una ejecución	«Relationship»	CP-01.Subsistema general
	RF-02.Crear reporte de una ejecución	«Relationship»	PS-40.Crear reporte de una ejecución
	RF-02.Crear reporte de una ejecución	«Relationship»	PS-39.Crear reporte de una ejecución
	OBJ-03.Aglutinar toda la información extraída en forma de reporte	«Relationship»	RF-02.Crear reporte de una ejecución
	OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA	«Relationship»	RF-02.Crear reporte de una ejecución
	OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis	«Relationship»	RF-02.Crear reporte de una ejecución
	AC-01.Usuario		RF-02.Crear reporte de una ejecución
Descripción	El objetivo de este requisito es permitir al usuario configurar, y con esta configuración generar un reporte de una ejecución de un caso de estudio.		
Pre-condición y Post-condición	Post-condición: <i>Creación</i> . Se crea un nuevo reporte con la configuración escogida Pre-condición: <i>Caso de Estudio</i> . El usuario tiene algún caso de estudio asociado Pre-condición: <i>Sesión iniciada</i> . El usuario ha iniciado sesión		
Estado	Proposed		

Tabla 3434: «RF» RF-21.Crear reporte de una ejecución

6.1.2.4 Requisitos no funcionales

6.1.2.4.1 Diagramas de requisitos no funcionales

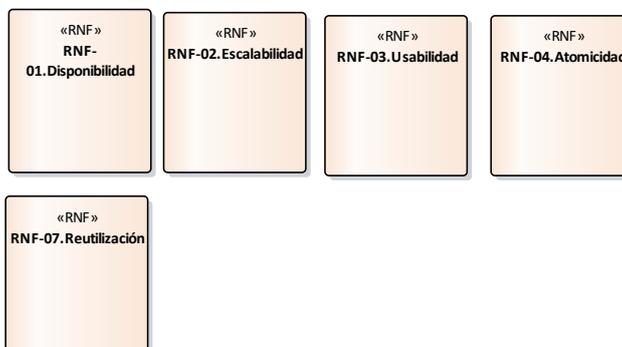


Figura 25: Diagrama de requisitos no funcionales

6.1.2.4.2 Definición de requisitos no funcionales

«RNF» RNF-01.Disponibilidad		
Versión	1.0	26/03/2024 11:27:48
Autor	User	
Objetivos asociados	OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA	
Descripción	El Sistema debe estar disponible 100% o muy cercano a esta disponibilidad.	
Estado	Proposed	

Tabla 3535: «RNF» RNF-01.Disponibilidad

«RNF» RNF-02.Escalabilidad		
Versión	1.0	26/03/2024 11:28:50
Autor	User	
Objetivos asociados	OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA OBJ-03.Aglutinar toda la información extraída en forma de reporte OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma	
Descripción	El sistema debe ser construido sobre la base de un desarrollo evolutivo e incremental, de manera tal que nuevas funcionalidades y requerimientos relacionados puedan ser incorporados afectando el código existente de la menor manera posible; para ello deben incorporarse aspectos de reutilización de componentes. El sistema debe estar en capacidad de permitir en el futuro el desarrollo de nuevas funcionalidades, modificar o eliminar funcionalidades después de su construcción y puesta en marcha inicial	
Estado	Proposed	

Tabla 3636: «RNF» RNF-02.Escalabilidad

«RNF» RNF-03.Usabilidad		
Versión	1.0	26/03/2024 11:29:46
Autor	User	
Objetivos asociados	OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis OBJ-03.Aglutinar toda la información extraída en forma de reporte OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma	
Descripción	El sistema debe ser de fácil uso por parte de los usuarios, así como de fácil adaptación para nuevos usuarios. El sistema no debe permitir el cierre de una operación hasta que todos sus procesos, subprocesos y tareas relacionados, hayan sido terminados y cerrados satisfactoriamente.	

	El sistema debe presentar mensajes de error que permitan al usuario identificar el tipo de error y comunicarse con el administrador del sistema.
Estado	Proposed

Tabla 3737: «RNF» RNF-03.Usabilidad

«RNF» RNF-04.Atomicidad	
Versión	1.0 26/03/2024 11:30:11
Autor	User
Objetivos asociados	OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA OBJ-04.Aplicar nuevas técnicas de minería de decisiones para análisis más complejos OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis OBJ-03.Aglutinar toda la información extraída en forma de reporte
Descripción	Debe de estar ligado a la resolución de una problemática concreta, con unas entradas y unas salidas claras.
Estado	Proposed

Tabla 3838: «RNF» RNF-04.Atomicidad

«RNF» RNF-07.Reutilización	
Versión	1.0 26/03/2024 11:33:08
Autor	User
Objetivos asociados	OBJ-04.Aplicar nuevas técnicas de minería de decisiones para análisis más complejos OBJ-05.Habilitar la visualización de los resultados del análisis extraído por la plataforma OBJ-03.Aglutinar toda la información extraída en forma de reporte OBJ-02.Aumentar la comprensibilidad de la información extraída a partir del análisis OBJ-01.Mejorar una plataforma de task mining aplicado al contexto de RPA
Descripción	Se deben diseñar potenciando la reutilización en los diversos escenarios e infraestructuras de los potenciales sistemas de información a complementar, y no de forma ligado a un determinado ecosistema particular
Estado	Proposed

Tabla 3939: «RNF» RNF-07.Reutilización

6.2 Análisis

La fase de análisis contendrá los productos resultantes de analizar, definir y estructurar los requisitos establecidos en la fase de Requisitos, según establece NDT [36], [37].

6.2.1 Modelo de clases de contenido

El modelo de clases de contenido representa el diagrama de clases derivado de los requisitos de almacenamiento de información, y permite modelar cómo se estructura la información que maneja la plataforma.

6.2.1.1 Diagrama de clases de contenido

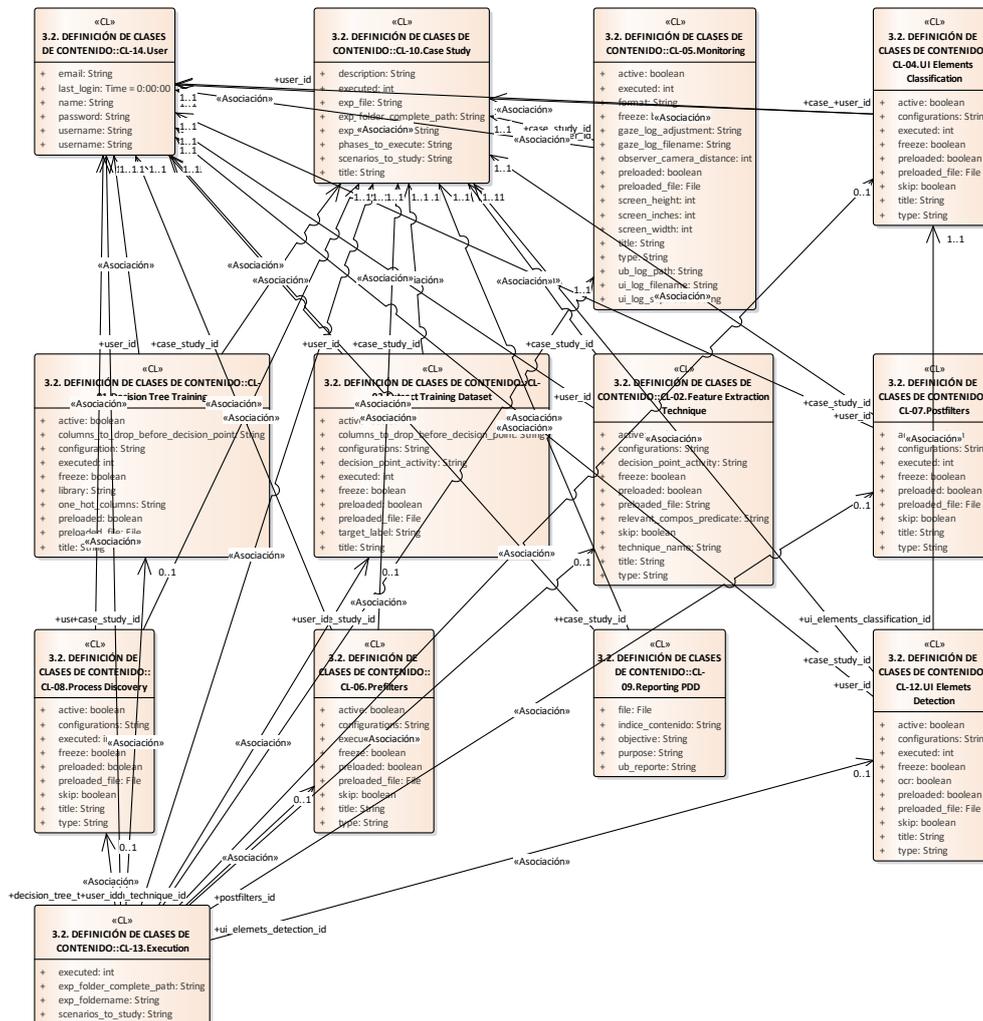


Figura 26: Diagrama de clases de contenido

6.2.1.2 Definición de clases de contenido

«CL» CL-01.Decision Tree Training			
Versión	1.0	12/04/2024 0:00:00	
Autor	NDT - Driver 3.0		
Conexiones	Fuente	Estereotipo	Destino
	CL-01.Decision Tree Training	«Asociación»	CL-10.Case Study
	CL-01.Decision Tree Training	«Asociación»	CL-14.User
	CL-13.Execution	«Asociación»	CL-01.Decision Tree Training
	RA-01.Decision Tree Training	«Relationship»	CL-01.Decision Tree Training

Artefacto DRS asociado	RA-01.Decision Tree Training
Descripción	Configuración para el entrenamiento de un arbol de decisión en base a un log enriquecido de eventos
Atributos	<p><i>boolean</i> active Atributo asociado: RA-01.Decision Tree Training@active Cardinalidad: [1, 1] Descripción: Si la configuración es la configuración a ejecutar al ejecutar el caso de estudio asociado</p> <hr/> <p><i>String</i> columns_to_drop_before_decision_point Atributo asociado: RA-01.Decision Tree Training@columns_to_drop_before_decision_point Cardinalidad: [1, 1] Descripción: Columnas a eliminar antes del punto de decisión</p> <hr/> <p><i>String</i> configuration Atributo asociado: RA-01.Decision Tree Training@configuration Cardinalidad: [1, 1] Descripción: JSON para configuracion adicional</p> <hr/> <p><i>int</i> executed Atributo asociado: RA-01.Decision Tree Training@executed Cardinalidad: [1, 1] Descripción: Número de ejecuciones con la configuración</p> <hr/> <p><i>boolean</i> freeze Atributo asociado: RA-01.Decision Tree Training@freeze Cardinalidad: [1, 1] Descripción: Si la configuración es de solo lectura</p> <hr/> <p><i>String</i> library Atributo asociado: RA-01.Decision Tree Training@library Cardinalidad: [1, 1] Descripción: Librería a la que pertenecen los algoritmos utilizados en la configuración</p> <hr/> <p><i>String</i> one_hot_columns Atributo asociado: RA-01.Decision Tree Training@one_hot_columns Cardinalidad: [1, 1] Descripción: Número de columnas tras transformar variables categóricas a numéricas</p> <hr/> <p><i>boolean</i> preloaded Atributo asociado: RA-01.Decision Tree Training@preloaded Cardinalidad: [1, 1] Descripción: Si los resultados de la fase están precargados</p> <hr/> <p><i>File</i> preloaded_file Atributo asociado: RA-01.Decision Tree Training@preloaded_file Cardinalidad: [1, 1] Descripción: Resultados precargados de la fase</p> <hr/> <p><i>String</i> title Atributo asociado: RA-01.Decision Tree Training@title Cardinalidad: [1, 1] Descripción: Título de la configuración de la fase</p>
Estado	Propuesto
Comentarios	

Tabla 40. «CL» CL-01.Decision Tree Training

«CL» CL-02.Feature Extraction Technique		
Versión	1.0	12/04/2024 0:00:00

Autor	NDT - Driver 3.0		
Conexiones	Fuente	Estereotipo	Destino
	CL-02.Feature Extraction Technique	«Asociación»	CL-14.User
	CL-02.Feature Extraction Technique	«Asociación»	CL-10.Case Study
	CL-13.Execution	«Asociación»	CL-02.Feature Extraction Technique
	RA-02.Feature Extraction Technique	«Relationship»	CL-02.Feature Extraction Technique
Artefacto DRS asociado	RA-02.Feature Extraction Technique		
Descripción	Configuración de la técnica para extraer valor a un log de eventos		
Atributos	<p><i>boolean</i> active Atributo asociado: RA-02.Feature Extraction Technique@active Cardinalidad: [1, 1] Descripción: Si la configuración es la configuración a ejecutar al ejecutar el caso de estudio asociado</p>		
	<p><i>String</i> configurations Atributo asociado: RA-02.Feature Extraction Technique@configurations Cardinalidad: [1, 1] Descripción: JSON para configuracion adicional</p>		
	<p><i>String</i> decision_point_activity Atributo asociado: RA-02.Feature Extraction Technique@decision_point_activity Cardinalidad: [1, 1] Descripción: Punto de decisión a estudiar</p>		
	<p><i>boolean</i> freeze Atributo asociado: RA-02.Feature Extraction Technique@freeze Cardinalidad: [1, 1] Descripción: Si la configuración es de solo lectura</p>		
	<p><i>boolean</i> preloaded Atributo asociado: RA-02.Feature Extraction Technique@preloaded Cardinalidad: [1, 1] Descripción: Si los resultados de la fase están precargados</p>		
	<p><i>String</i> preloaded_file Atributo asociado: RA-02.Feature Extraction Technique@preloaded_file Cardinalidad: [1, 1] Descripción: Resultados precargados de la fase</p>		
	<p><i>String</i> relevant_compos_predicate Atributo asociado: RA-02.Feature Extraction Technique@relevant_compos_predicate Cardinalidad: [1, 1] Descripción: Condición para que un componente UI sea relevante</p>		
	<p><i>boolean</i> skip Atributo asociado: RA-02.Feature Extraction Technique@skip Cardinalidad: [1, 1] Descripción: Si la fase se salta (no se ejecuta) al realizar la ejecución del caso de estudio</p>		
	<p><i>String</i> technique_name Atributo asociado: RA-02.Feature Extraction Technique@technique_name Cardinalidad: [1, 1] Descripción: Nombre de la técnica configurada</p>		
	<p><i>String</i> title Atributo asociado: RA-02.Feature Extraction Technique@title Cardinalidad: [1, 1] Descripción: Título de la configuración de la fase</p>		
	<p><i>String</i> type Atributo asociado: RA-02.Feature Extraction Technique@type Cardinalidad: [1, 1]</p>		

	Descripción: Técnica para extraer conocimiento de un log de eventos
Estado	Propuesto
Comentarios	

Tabla 4140: «CL» CL-02.Feature Extraction Technique

«CL» CL-03.Extract Training Dataset			
Versión	1.0	12/04/2024 0:00:00	
Autor	NDT - Driver 3.0		
Conexiones	Fuente	Estereotipo	Destino
	CL-03.Extract Training Dataset	«Asociación»	CL-10.Case Study
	CL-03.Extract Training Dataset	«Asociación»	CL-14.User
	RA-03.Extract Training Dataset	«Relationship»	CL-03.Extract Training Dataset
	CL-13.Execution	«Asociación»	CL-03.Extract Training Dataset
Artefacto DRS asociado	RA-03.Extract Training Dataset		
Descripción	Configuración para la extracción de datos para entrenar el arbol de decision		
Atributos	<i>boolean</i> active Atributo asociado: RA-03.Extract Training Dataset@active Cardinalidad: [1, 1] Descripción: Si la configuración es la configuración a ejecutar al ejecutar el caso de estudio asociado		
	<i>String</i> columns_to_drop_before_decision_point Atributo asociado: RA-03.Extract Training Dataset@columns_to_drop_before_decision_point Cardinalidad: [1, 1] Descripción: Columnas a eliminar antes del punto de decisión		
	<i>String</i> configurations Atributo asociado: RA-03.Extract Training Dataset@configurations Cardinalidad: [1, 1] Descripción: JSON para configuracion adicional		
	<i>String</i> decision_point_activity Atributo asociado: RA-03.Extract Training Dataset@decision_point_activity Cardinalidad: [1, 1] Descripción: Punto de decisión a estudiar		
	<i>int</i> executed Atributo asociado: RA-03.Extract Training Dataset@executed Cardinalidad: [1, 1] Descripción: Número de ejecuciones con la configuración		
	<i>boolean</i> freeze Atributo asociado: RA-03.Extract Training Dataset@freeze Cardinalidad: [1, 1] Descripción: Si la configuración es de solo lectura		
	<i>boolean</i> preloaded Atributo asociado: RA-03.Extract Training Dataset@preloaded Cardinalidad: [1, 1] Descripción: Si los resultados de la fase están precargados		
	<i>File</i> preloaded_file Atributo asociado: RA-03.Extract Training Dataset@preloaded_file Cardinalidad: [1, 1]		

	Descripción: Resultados precargados de la fase <i>String</i> target_label Atributo asociado: RA-03.Extract Training Dataset@target_label Cardinalidad: [1, 1] Descripción: Variable objetivo a predecir
	Descripción: Variable objetivo a predecir <i>String</i> title Atributo asociado: RA-03.Extract Training Dataset@title Cardinalidad: [1, 1] Descripción: Título de la configuración de la fase
Estado	Propuesto
Comentarios	

Tabla 4241: «CL» CL-03.Extract Training Dataset

«CL» CL-04.UI Elements Classification			
Versión	1.0	12/04/2024 0:00:00	
Autor	NDT - Driver 3.0		
Conexiones	Fuente	Estereotipo	Destino
	CL-04.UI Elements Classification	«Asociación»	CL-14.User
	CL-04.UI Elements Classification	«Asociación»	CL-10.Case Study
	RA-04.UI Elements Classification	«Relationship»	CL-04.UI Elements Classification
	CL-13.Execution	«Asociación»	CL-04.UI Elements Classification
	CL-12.UI Elemets Detection	«Asociación»	CL-04.UI Elements Classification
Artefacto DRS asociado	RA-04.UI Elements Classification		
Descripción	Configuración para la técnica de clasificación de elementos UI		
Atributos	<i>boolean</i> active Atributo asociado: RA-04.UI Elements Classification@active Cardinalidad: [1, 1] Descripción: Si la configuración es la configuración a ejecutar al ejecutar el caso de estudio asociado		
	<i>String</i> configurations Atributo asociado: RA-04.UI Elements Classification@configurations Cardinalidad: [1, 1] Descripción: JSON para configuracion adicional		
	<i>int</i> executed Atributo asociado: RA-04.UI Elements Classification@executed Cardinalidad: [1, 1] Descripción: Número de ejecuciones con la configuración		
	<i>boolean</i> freeze Atributo asociado: RA-04.UI Elements Classification@freeze Cardinalidad: [1, 1] Descripción: Si la configuración es de solo lectura		
	<i>boolean</i> preloaded Atributo asociado: RA-04.UI Elements Classification@preloaded Cardinalidad: [1, 1] Descripción: Si los resultados de la fase están precargados		
	<i>File</i> preloaded_file Atributo asociado: RA-04.UI Elements Classification@preloaded_file Cardinalidad: [1, 1] Descripción: Resultados precargados de la fase		
	<i>boolean</i> skip		

	<p>Atributo asociado: RA-04.UI Elements Classification@skip Cardinalidad: [1, 1] Descripción: Si la fase se salta (no se ejecuta) al realizar la ejecución del caso de estudio</p> <hr/> <p><i>String</i> title Atributo asociado: RA-04.UI Elements Classification@title Cardinalidad: [1, 1] Descripción: Título de la configuración de la fase</p> <hr/> <p><i>String</i> type Atributo asociado: RA-04.UI Elements Classification@type Cardinalidad: [1, 1] Descripción: Técnica usada para la clasificación de elementos UI</p>
Estado	Propuesto
Comentarios	

Tabla 4342: «CL» CL-04.UI Elements Classification

«CL» CL-05.Monitoring			
Versión	1.0	12/04/2024 0:00:00	
Autor	NDT - Driver 3.0		
Conexiones	Fuente	Estereotipo	Destino
	CL-05.Monitoring	«Asociación»	CL-14.User
	CL-05.Monitoring	«Asociación»	CL-10.Case Study
	CL-13.Execution	«Asociación»	CL-05.Monitoring
	RA-05.Monitoring	«Relationship»	CL-05.Monitoring
Artefacto DRS asociado	RA-05.Monitoring		
Descripción	Configuración sobre la información del log de eventos con las acciones de un usuario (keystrokes, clicks, etc.)		
Atributos	<i>boolean</i> active Atributo asociado: RA-05.Monitoring@active Cardinalidad: [1, 1] Descripción: Configuración activa a ejecutar al ejecutar un caso de estudio		
	<i>int</i> executed Atributo asociado: RA-05.Monitoring@executed Cardinalidad: [1, 1] Descripción: Número ejecuciones con esta configuración de la fase		
	<i>String</i> format Atributo asociado: RA-05.Monitoring@format Cardinalidad: [1, 1] Descripción: Si esta en formato .mht u otro. Depende de la herramienta con la que se haya grabado el log		
	<i>boolean</i> freeze Atributo asociado: RA-05.Monitoring@freeze Cardinalidad: [1, 1] Descripción: Si la configuración está en modo de solo lectura		
	<i>String</i> gaze_log_adjustment Atributo asociado: RA-05.Monitoring@gaze_log_adjustment Cardinalidad: [1, 1] Descripción: Ajuste al gaze log		
	<i>String</i> gaze_log_filename Atributo asociado: RA-05.Monitoring@gaze_log_filename		

	<p>Cardinalidad: [1, 1] Descripción: Nombre del log de eyetracking</p> <hr/> <p><i>int</i> observer_camera_distance Atributo asociado: RA-05.Monitoring@observer_camera_distance Cardinalidad: [1, 1] Descripción: Distancia desde la cámara al usuario monitorizado</p> <hr/> <p><i>boolean</i> preloaded Atributo asociado: RA-05.Monitoring@preloaded Cardinalidad: [1, 1] Descripción: Si los resultados de la fase están precargados</p> <hr/> <p><i>File</i> preloaded_file Atributo asociado: RA-05.Monitoring@preloaded_file Cardinalidad: [1, 1] Descripción: Resultados precargados de la fase</p> <hr/> <p><i>int</i> screen_height Atributo asociado: RA-05.Monitoring@screen_height Cardinalidad: [1, 1] Descripción: Altura de la pantalla monitorizada</p> <hr/> <p><i>int</i> screen_inches Atributo asociado: RA-05.Monitoring@screen_inches Cardinalidad: [1, 1] Descripción: Pulgadas de la pantalla monitorizada</p> <hr/> <p><i>int</i> screen_width Atributo asociado: RA-05.Monitoring@screen_width Cardinalidad: [1, 1] Descripción: Ancho de la pantalla monitorizada</p> <hr/> <p><i>String</i> title Atributo asociado: RA-05.Monitoring@title Cardinalidad: [1, 1] Descripción: Título de la configuración de la fase</p> <hr/> <p><i>String</i> type Atributo asociado: RA-05.Monitoring@type Cardinalidad: [1, 1] Descripción: Herramienta usada para monitorear una actividad</p> <hr/> <p><i>String</i> ub_log_path Atributo asociado: RA-05.Monitoring@ub_log_path Cardinalidad: [1, 1] Descripción: Ubicación en la jerarquía de carpeta donde se guarda el log</p> <hr/> <p><i>String</i> ui_log_filename Atributo asociado: RA-05.Monitoring@ui_log_filename Cardinalidad: [1, 1] Descripción: Nombre del archivo (log) contenedor de las acciones del usuario</p> <hr/> <p><i>String</i> ui_log_separator Atributo asociado: RA-05.Monitoring@ui_log_separator Cardinalidad: [1, 1] Descripción: Separador usado en el log (,;)</p>
Estado	Propuesto
Comentarios	

Tabla 4443: «CL» CL-05.Monitoring

«CL» CL-06.Prefilters			
Versión	1.0	12/04/2024 0:00:00	
Autor	NDT - Driver 3.0		
Conexiones	Fuente	Estereotipo	Destino
	CL-06.Prefilters	«Asociación»	CL-14.User
	CL-06.Prefilters	«Asociación»	CL-10.Case Study
	CL-13.Execution	«Asociación»	CL-06.Prefilters
	RA-06.Prefilters	«Relationship»	CL-06.Prefilters
Artefacto DRS asociado	RA-06.Prefilters		
Descripción	Configuración para el filtrado de información del log de eventos		
Atributos	<i>boolean</i> active Atributo asociado: RA-06.Prefilters@active Cardinalidad: [1, 1] Descripción: Si la configuración es la configuración a ejecutar al ejecutar el caso de estudio asociado		
	<i>String</i> configurations Atributo asociado: RA-06.Prefilters@configurations Cardinalidad: [1, 1] Descripción: JSON para configuracion adicional		
	<i>int</i> executed Atributo asociado: RA-06.Prefilters@executed Cardinalidad: [1, 1] Descripción: Número de ejecuciones con la configuración		
	<i>boolean</i> freeze Atributo asociado: RA-06.Prefilters@freeze Cardinalidad: [1, 1] Descripción: Si la configuración es de solo lectura		
	<i>boolean</i> preloaded Atributo asociado: RA-06.Prefilters@preloaded Cardinalidad: [1, 1] Descripción: Si los resultados de la fase están precargados		
	<i>File</i> preloaded_file Atributo asociado: RA-06.Prefilters@preloaded_file Cardinalidad: [1, 1] Descripción: Resultados precargados de la fase		
	<i>boolean</i> skip Atributo asociado: RA-06.Prefilters@skip Cardinalidad: [1, 1] Descripción: Si la fase se salta (no se ejecuta) al realizar la ejecución del caso de estudio		
	<i>String</i> title Atributo asociado: RA-06.Prefilters@title Cardinalidad: [1, 1] Descripción: Título de la configuración de la fase		
	<i>String</i> type Atributo asociado: RA-06.Prefilters@type Cardinalidad: [1, 1] Descripción: Herramienta usada para filtrar el contenido de un log		
	Estado	Propuesto	
Comentarios			

Tabla 45 44: «CL» CL-06.Prefilters

«CL» CL-07.Postfilters			
Versión	1.0	12/04/2024 0:00:00	
Autor	NDT - Driver 3.0		
Conexiones	Fuente	Estereotipo	Destino
	CL-07.Postfilters	«Asociación»	CL-14.User
	CL-07.Postfilters	«Asociación»	CL-10.Case Study
	RA-07.Postfilters	«Relationship»	CL-07.Postfilters
	CL-13.Execution	«Asociación»	CL-07.Postfilters
Artefacto DRS asociado	RA-07.Postfilters		
Descripción	Configuración para el filtrado de información del log de eventos enriquecido		
Atributos	<i>boolean</i> active Atributo asociado: RA-07.Postfilters@active Cardinalidad: [1, 1] Descripción: Si la configuración es la configuración a ejecutar al ejecutar el caso de estudio asociado		
	<i>String</i> configurations Atributo asociado: RA-07.Postfilters@configurations Cardinalidad: [1, 1] Descripción: JSON para configuracion adicional		
	<i>int</i> executed Atributo asociado: RA-07.Postfilters@executed Cardinalidad: [1, 1] Descripción: Número de ejecuciones con la configuración		
	<i>boolean</i> freeze Atributo asociado: RA-07.Postfilters@freeze Cardinalidad: [1, 1] Descripción: Si la configuración es de solo lectura		
	<i>boolean</i> preloaded Atributo asociado: RA-07.Postfilters@preloaded Cardinalidad: [1, 1] Descripción: Si los resultados de la fase están precargados		
	<i>File</i> preloaded_file Atributo asociado: RA-07.Postfilters@preloaded_file Cardinalidad: [1, 1] Descripción: Resultados precargados de la fase		
	<i>boolean</i> skip Atributo asociado: RA-07.Postfilters@skip Cardinalidad: [1, 1] Descripción: Si la fase se salta (no se ejecuta) al realizar la ejecución del caso de estudio		
	<i>String</i> title Atributo asociado: RA-07.Postfilters@title Cardinalidad: [1, 1] Descripción: Título de la configuración de la fase		
	<i>String</i> type Atributo asociado: RA-07.Postfilters@type Cardinalidad: [1, 1] Descripción: Técnica para filtrar el contenido de un log enriquecido		
	Estado	Propuesto	
Comentarios			

Tabla 46 45: «CL» CL-07.Postfilters

«CL» CL-08.Process Discovery			
Versión	1.0	12/04/2024 0:00:00	
Autor	NDT - Driver 3.0		
Conexiones	Fuente	Estereotipo	Destino
	CL-08.Process Discovery	«Asociación»	CL-14.User
	CL-08.Process Discovery	«Asociación»	CL-10.Case Study
	RA-08.Process Discovery	«Relationship»	CL-08.Process Discovery
	CL-13.Execution	«Asociación»	CL-08.Process Discovery
Artefacto DRS asociado	RA-08.Process Discovery		
Descripción	Configuración para el descubrimiento de procesos BPM		
Atributos	<i>boolean</i> active Atributo asociado: RA-08.Process Discovery@active Cardinalidad: [1, 1] Descripción: Si la configuración es la configuración a ejecutar al ejecutar el caso de estudio asociado		
	<i>String</i> configurations Atributo asociado: RA-08.Process Discovery@configurations Cardinalidad: [1, 1] Descripción: JSON para configuracion adicional		
	<i>int</i> executed Atributo asociado: RA-08.Process Discovery@executed Cardinalidad: [1, 1] Descripción: Número de ejecuciones con la configuración		
	<i>boolean</i> freeze Atributo asociado: RA-08.Process Discovery@freeze Cardinalidad: [1, 1] Descripción: Si la configuración es de solo lectura		
	<i>boolean</i> preloaded Atributo asociado: RA-08.Process Discovery@preloaded Cardinalidad: [1, 1] Descripción: Si los resultados de la fase están precargados		
	<i>File</i> preloaded_file Atributo asociado: RA-08.Process Discovery@preloaded_file Cardinalidad: [1, 1] Descripción: Resultados precargados de la fase		
	<i>boolean</i> skip Atributo asociado: RA-08.Process Discovery@skip Cardinalidad: [1, 1] Descripción: Si la fase se salta (no se ejecuta) al realizar la ejecución del caso de estudio		
	<i>String</i> title Atributo asociado: RA-08.Process Discovery@title Cardinalidad: [1, 1] Descripción: Título de la configuración de la fase		
	<i>String</i> type Atributo asociado: RA-08.Process Discovery@type Cardinalidad: [1, 1] Descripción: Técnica para descubrir el proceso de negocio		
	Estado	Propuesto	
Comentarios			

Tabla 4746: «CL» CL-08.Process Discovery

«CL» CL-09.Reporting PDD			
Versión	1.0	12/04/2024 0:00:00	
Autor	NDT - Driver 3.0		
Conexiones	Fuente	Estereotipo	Destino
	CL-09.Reporting PDD	«Asociación»	CL-14.User
	CL-09.Reporting PDD	«Asociación»	CL-10.Case Study
	RA-09.Reporting PDD	«Relationship»	CL-09.Reporting PDD
Artefacto DRS asociado	RA-09.Reporting PDD		
Descripción	Reporte de una ejecución de un caso de estudio		
Atributos	<i>File</i> file Atributo asociado: RA-09.Reporting PDD@file Cardinalidad: [1, 1] Descripción: Documento de reporte		
	<i>String</i> indice_contenido Atributo asociado: RA-09.Reporting PDD@indice_contenido Cardinalidad: [1, 1] Descripción: Índice de contenido configurado por el usuario a mostrar en el reporte		
	<i>String</i> objective Atributo asociado: RA-09.Reporting PDD@Objective Cardinalidad: [1, 1] Descripción: Objetivo del estudio		
	<i>String</i> purpose Atributo asociado: RA-09.Reporting PDD@Purpose Cardinalidad: [1, 1] Descripción: Propósito del estudio		
	<i>String</i> ub_reporte Atributo asociado: RA-09.Reporting PDD@ub_reporte Cardinalidad: [1, 1] Descripción: Ubicación en la jerarquía de carpetas donde se almacena el reporte		
Estado	Propuesto		
Comentarios			

Tabla 4847: «CL» CL-09.Reporting PDD

«CL» CL-10.Case Study			
Versión	1.0	12/04/2024 0:00:00	
Autor	NDT - Driver 3.0		
Conexiones	Fuente	Estereotipo	Destino
	CL-10.Case Study	«Asociación»	CL-14.User
	CL-01.Decision Tree Training	«Asociación»	CL-10.Case Study
	RA-10.Case Study	«Relationship»	CL-10.Case Study
	CL-04.UI Elements Classification	«Asociación»	CL-10.Case Study
	CL-12.UI Elemets Detection	«Asociación»	CL-10.Case Study
	CL-09.Reporting PDD	«Asociación»	CL-10.Case Study
	CL-02.Feature Extraction Technique	«Asociación»	CL-10.Case Study
	CL-05.Monitoring	«Asociación»	CL-10.Case Study
	CL-07.Postfilters	«Asociación»	CL-10.Case Study
	CL-03.Extract Training Dataset	«Asociación»	CL-10.Case Study
	CL-06.Prefilters	«Asociación»	CL-10.Case Study
CL-08.Process Discovery	«Asociación»	CL-10.Case Study	

	CL-13.Execution	«Asociación»	CL-10.Case Study
Artefacto DRS asociado	RA-10.Case Study		
Descripción	Caso de estudio		
Atributos	<i>String</i> description Atributo asociado: RA-10.Case Study@description Cardinalidad: [1, 1] Descripción: Descripción del caso de estudio		
	<i>int</i> executed Atributo asociado: RA-10.Case Study@executed Cardinalidad: [1, 1] Descripción: Número de veces que ha sido ejecutado el caso de estudio		
	<i>String</i> exp_file Atributo asociado: RA-10.Case Study@exp_file Cardinalidad: [1, 1] Descripción: Archivo que contiene toda la información del caso de estudio		
	<i>String</i> exp_folder_complete_path Atributo asociado: RA-10.Case Study@exp_folder_complete_path Cardinalidad: [1, 1] Descripción: Ruta completa donde se almacena el contenido relacionado con el caso de estudio		
	<i>String</i> exp_foldername Atributo asociado: RA-10.Case Study@exp_foldername Cardinalidad: [1, 1] Descripción: Nombre del directorio donde se almacena el contenido relacionado con el caso de estudio		
	<i>String</i> phases_to_execute Atributo asociado: RA-10.Case Study@phases_to_execute Cardinalidad: [1, 1] Descripción: Fases a ejecutar cuando se ejecute el caso de estudio		
	<i>String</i> escenarios_to_study Atributo asociado: RA-10.Case Study@escenarios_to_study Cardinalidad: [1, 1] Descripción: Escenarios del caso de estudio a ejecutar cuando se ejecute el caso de estudio		
	<i>String</i> title Atributo asociado: RA-10.Case Study@title Cardinalidad: [1, 1] Descripción: Título del caso de estudio		
	Estado	Propuesto	
Comentarios			

Tabla 4948: «CL» CL-10.Case Study

«CL» CL-12.UI Elemets Detection			
Versión	1.0	12/04/2024 0:00:00	
Autor	NDT - Driver 3.0		
Conexiones	Fuente	Estereotipo	Destino
	CL-12.UI Elemets Detection	«Asociación»	CL-10.Case Study
	CL-12.UI Elemets Detection	«Asociación»	CL-14.User
	CL-12.UI Elemets Detection	«Asociación»	CL-04.UI Elements Classification
	RA-12.UI Elemets Detection	«Relationship»	CL-12.UI Elemets Detection
	CL-13.Execution	«Asociación»	CL-12.UI Elemets Detection
Artefacto DRS asociado	RA-12.UI Elemets Detection		

Descripción	Configuración para la técnica de detección de elementos UI
Atributos	<p><i>boolean</i> active Atributo asociado: RA-12.UI Elemets Detection@active Cardinalidad: [1, 1] Descripción: Si la configuración es la configuración a ejecutar al ejecutar el caso de estudio asociado</p>
	<p><i>String</i> configurations Atributo asociado: RA-12.UI Elemets Detection@configurations Cardinalidad: [1, 1] Descripción: JSON para configuracion adicional</p>
	<p><i>int</i> executed Atributo asociado: RA-12.UI Elemets Detection@executed Cardinalidad: [1, 1] Descripción: Número de ejecuciones con la configuración</p>
	<p><i>boolean</i> freeze Atributo asociado: RA-12.UI Elemets Detection@freeze Cardinalidad: [1, 1] Descripción: Si la configuración es de solo lectura</p>
	<p><i>boolean</i> ocr Atributo asociado: RA-12.UI Elemets Detection@ocr Cardinalidad: [1, 1] Descripción: Si se aplica OCR</p>
	<p><i>boolean</i> preloaded Atributo asociado: RA-12.UI Elemets Detection@preloaded Cardinalidad: [1, 1] Descripción: Si los resultados de la fase están precargados</p>
	<p><i>File</i> preloaded_file Atributo asociado: RA-12.UI Elemets Detection@preloaded_file Cardinalidad: [1, 1] Descripción: Resultados precargados de la fase</p>
	<p><i>boolean</i> skip Atributo asociado: RA-12.UI Elemets Detection@skip Cardinalidad: [1, 1] Descripción: Si la fase se salta (no se ejecuta) al realizar la ejecución del caso de estudio</p>
	<p><i>String</i> title Atributo asociado: RA-12.UI Elemets Detection@title Cardinalidad: [1, 1] Descripción: Título de la configuración de la fase</p>
	<p><i>String</i> type Atributo asociado: RA-12.UI Elemets Detection@type Cardinalidad: [1, 1] Descripción: Técnica usada para detectar elementos UI</p>
Estado	Propuesto
Comentarios	

Tabla 5049: «CL» CL-12.UI Elements Detection

«CL» CL-13.Execution			
Versión	1.0	12/04/2024 0:00:00	
Autor	NDT - Driver 3.0		
Conexiones	Fuente	Estereotipo	Destino

	CL-13.Execution	«Asociación»	CL-05.Monitoring
	CL-13.Execution	«Asociación»	CL-01.Decision Tree Training
	CL-13.Execution	«Asociación»	CL-07.Postfilters
	CL-13.Execution	«Asociación»	CL-04.UI Elements Classification
	CL-13.Execution	«Asociación»	CL-03.Extract Training Dataset
	CL-13.Execution	«Asociación»	CL-02.Feature Extraction Technique
	CL-13.Execution	«Asociación»	CL-06.Prefilters
	CL-13.Execution	«Asociación»	CL-08.Process Discovery
	CL-13.Execution	«Asociación»	CL-12.UI Elemets Detection
	CL-13.Execution	«Asociación»	CL-14.User
	CL-13.Execution	«Asociación»	CL-10.Case Study
	RA-13.Execution	«Relationship»	CL-13.Execution
Artefacto DRS asociado	RA-13.Execution		
Descripción	Ejecucion de un caso de estudio		
Atributos	<p><i>int</i> executed Atributo asociado: RA-13.Execution@executed Cardinalidad: [1, 1] Descripción: Porcentaje que representa si la ejecucion se ha completado satisfactoriamente o en que porcentaje se ha quedado</p> <hr/> <p><i>String</i> exp_folder_complete_path Atributo asociado: RA-13.Execution@exp_folder_complete_path Cardinalidad: [1, 1] Descripción: Ruta completa donde se almacena el contenido relacionado con la ejecución</p> <hr/> <p><i>String</i> exp_foldername Atributo asociado: RA-13.Execution@exp_foldername Cardinalidad: [1, 1] Descripción: Nombre del directorio donde se almacena el contenido relacionado con la ejecución</p> <hr/> <p><i>String</i> escenarios_to_study Atributo asociado: RA-13.Execution@escenarios_to_study Cardinalidad: [1, 1] Descripción: Escenarios ejecutados del caso de estudio</p>		
Estado	Propuesto		
Comentarios			

Tabla 5150: «CL» CL-13.Execution

«CL» CL-14.User			
Versión	1.0	12/04/2024 0:00:00	
Autor	NDT - Driver 3.0		
Conexiones	Fuente	Estereotipo	Destino
	CL-02.Feature Extraction Technique	«Asociación»	CL-14.User
	CL-01.Decision Tree Training	«Asociación»	CL-14.User
	CL-05.Monitoring	«Asociación»	CL-14.User
	CL-04.UI Elements Classification	«Asociación»	CL-14.User
	CL-07.Postfilters	«Asociación»	CL-14.User
	CL-06.Prefilters	«Asociación»	CL-14.User
	CL-09.Reporting PDD	«Asociación»	CL-14.User
	CL-08.Process Discovery	«Asociación»	CL-14.User
	CL-03.Extract Training Dataset	«Asociación»	CL-14.User
	CL-10.Case Study	«Asociación»	CL-14.User

	CL-13.Execution	«Asociación»	CL-14.User
	CL-12.UI Elemets Detection	«Asociación»	CL-14.User
	RA-14.User	«Relationship»	CL-14.User
Artefacto DRS asociado	RA-14.User		
Descripción	Usuario del sistema		
Atributos	<i>String</i> email Atributo asociado: RA-14.User@email Cardinalidad: [1, 1] Descripción: Email del usuario		
	<i>Time</i> last_login Atributo asociado: RA-14.User@last_login Cardinalidad: [1, 1] Descripción: Hora del último inicio de sesión del usuario		
	<i>String</i> name Atributo asociado: RA-14.User@name Cardinalidad: [1, 1] Descripción: Apodo del usuario		
	<i>String</i> password Atributo asociado: RA-14.User@password Cardinalidad: [1, 1] Descripción: Contraseña del usuario		
	<i>String</i> username Atributo asociado: RA-14.User@username Cardinalidad: [1, 1] Descripción: Apellidos del usuario		
	<i>String</i> username Atributo asociado: RA-14.User@username Cardinalidad: [1, 1] Descripción: Nombre del usuario		
Estado	Propuesto		
Comentarios			

Tabla 5251: «CL» CL-14.User

6.2.2 Modelos de clases de procesos

6.2.2.1 Diagrama de clases de procesos

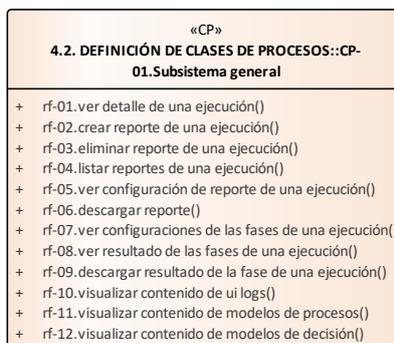


Figura 27 : Diagrama de clases de procesos

6.2.2.2 Definición de clases de procesos

A continuación, sólo se mostrarán los diagramas de secuencia considerados como los más relevantes para el proyecto. Para más detalle, todos los diagramas y definiciones de clases de procesos se encontrarán en el anexo *Memoria técnica - Anexo Modelos de clases de procesos.docx* incluido en el mismo directorio que el presente documento.

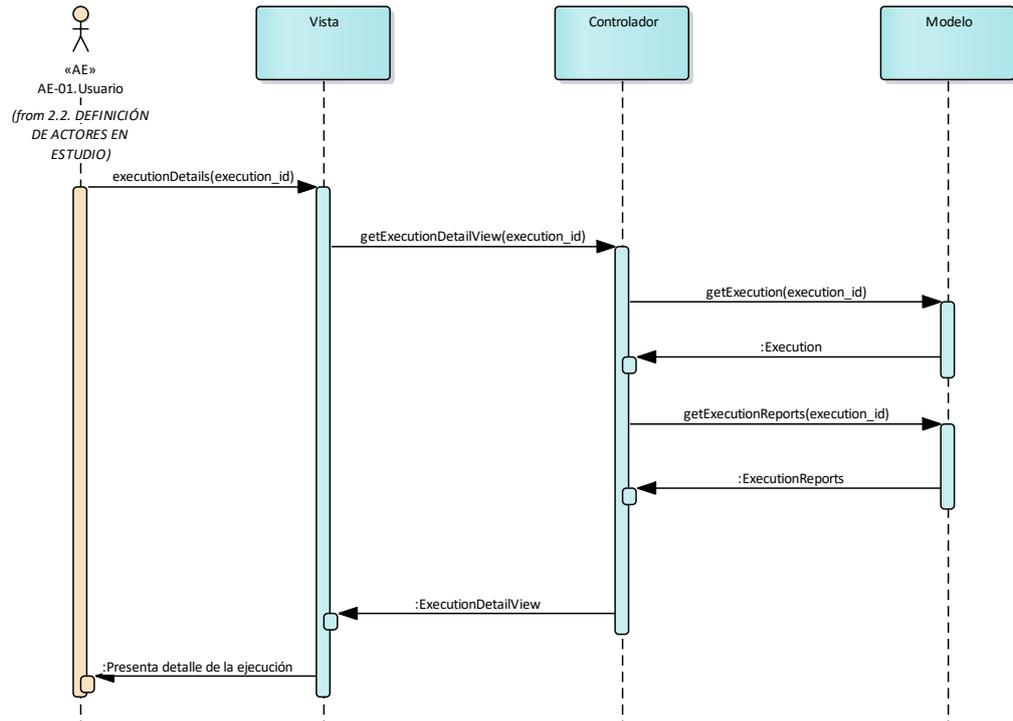


Figura 28: rf-01. ver detalle de una ejecución

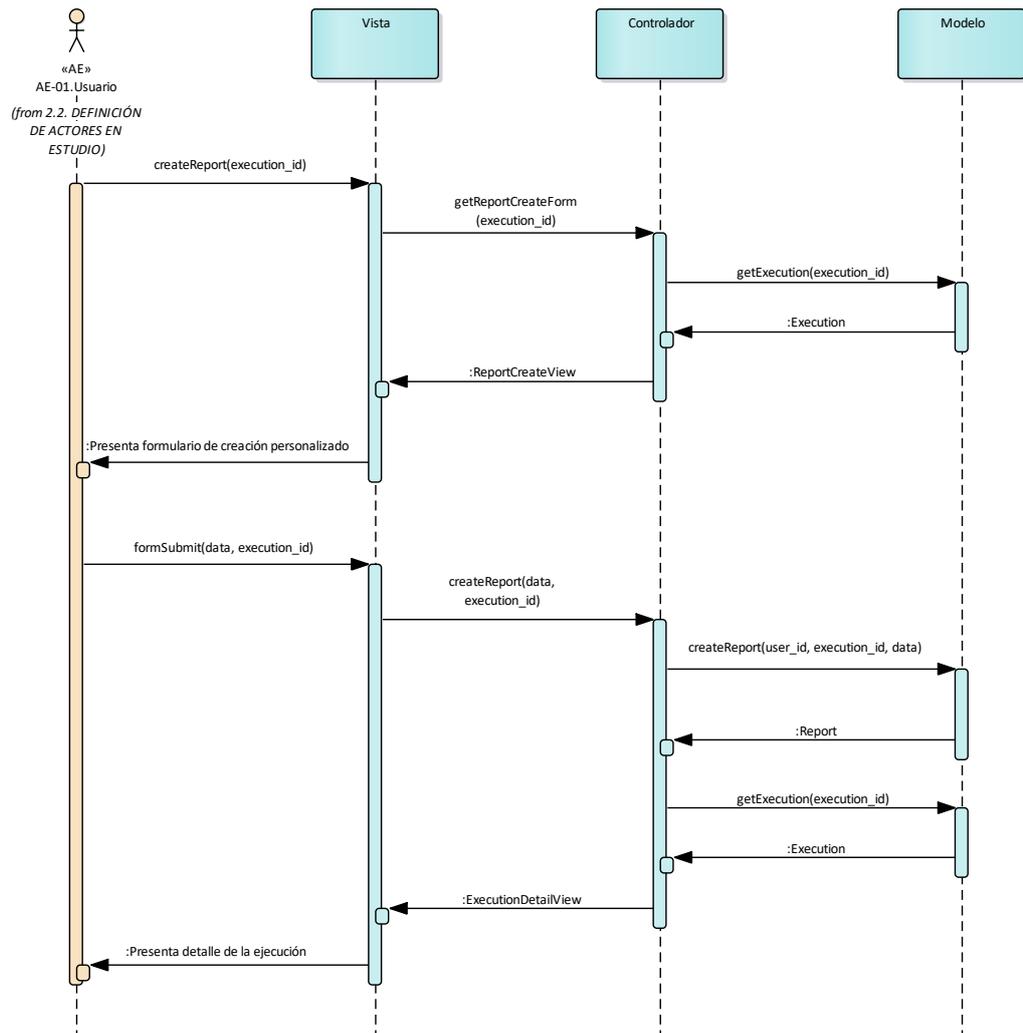


Figura 29: rf-02.crear reporte de una ejecución

6.3 Diseño

6.3.2 Arquitectura

Se detalla a continuación con que patrón de arquitectura se identifica este proyecto y cuál es su funcionamiento.

6.3.2.1 MVC

El Modelo Vista Controlador es un patrón arquitectónico que fundamenta la estructura MTV de Django, por lo que su comprensión resulta clave para entender la operativa de dicha plataforma. Este patrón se caracteriza por la división de la aplicación en tres niveles distintos, cada uno con responsabilidades claramente definidas que permiten separar la interfaz de usuario de la lógica de negocio y la gestión de datos.

Los niveles de este patrón arquitectónico son:

Capa de Presentación: Involucra todo el código que se encarga tanto de recoger los datos del usuario como de mostrarle los resultados. Este nivel se ocupa de:

- La navegación del sistema.
- La validación de los datos ingresados por los usuarios.
- El formateo de los datos para su presentación.
- La generación de las interfaces gráficas.

Capa de Negocio: Este nivel gestiona las operaciones clave que define el propósito funcional de la aplicación. Incluye:

- Las reglas y procesos de negocio.
- El manejo de las transacciones.
- Las decisiones operativas basadas en los datos de la aplicación.

Capa de Persistencia: Este estrato se encarga de todas las operaciones relacionadas con el almacenamiento y recuperación de datos. Las actividades típicas incluyen:

- Inserción de nuevos registros.
- Eliminación de registros existentes.
- Actualización de datos.
- Realización de consultas y recuperación de datos.

Esta organización en capas promueve una clara separación de responsabilidades, asegurando que cada componente del sistema se especialice en una sola función, lo que mejora la mantenibilidad y la escalabilidad del software.

6.3.2.2 MTV

El Framework Django se distingue por su arquitectura Modelo-Template-Vista (MTV), un diseño que automatiza el manejo de datos y cuyo núcleo operativo reside en las plantillas y las vistas. Este patrón arquitectónico es aplicado en la plataforma RPA y por tanto en su mejora, siguiendo los principios y estructura de Django.

Las tres capas que componen la arquitectura MTV de Django son:

Capa Modelo (Model): Representa la estructura de la base de datos dentro del framework. Los modelos en Django pueden generarse automáticamente con herramientas integradas en el framework, facilitando el acceso y la manipulación de datos a través de la API de Django. Este nivel permite interactuar con la base de datos de manera abstracta, evitando en muchos casos el uso directo de SQL.

Capa Template (Plantilla): Corresponde a la presentación de la información al usuario y está constituida principalmente por archivos HTML, CSS y JavaScript, a menudo enriquecidos con jQuery. Las plantillas en Django deben mantenerse libres de lógica del sistema, limitándose a presentar datos y recibir instrucciones del usuario mediante métodos GET y POST que interactúan con la capa de Vista.

Capa Vista (View): Funciona como el intermediario entre los modelos y las plantillas, encargándose de la lógica necesaria para recuperar datos del modelo y enviarlos a la plantilla adecuada. Además, gestiona las solicitudes y respuestas dentro del framework, estableciendo conexiones con la base de datos y definiendo cómo se relacionan las plantillas con las vistas a través de los archivos de configuración de URL y las vistas específicas, que dictan las operaciones necesarias para procesar y responder a las peticiones del usuario.

Este diseño de tres capas asegura una organización clara y una separación de responsabilidades, facilitando el desarrollo y mantenimiento de aplicaciones web complejas.

6.3.2.3 Patrón Modelo-Vista-Controlador MVC

El patrón Modelo-Vista-Controlador (MVC) estructura la arquitectura de un sistema en tres componentes principales, cada uno con funciones específicas, facilitando así la organización y la escalabilidad del sistema.

Los componentes de este patrón son:

Modelo: Constituye el núcleo de la información con la que el sistema opera. Este componente se encarga de gestionar los datos, incluyendo el acceso y las actualizaciones a la base de datos.

Controlador: Actúa como el intermediario entre el Modelo y la Vista. Su principal función es recibir las entradas del usuario y solicitar al Modelo que actualice su estado conforme a estas entradas.

Vista: Es el componente que presenta los datos al usuario, mostrando la información que reside en el Modelo.

El flujo de operaciones en el patrón MVC se desarrolla de la siguiente manera:

1. El usuario interactúa con la aplicación a través de la Vista, ejecutando acciones que son gestionadas por el Controlador.
2. El Controlador procesa estas acciones, modificando o consultando el estado del Modelo según sea necesario.
3. Después de que el Modelo ha sido actualizado o consultado, el Controlador dirige la respuesta a la Vista para que esta pueda actualizar la presentación al usuario según los cambios realizados.

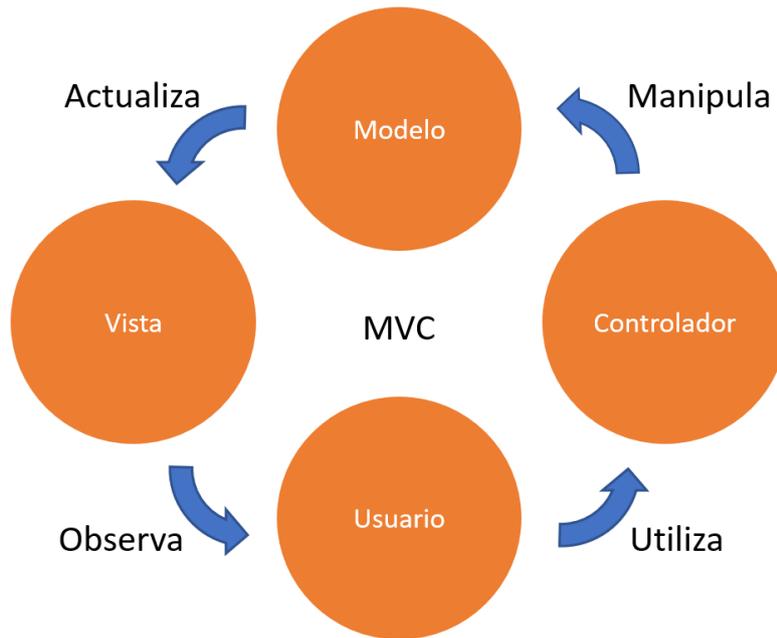


Figura 31: Arquitectura de MVC

Este ciclo de interacción garantiza que la arquitectura del sistema se mantenga clara y organizada, permitiendo una fácil mantenibilidad y la posibilidad de escalar el sistema conforme crezcan las necesidades del usuario. La Figura 40 ilustra todo el proceso de interacción y comunicación dentro de esta estructura MVC.

6.3.3 Análisis de alternativas tecnológicas

A continuación, se muestran cada una de las herramientas y tecnologías que se utilizarán de manera transversal en la mejora de la plataforma.

6.3.3.1 Control de versiones

Para el control de versiones se ha seleccionado Git, que es la herramienta más popular para esta función hasta la fecha. La elección se debe a la experiencia previa positiva del autor con Git en proyectos anteriores, lo que asegura una integración efectiva y eficiente en el desarrollo actual.

Git es un sistema de control de versiones distribuido y diseñado para manejar proyectos con gran cantidad de archivos de código fuente de manera eficiente y confiable. Proporciona funcionalidades completas para el seguimiento de cada cambio realizado en el código, lo que facilita la gestión de distintas versiones y la colaboración entre desarrolladores.



Figura 32. Git

En cuanto al alojamiento y manejo de versiones en un entorno remoto, se optará por GitHub. GitHub es una plataforma de desarrollo colaborativo de software que utiliza Git como su sistema de control de versiones. Al igual que Git, GitHub es ampliamente reconocido y utilizado por desarrolladores de todo el mundo, lo que facilita la colaboración y el seguimiento de proyectos.

Además de ser un gestor de repositorios, GitHub ofrece funcionalidades como alojamiento de wikis y un sistema de seguimiento de incidencias, lo que lo convierte en una herramienta integral para la gestión de proyectos de software. La plataforma también permite la creación de diferentes organizaciones y subgrupos dentro de las mismas, lo cual es esencial para administrar distintos niveles de acceso y roles dentro del equipo de desarrollo.



Figura 33. GitHub

6.3.3.2 Frontend

El desarrollo de la interfaz de la aplicación web se llevará a cabo utilizando HTML5, CSS3 y JavaScript. Estos lenguajes son la base estándar para la creación de aplicaciones web, ya que ofrecen las

herramientas y técnicas necesarias para el diseño del frontend, permitiendo una interacción usuario-sistema fluida y agradable.

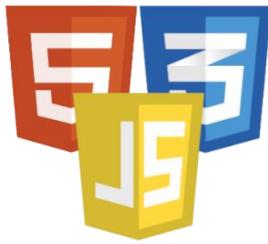


Figura 34. HTML5, CSS3 y Javascript.

HTML5 es un lenguaje de marcado que los navegadores interpretan para construir las páginas web. Este lenguaje trabaja en sinergia con CSS, lo que permite añadir estilos sofisticados y gestionar la presentación visual. Además, la integración con JavaScript enriquece las páginas con funcionalidades dinámicas, controlando eventos e interacciones en tiempo real.

En este contexto, también se incorporarán jQuery y Bootstrap, que son librerías y frameworks que extienden las capacidades de JavaScript y CSS. Estas herramientas son esenciales para intensificar la profesionalidad y el acabado estético de la aplicación, facilitando la creación de interfaces de usuario robustas y elegantes.

6.3.3.2.1 JQuery

Se toma jQuery [38] como solución para interactuar de manera versátil y extensible con documentos HTML. Esta biblioteca de JavaScript es compacta y rápida, dotada de un amplio abanico de funcionalidades que simplifican tareas como la manipulación de documentos HTML, gestión de eventos, animaciones y el uso de AJAX. Su API es intuitiva y está diseñada para ser compatible con la mayoría de los navegadores web. jQuery [39] ha revolucionado la manera en que millones de desarrolladores programan en JavaScript.



Figura 35. JQuery

6.3.3.2.2 Bootstrap

Bootstrap [40] es una herramienta fundamental para el diseño y la personalización ágil de sitios web, utilizando un conjunto de herramientas de código abierto. Es especialmente valorado por su capacidad de adaptación responsive, que asegura una óptima visualización en dispositivos móviles y tabletas. Ofrece una extensa gama de componentes predefinidos, incluyendo plugins de JavaScript, lo que lo convierte en la herramienta más popular en su categoría.



Figura 36. Bootstrap

Para esto, agrupa una serie de plantillas de diseño con tipografía, formularios, botones, cuadros, menús de navegación y otros elementos de diseño basado en HTML y CSS.

6.3.3.3 Backend

Django es un framework para desarrollo web que sigue patrón MVC y utiliza Python como lenguaje de programación. Este framework promueve un desarrollo rápido y un diseño de aplicaciones web limpio y práctico. Atiende algunas de las necesidades más comunes del desarrollo web, lo que permite que los desarrolladores se enfoquen en aspectos más específicos de la aplicación. Es de código abierto y gratuito. Se distingue por su capacidad para implementar herramientas avanzadas con una inversión mínima de recursos, además de su seguridad y escalabilidad. Facilita la implementación de diversas aplicaciones simultáneamente, ofreciendo varias opciones de arquitectura y modularidad.

django

Figura 37. Django

Entre sus principales características están:

- Escrito en Python: Por estar basado en Python, Django permite escribir código claro y sencillo, y facilita el desarrollo de aplicaciones rápidas y eficientes.
- Eficiencia: Django está diseñado para minimizar el tiempo desde el inicio del desarrollo hasta el despliegue de una aplicación web.
- Principio DRY (Don't Repeat Yourself): Este principio se aplica en Django para evitar la duplicidad en el código y fomentar su reutilización, disponiendo para ello de cientos de paquetes con variadas funcionalidades que van desde la seguridad de archivos hasta aplicaciones de modelos de machine learning.
- Administración incluida: Django incorpora un sistema de administración que viene activo por defecto, utilizable inmediatamente sin configuraciones adicionales.
- ORM (Mapeo objeto-relacional): Proporciona herramientas que permiten la interacción con la base de datos sin utilizar SQL directamente.

La adopción de Django en la plataforma se justifica por las ventajas que ofrece en el desarrollo, como el soporte a través de bibliotecas en Python que simplifican el manejo de datos y una amplia comunidad de desarrolladores que asegura soporte continuo y resolución de problemas. Además, Django incluye módulos fundamentales para funciones como el registro de usuarios y la gestión de bases de datos.

6.3.3.4 Base de datos

Para la gestión de datos en el framework Django, es fundamental contar con un sistema de almacenamiento que garantice la persistencia de la información. Se utiliza un modelo de datos híbrido que combina las ventajas de las bases de datos relacionales para dar soporte a diversas funciones como la gestión del sistema, administración de la plataforma, y registro y autenticación de usuarios. Este modelo se emplea para almacenar información con estructura fija y que requiere consultas frecuentes, características esenciales para la plataforma RPA.

Dentro de las opciones disponibles, PostgreSQL se destaca como la alternativa preferida por varias razones clave, que lo hacen idóneo para este proyecto:

- Compatibilidad con Django: PostgreSQL es la base de datos predeterminada para Django. El desarrollo del código en Django se realiza primero para PostgreSQL antes de adaptarlo a otras bases de datos.
- Módulo contrib de PostgreSQL: Este módulo se instala con Django y ofrece funcionalidades avanzadas. Aunque MySQL también ofrece un módulo similar, no es el principal.
- Gestión de cambios de esquema: PostgreSQL permite realizar cambios de esquema dentro de las transacciones, lo que facilita enormemente la implementación de migraciones en Django.
- Búsqueda de texto completo: PostgreSQL incluye capacidades de búsqueda de texto completo integradas, ofreciendo un mejor rendimiento que MySQL.
- Extensión PostGIS: Esta extensión es estándar para GeoDjango, que maneja datos geográficos y mapas.
- Soporte avanzado para JSON: PostgreSQL maneja eficazmente el almacenamiento y consulta de JSON con JSONB, una funcionalidad también presente en versiones recientes de MySQL.
- HStore: Permite un almacenamiento de clave-valor sencillo.
- Índices parciales: Facilitan la implementación rápida de restricciones de unicidad directamente en la base de datos.
- Autoincremento de identificadores: PostgreSQL utiliza sequences para incrementar automáticamente los identificadores de claves primarias.
- Foreign data wrappers: Esta funcionalidad permite a PostgreSQL interactuar con diversas fuentes de datos, incluidas otras bases de datos y sistemas como Redis.
- Tipado fuerte: PostgreSQL es estricto en el tipado de datos, rechazando entradas que no coincidan con el tipo esperado, a diferencia de MySQL que puede emitir solo una advertencia.
- PostgreSQL se distribuye bajo una licencia de código abierto, similar a las licencias BSD o MIT, lo que facilita su uso y distribución.



Figura 38. PostgreSQL.

6.3.4 Prototipos

En esta sección de la memoria, se detallan los prototipos diseñados para las interfaces de usuario correspondientes a distintos casos de uso. Dado que algunos prototipos son bastante similares entre sí, solo se incluyen aquellos que ofrecen una perspectiva diferente y son representativos de las mejoras que se desean implementar.

En particular, para la visualización de la configuración de las fases de una ejecución se presentan dos prototipos: los correspondientes a las configuraciones de las fases "Log processing" y "Extract training dataset", omitiendo los prototipos detallados de las configuraciones de las otras seis fases del pipeline.

De manera similar, para la visualización de resultados cuyos contenidos son datasets se indica solo el prototipo de la fase de "Log Processing", dado que los prototipos de las otras cinco fases resultan prácticamente idénticos.

A continuación, se incluye el prototipo de una interfaz de usuario diseñada para la vista en detalle de una ejecución. En la sección superior, se encuentran los detalles de configuración de la ejecución, incluyendo título, descripción, el caso de estudio asociado y el tiempo total de ejecución. En la parte central de la interfaz, se visualiza el 'pipeline' de ejecución con las diferentes fases. Las fases no aplicadas en la ejecución y, por lo tanto, inactivas, están marcadas en gris, mientras que las fases que se han llevado a cabo están resaltadas en color y son interactivas mediante botones. Estos botones permiten acceder a la configuración específica y a los resultados obtenidos en cada fase. Finalmente, en la parte inferior de la interfaz, se listan los informes generados, los cuales ofrecen información detallada sobre la ejecución. Los usuarios pueden interactuar con estos informes para ver detalles de su configuración, descargarlos o eliminarlos. Además, la interfaz cuenta con un botón para generar un nuevo reporte con la información de dicha ejecución.

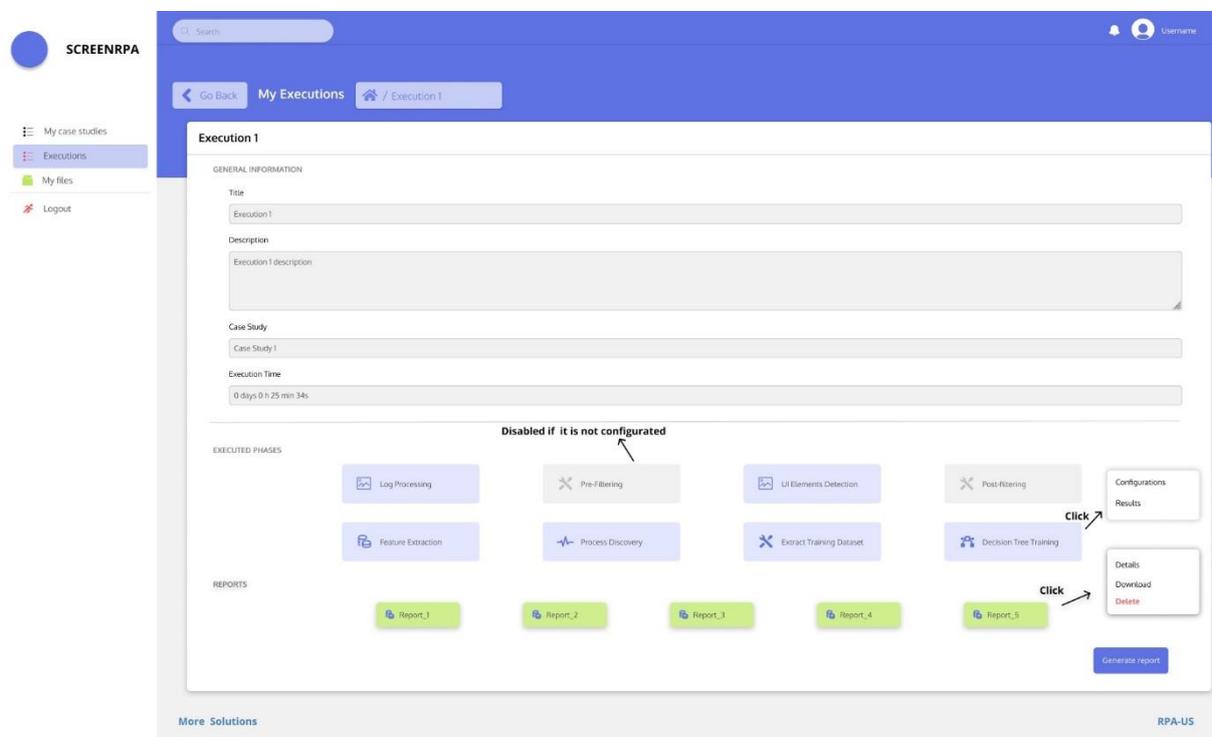


Figura 39. Prototipo de vista en detalle de una ejecución.

También se incluye la interfaz de usuario diseñada para la vista en detalle de la configuración de la fase monitorización del comportamiento. En la parte superior se encuentra información general sobre la configuración, como el título y el tipo de log de entrada. En la parte central se encuentran los parámetros de configuración específicos para el procesamiento del log, apareciendo, por ejemplo, el nombre del archivo de log y el separador de columnas utilizado. Por último, en la parte inferior se detallan configuraciones adicionales en formato JSON, que están personalizadas más allá de los parámetros estándar.

The screenshot displays the 'Log Processing Configuration' page in the SCREENRPA application. The page has a blue header with a search bar and user profile. A sidebar on the left contains navigation options: 'My case studies', 'Executions', 'My files', and 'Logout'. The main content area is titled 'Log Processing Configuration' and is divided into three sections:

- GENERAL INFORMATION:** Contains a 'Title' text input field and a 'Type' dropdown menu.
- CONFIGURATION:** Contains several input fields:
 - 'URI Log Filename' with the value 'Recording_20230424_1022.mim'.
 - 'Column Separator' (empty).
 - 'Gaze Log' with the value 'ET_REXAPI-GazeAnalysis.csv'. A note above it says 'Greyed out if no gaze info in case study'.
 - 'Gaze Log Adjustment' with the value '0'.
 - 'System Info. Log' with the value 'Native_SlideEvents.csv'.
- Additional Configuration (JSON):** A text area containing a JSON snippet:

```
{  
  "param1": "value1",  
  "param2": "value2"  
}
```

At the bottom of the page, there is a 'More Solutions' link on the left and the 'RPA-US' logo on the right.

Figura 40. Prototipo de vista en detalle de configuración de la fase “Log Processing”.

Se incluye, además, la interfaz de usuario diseñada para la vista en detalle de la configuración de la fase de extracción de un dataset. En la parte superior se encuentra el título de dicha ejecución. En la parte central se encuentran los parámetros de configuración específicos para la extracción del “dataset” de entrenamiento de los algoritmos, incluyendo, por ejemplo, el puntos de decisión a estudiar. De nuevo, en la parte inferior se detallan configuraciones adicionales en formato JSON, que contienen configuraciones personalizadas más allá de los parámetros estándar.

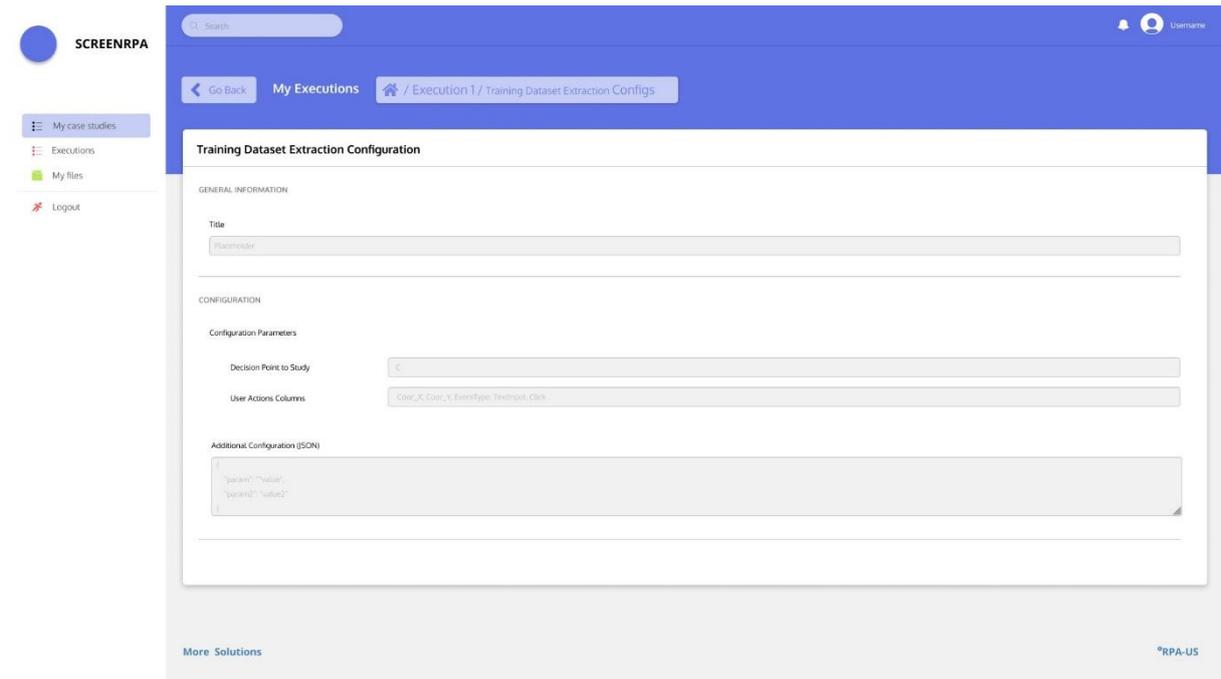


Figura 41. Prototipo de vista en detalle de configuración de la fase “Extract Training Dataset”.

Se incluye una vista para interpretar de forma visual los resultados de la fase de monitorización del comportamiento en forma de Log. En la parte superior, hay un menú desplegable que permite seleccionar el escenario cuyos resultados se desean visualizar. En el centro de la interfaz se muestra una tabla con los resultados del escenario seleccionado, la cual puede manipularse para facilitar su visualización, desplazándose hacia abajo para ver más filas o hacia los lados para ver más columnas.

Finalmente, en la parte inferior, la interfaz incluye un botón para descargar los resultados del escenario específico seleccionado en formato CSV.

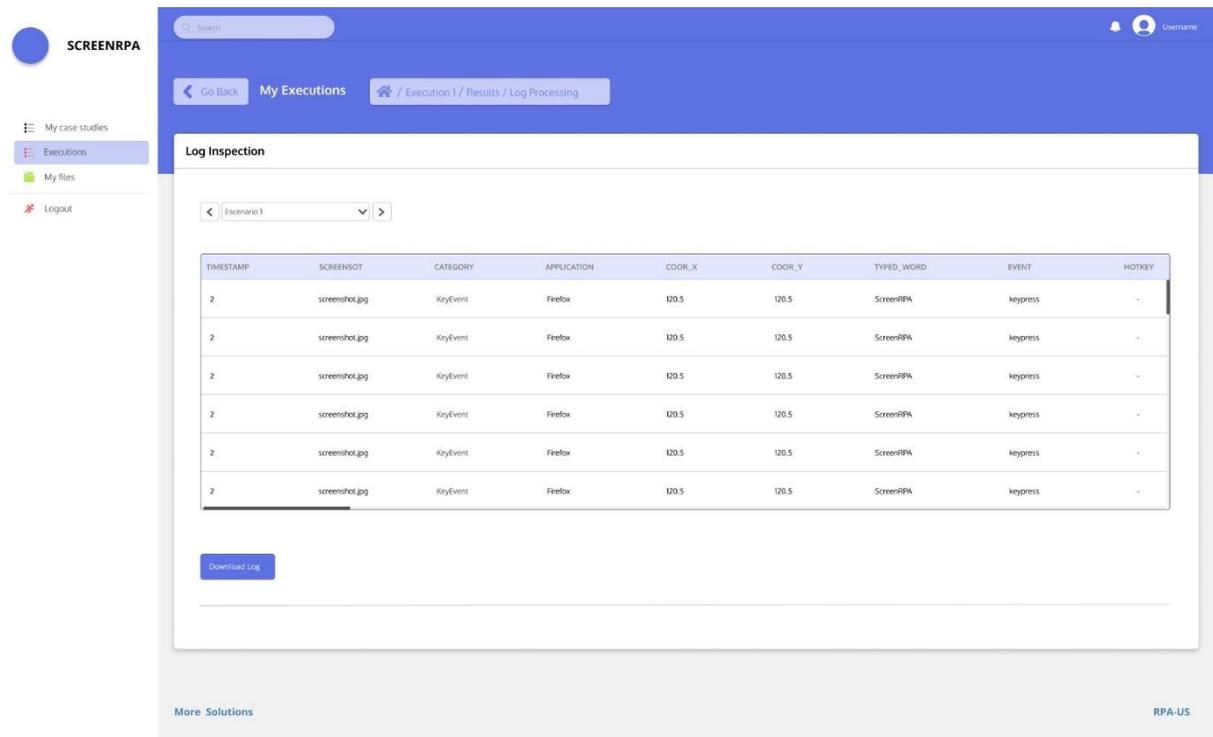


Figura 42. Prototipo de vista en detalle de resultados de la fase “Log Processing”.

Se incluye una vista para interpretar de forma visual los resultados de la fase de descubrimiento de proceso en forma de diagrama de procesos BPMN. En la parte superior, hay un menú desplegable que permite seleccionar el escenario cuyos resultados se desean visualizar. En el centro de la interfaz se muestra el diagrama BPMN correspondiente al resultado del escenario seleccionado.

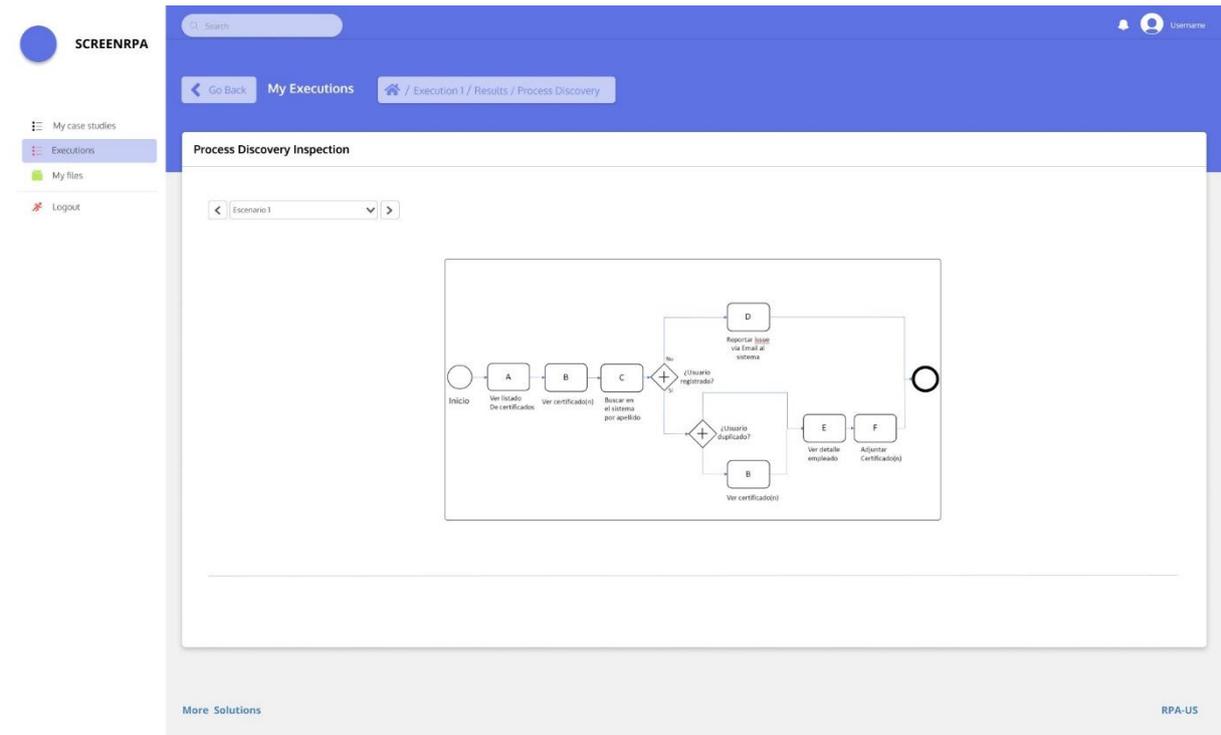


Figura 43. Prototipo de vista en detalle de resultados de la fase “Process Discovery”.

Se incorpora una vista para interpretar de forma visual los resultados de la fase de entrenamiento de un árbol de decisión. En la parte superior, hay un menú desplegable que permite seleccionar el escenario cuyos resultados se desean visualizar. En el centro de la interfaz se muestra el árbol de clasificación correspondiente al resultado del escenario seleccionado.

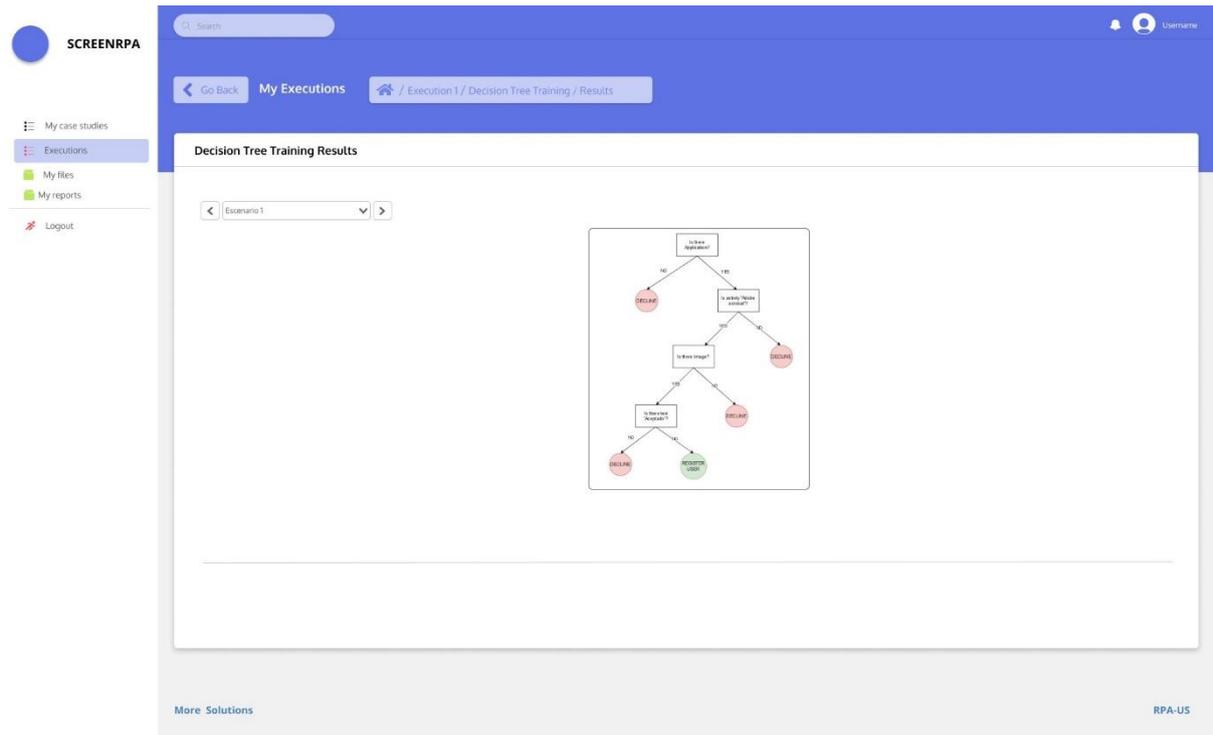


Figura 44. Prototipo de vista en detalle de resultados de la fase “Extract Training Dataset”.

Se incorpora también una vista para configurar la generación de reportes basada en los resultados de una ejecución, permitiendo personalizar su contenido. En la sección superior, se encuentran áreas dedicadas a añadir información general sobre el reporte, así como su propósito y objetivos. En la parte central de la interfaz, se dispone de un 'checkboxlist' que permite seleccionar los elementos a incluir en el reporte. Dependiendo de las fases realizadas durante la ejecución, las opciones disponibles en el 'checkboxlist' variarán para adaptarse al tipo de contenido que puede generarse con la información obtenida. En la parte inferior de la interfaz, se encuentra un botón para generar el informe de la ejecución con las configuraciones especificadas.

The screenshot shows the 'Generate Report Configuration' page. At the top, there is a search bar and a user profile icon. Below that, a navigation bar includes 'Go Back', 'My Executions', and 'Execution 1 / Generate Report'. The main content area is titled 'Generate Report Configuration' and is divided into two sections. The first section, 'INTRODUCTION', has two text input fields labeled 'Purpose' and 'Objective'. The second section, 'AS IS PROCESS DESCRIPTION', contains a checklist with five items, each with a checkbox and a corresponding output: 'Process Overview' (unchecked, 'Blog: Process Discovery'), 'Applications Used' (checked, 'No blog'), 'AS IS Process Map' (checked, 'Blog: Process Discovery'), 'Detailed As Is Process Actions' (checked, 'Blog: Process Discovery'), and 'Input Data Description' (checked, 'No blog'). A 'Generate' button is located at the bottom right of the configuration area.

Figura 45. Prototipo de vista en detalle de configuración para la generación de reporte.

Finalmente, se incluye en los prototipos de interfaz una vista en detalle de la configuración de generación de un reporte. La vista se divide en dos partes principales: en la sección superior, se presentan los detalles de la información general sobre el reporte, incluyendo su propósito y objetivos. En la parte inferior de la interfaz, se muestra un “checklist” con los campos deshabilitados que reflejan el contenido del reporte seleccionado a través de los títulos del índice.

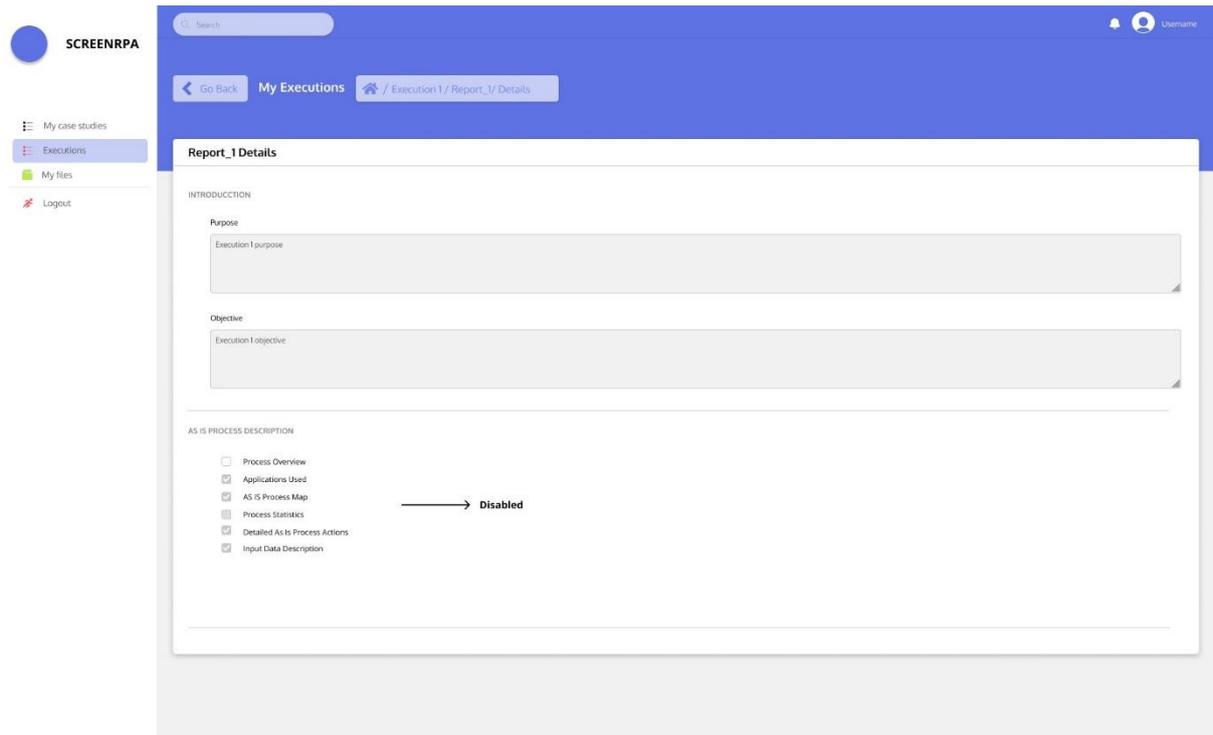


Figura 46. Prototipo de vista en detalle de configuración de reporte.

6.4 Implementación

6.4.1 Entorno de desarrollo

Se exponen los elementos y herramientas utilizados para el desarrollo de la mejora de la plataforma.

6.4.1.1 Visual Code

Visual Studio Code (VS Code) es un entorno de desarrollo integrado (IDE) ligero pero muy poderoso, desarrollado por Microsoft. Es ampliamente utilizado para programar en una variedad de lenguajes, incluido Python, gracias a su extensible soporte mediante extensiones. Algunas de las funcionalidades clave de Visual Studio Code incluyen:

- **Edición de Código Avanzada:** Incluye soporte para la depuración, realce de sintaxis, autocompletado inteligente (IntelliSense), refactorización de código y vista previa de definiciones sin necesidad de abrir el archivo.
- **Soporte Integrado para Git:** Permite realizar operaciones de Git directamente desde el editor, incluyendo commit, pull, push, y otras funcionalidades esenciales de control de versiones.
- **Nuevas funcionalidades mediante Extensiones:** Los usuarios pueden personalizar y ampliar las capacidades de VS Code con extensiones disponibles en el Marketplace, lo que permite agregar soporte para una variedad de lenguajes, herramientas y frameworks, incluyendo Python y Django.

- **Depuración Integrada:** Ofrece un depurador potente que soporta la configuración de puntos de interrupción, inspección de variables, visualización de llamadas a funciones y muchas otras características esenciales para la depuración eficiente de aplicaciones.
- **Terminal Integrado:** Viene con un terminal integrado que permite a los usuarios ejecutar comandos de shell directamente desde el IDE, facilitando tareas como la gestión de paquetes, ejecución de scripts y manejo de versiones.
- **Multiplataforma:** Compatible con Windows, macOS y Linux, lo que lo convierte en una opción versátil para desarrolladores en diversas plataformas.

6.4.1.2 Python

Se emplea la versión 3.10.7 del lenguaje de programación Python. Esta versión facilita la gestión y ejecución de archivos .py y es crucial para el uso del framework web Django.

Python es un lenguaje de programación interpretado, de alto nivel y multiparadigma, que soporta programación orientada a objetos, funcional y procedural. Es conocido por su sintaxis clara y legible, lo que facilita la escritura y mantenimiento del código.

6.4.1.3 Django 4.1.2

Se utiliza la versión 4.1.2 de Django, un framework de alto nivel para Python diseñado para el desarrollo rápido y limpio de proyectos web. Django simplifica la creación de aplicaciones web robustas, proporcionando una estructura que incluye gestión de base de datos, sistema de plantillas (Django Templates), y herramientas preconstruidas para manejar la lógica y presentación de datos. El framework se integra en todos los niveles de la aplicación, apoyando tanto la definición de modelos como la lógica de negocio y la interfaz de usuario. Django sigue el patrón de arquitectura Modelo-Vista-Controlador (MVC), aunque en la terminología de Django, se refiere a Modelo-Template-Vista (MTV).

6.4.1.4 Django Template

Django Template es el motor de plantillas incorporado de Django, utilizado para crear interfaces de usuario dinámicas y eficientes. Facilita la separación entre la lógica de la aplicación y la presentación de los datos, permitiendo a los desarrolladores y diseñadores trabajar de manera más eficiente y flexible. Los templates se encargan de presentar la información en formato HTML, facilitando la integración con tecnologías de frontend. El código que se incrusta en las plantillas de Django se ejecuta en el servidor, generando HTML dinámico antes de enviarlo al cliente. Esto permite que los datos dinámicos y la lógica de negocio sean procesados en el servidor, asegurando una mayor seguridad y eficiencia.

6.4.1.5 Gestor de paquetes Anaconda

Anaconda es utilizado como gestor de paquetes y entornos virtuales para Python, proporcionando una plataforma robusta para la gestión de dependencias y entornos de desarrollo. Este gestor facilita la instalación de paquetes desde repositorios comunes y la creación de entornos reproducibles, asegurando la coherencia entre los entornos de desarrollo y producción.

Anaconda es especialmente útil en proyectos de ciencia de datos, pero su utilidad se extiende a cualquier tipo de proyecto Python, incluido la plataforma.

6.4.1.6 PgAdmin

PgAdmin sirve como la herramienta de gestión para bases de datos PostgreSQL en CRPAsite. Este software libre se distribuye bajo la licencia PostgreSQL/Artistic y proporciona una interfaz gráfica para la administración de bases de datos PostgreSQL. Facilita la ejecución de consultas, administración de esquemas, configuración de permisos y otras tareas de gestión de base de datos necesarias para el mantenimiento y desarrollo del proyecto.

6.4.2 Elección de librerías

Con el objetivo de cubrir los requisitos que se han detallado en los puntos anteriores, de la mejor forma posible, se ha llevado a cabo una investigación sobre distintas librerías disponibles para resolver los problemas que se han presentado. Para ello se describirán en los siguientes apartados las principales librerías escogidas.

6.4.2.1 Scikit-learn

Para el manejo y análisis de árboles de decisión, se ha optado por utilizar Scikit-learn [42], una librería robusta de aprendizaje automático para Python. Esta librería ofrece una amplia gama de algoritmos estadísticos y de aprendizaje, incluyendo clasificación, regresión, agrupamiento y reducción de dimensionalidad.

6.4.2.2 Python-docx

La librería Python-docx [43] se ha seleccionado para la creación de reportes en formato DOCX. Esta herramienta permite la manipulación de documentos Word desde Python, facilitando la creación, modificación y el almacenamiento de documentos complejos de manera programática.

6.4.2.3 Graphviz

Graphviz [44] es una herramienta de visualización gráfica que se utiliza para representar estructuras de información estructurada como diagramas de grafos. En nuestro proyecto, Graphviz se utiliza para visualizar árboles de clasificación y modelos de diagramas procesos BPMN (Business Process Model and Notation) proporcionando una representación clara y comprensible de modelos complejos.

6.4.2.4 Pydotplus

Pydotplus [45] es una librería de Python que sirve de interfaz para Graphviz, facilitando la creación de gráficos de forma programática. En combinación con Graphviz, permite una manipulación más directa y detallada de los gráficos desde código Python, lo que es particularmente útil para tareas que requieren la generación automática de visualizaciones complejas. Pydotplus se integra de manera efectiva en nuestro flujo de trabajo para mejorar la visualización de árboles de clasificación, permitiendo una personalización extensa y facilitando la integración con otros componentes Python del proyecto.

6.4.3 Pruebas

En este proyecto se han desarrollado 45 pruebas, todas con resultados satisfactorios. Estas pruebas, estructuradas como las tablas que se presentan a continuación, fueron seleccionadas por ser las más representativas de los casos de uso destacados en las secciones "Definición de casos de uso del sistema" y "Definición de clases de procesos". Debido al volumen de las pruebas, se ha optado por incluirlas en un documento separado titulado "Memoria técnica - Anexo Pruebas.docx". Para acceder a la descripción completa de todas las pruebas, se puede consultar este archivo, que se encuentra en el mismo directorio que este documento.

PS-39.Crear reporte de una ejecución		
Versión	1.0	25/04/2024 0:00:00
Autores	NDT - Driver 2.0	
Conexiones	Fuente	Destino
	PS-39.Crear reporte de una ejecución	PS-43.Ver detalle de una ejecución
	RF-02.Crear reporte de una ejecución	PS-39.Crear reporte de una ejecución
	AC-01.Usuario	PS-39.Crear reporte de una ejecución
	Mostrar detalles de la ejecución	PS-39.Crear reporte de una ejecución
Descripción	Esta Prueba de Sistema probará el caso de uso'RF-02.Crear reporte de una ejecución\' cuya descripción es:El objetivo de este requisito es permitir al usuario configurar, y con esta configuración generar un reporte de una ejecución de un caso de estudio.	
Pre-condición	Post-condición: <i>Creación</i> . Se crea un nuevo reporte con la configuración escogida	
Post-condición	Pre-condición: <i>Caso de Estudio</i> . El usuario tiene algún caso de estudio asociado Pre-condición: <i>Sesión iniciada</i> . El usuario ha iniciado sesión	
Estado	Propuesto	

Tabla 53: PS-39.Crear reporte de una ejecución

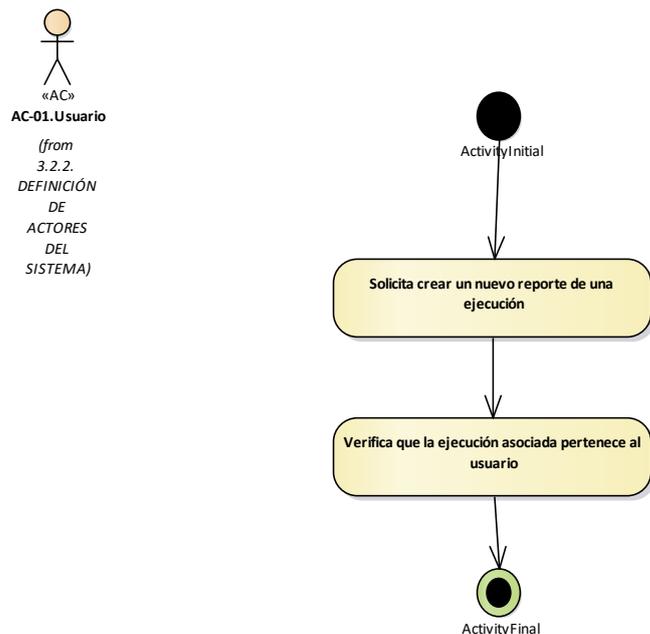


Figura 47: PS-40.Crear reporte de una ejecución

PS-40.Crear reporte de una ejecución		
Versión	1.0	25/04/2024 0:00:00
Autores	NDT - Driver 2.0	
Conexiones	Fuente	Destino
	RF-02.Crear reporte de una ejecución	PS-40.Crear reporte de una ejecución
	AC-01.Usuario	PS-40.Crear reporte de una ejecución
Descripción	Esta Prueba de Sistema probará el caso de uso'RF-02.Crear reporte de una ejecución\' cuya descripción es:El objetivo de este requisito es permitir al usuario configurar, y con esta configuración generar un reporte de una ejecución de un caso de estudio.	
Pre-condición	Post-condición: <i>Creación</i> . Se crea un nuevo reporte con la configuración escogida	
Post-condición	Pre-condición: <i>Sesión iniciada</i> . El usuario ha iniciado sesión Pre-condición: <i>Caso de Estudio</i> . El usuario tiene algún caso de estudio asociado	
Estado	Propuesto	

Tabla 54: PS-40.Crear reporte de una ejecución

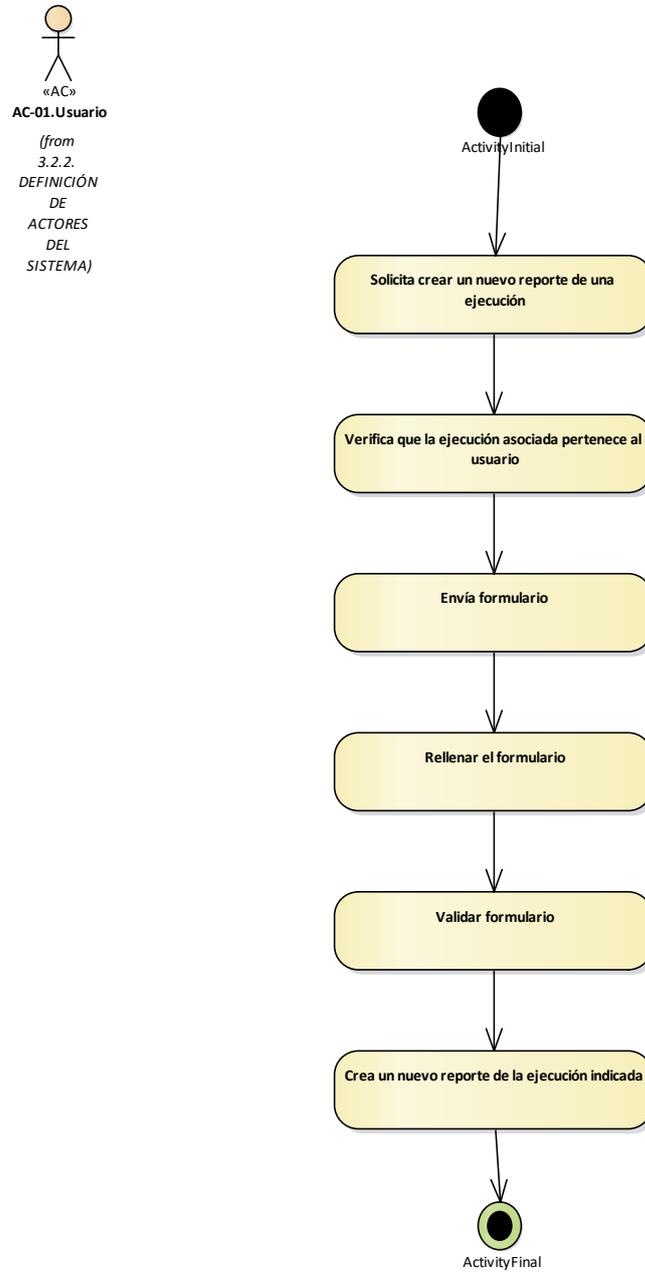


Figura 48: PS-41. Crear reporte de una ejecución

PS-41.Crear reporte de una ejecución		
Versión	1.0	25/04/2024 0:00:00
Autores	NDT - Driver 2.0	
Conexiones	Fuente	Destino
	AC-01.Usuario	PS-41.Crear reporte de una ejecución
	RF-02.Crear reporte de una ejecución	PS-41.Crear reporte de una ejecución
Descripción	Esta Prueba de Sistema probará el caso de uso'RF-02.Crear reporte de una ejecución\' cuya descripción es:El objetivo de este requisito es permitir al usuario configurar, y con esta configuración generar un reporte de una ejecución de un caso de estudio.	
Pre-condición y Post-condición	Post-condición: <i>Creación</i> . Se crea un nuevo reporte con la configuración escogida Pre-condición: <i>Caso de Estudio</i> . El usuario tiene algún caso de estudio asociado Pre-condición: <i>Sesión iniciada</i> . El usuario ha iniciado sesión	
Estado	Propuesto	

Tabla 55: PS-41.Crear reporte de una ejecución

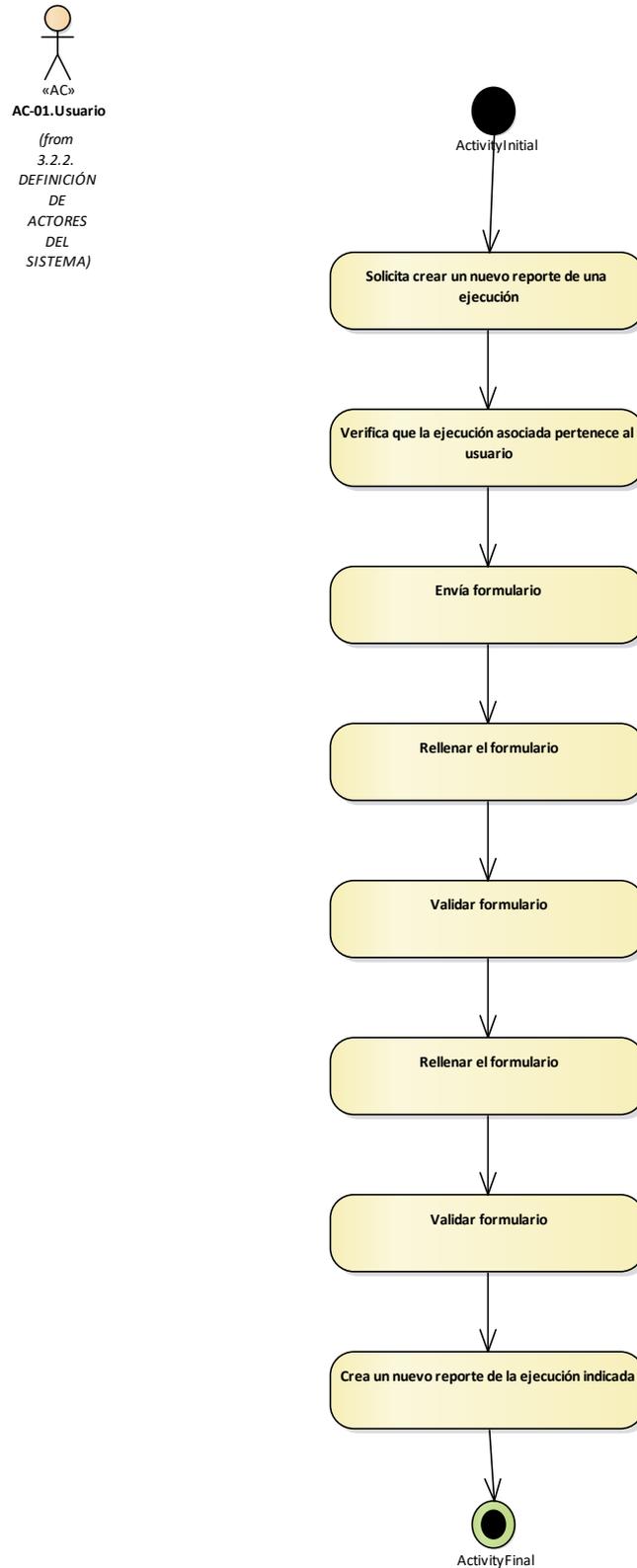


Figura 49: PS-42.Crear reporte de una ejecución

PS-42.Crear reporte de una ejecución		
Versión	1.0	25/04/2024 0:00:00
Autores	NDT - Driver 2.0	
Conexiones	Fuente	Destino
	AC-01.Usuario	PS-42.Crear reporte de una ejecución
	RF-02.Crear reporte de una ejecución	PS-42.Crear reporte de una ejecución
Descripción	Esta Prueba de Sistema probará el caso de uso'RF-02.Crear reporte de una ejecución\' cuya descripción es:El objetivo de este requisito es permitir al usuario configurar, y con esta configuración generar un reporte de una ejecución de un caso de estudio.	
Pre-condición y Post-condición	Post-condición: <i>Creación</i> . Se crea un nuevo reporte con la configuración escogida Pre-condición: <i>Caso de Estudio</i> . El usuario tiene algún caso de estudio asociado Pre-condición: <i>Sesión iniciada</i> . El usuario ha iniciado sesión	
Estado	Propuesto	

Tabla 56: PS-42.Crear reporte de una ejecución

PS-43.Ver detalle de una ejecución		
Versión	1.0	25/04/2024 0:00:00
Autores	NDT - Driver 2.0	
Conexiones	Fuente	Destino
	PS-43.Ver detalle de una ejecución	PS-33.Listar reportes de una ejecución
	PS-16.Descargar resultado de la fase de una ejecución	PS-43.Ver detalle de una ejecución
	RF-01.Ver detalle de una ejecución	PS-43.Ver detalle de una ejecución
	PS-19.Ver resultado de las fases de una ejecución	PS-43.Ver detalle de una ejecución
	AC-01.Usuario	PS-43.Ver detalle de una ejecución
	PS-24.Ver configuraciones de las fases de una ejecución	PS-43.Ver detalle de una ejecución
	PS-39.Crear reporte de una ejecución	PS-43.Ver detalle de una ejecución
Descripción	Esta Prueba de Sistema probará el caso de uso'RF-01.Ver detalle de una ejecución\' cuya descripción es:El usuario debe tener acceso a una vista de los detalles de una ejecución, vista central desde donde se centralizará la manipulación de los reportes y resultados asociados a esa ejecución.	
Pre-condición y Post-condición	Post-condición: <i>Detalles</i> . Se redirige al usuario a la vista de detalles del caso de estudio Pre-condición: <i>Casos de Estudio</i> . El usuario debe tener al menos un caso de estudio Pre-condición: <i>Sesión Iniciada</i> . El usuario debe de haber iniciado sesión	
Estado	Propuesto	

Tabla 57: PS-43.Ver detalle de una ejecución

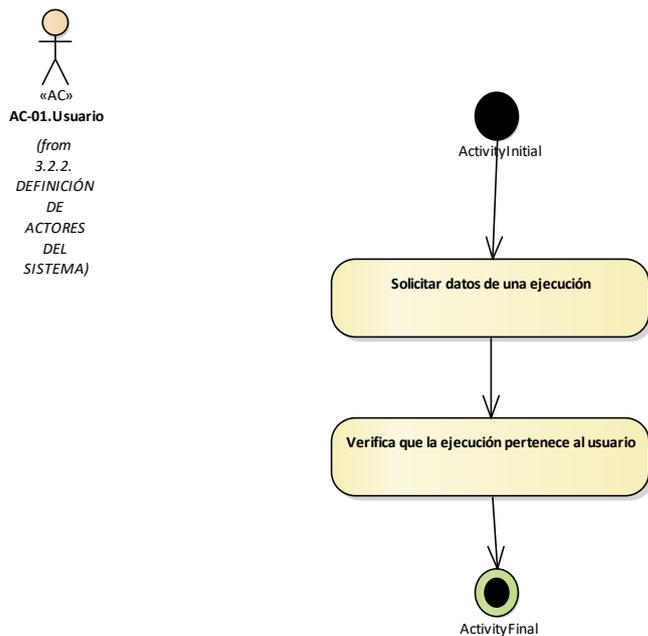


Figura 50. PS-44.Ver detalle de una ejecución

PS-44.Ver detalle de una ejecución		
Versión	1.0	25/04/2024 0:00:00
Autores	NDT - Driver 2.0	
Conexiones	Fuente	Destino
	RF-01.Ver detalle de una ejecución	PS-44.Ver detalle de una ejecución
	AC-01.Usuario	PS-44.Ver detalle de una ejecución
Descripción	Esta Prueba de Sistema probará el caso de uso 'RF-01.Ver detalle de una ejecución' \ cuya descripción es: El usuario debe tener acceso a una vista de los detalles de una ejecución, vista central desde donde se centralizará la manipulación de los reportes y resultados asociados a esa ejecución.	
Pre-condición	Post-condición: <i>Detalles</i> . Se redirige al usuario a la vista de detalles del caso de estudio	
Post-condición	Pre-condición: <i>Sesión Iniciada</i> . El usuario debe de haber iniciado sesión	
	Pre-condición: <i>Casos de Estudio</i> . El usuario debe tener al menos un caso de estudio	
Estado	Propuesto	

Tabla 58: PS-44.Ver detalle de una ejecución

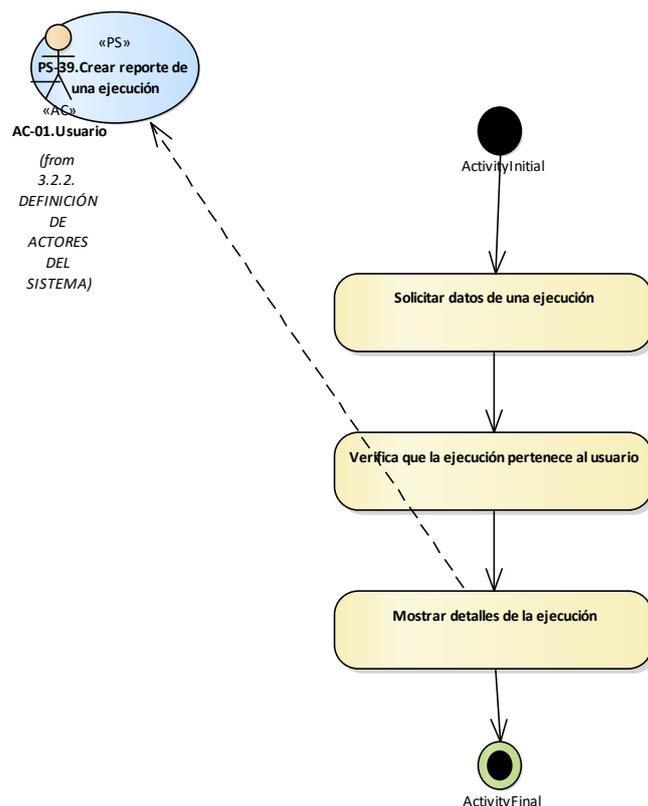


Figura 51: PS-45.Ver detalle de una ejecución

PS-45.Ver detalle de una ejecución		
Versión	1.0	25/04/2024 0:00:00
Autores	NDT - Driver 2.0	
Conexiones	Fuente	Destino
	RF-01.Ver detalle de una ejecución	PS-45.Ver detalle de una ejecución
	AC-01.Usuario	PS-45.Ver detalle de una ejecución
Descripción	Esta Prueba de Sistema probará el caso de uso 'RF-01.Ver detalle de una ejecución' cuya descripción es: El usuario debe tener acceso a una vista de los detalles de una ejecución, vista central desde donde se centralizará la manipulación de los reportes y resultados asociados a esa ejecución.	
Pre-condición y Post-condición	Post-condición: <i>Detalles</i> . Se redirige al usuario a la vista de detalles del caso de estudio Pre-condición: <i>Sesión Iniciada</i> . El usuario debe de haber iniciado sesión Pre-condición: <i>Casos de Estudio</i> . El usuario debe tener al menos un caso de estudio	
Estado	Propuesto	

Tabla 59: PS-45.Ver detalle de una ejecución

7 Conclusiones

El trabajo realizado en la plataforma aporta un gran valor tanto para los desarrolladores como para los analistas, facilitándoles el trabajo y proporcionando mayor comprensibilidad a una especialidad con un alto componente técnico.

7.1 Lecciones aprendidas

Durante el transcurso del trabajo, he aprendido que trabajar en una plataforma como ScreenRPA, que es compleja a nivel técnico y está en continuo desarrollo, es todo un reto. Además de la programación de las funcionalidades, es crucial tener en cuenta el trabajo en equipo y un constante control de versiones. Inicialmente, no consideraba la importancia de estos aspectos, pero a medida que el proyecto iba avanzando, se volvieron vitales para el éxito de este.

7.2 Consecución de los objetivos del proyecto

Se han alcanzado los objetivos iniciales planteados, y se ha descubierto que el estudio de las decisiones ha servido para conocer también el grado de certeza respecto a las variantes estudiadas dentro del RPA.

7.3 Trabajos futuros

A partir de lo que se ha conseguido con este trabajo, se abren líneas futuras como la extracción de fragmentos de código a partir del PDD, la optimización de los procesos actuales, y la incorporación de nuevas funcionalidades basadas en las necesidades emergentes de los usuarios y las tendencias del mercado.

8 Apéndices

8.1 Glosario de términos

- **Alcance del Proyecto.** El trabajo realizado para entregar un producto, servicio o resultado con las funciones y características especificadas. [41]
- **Diagrama de Gantt.** Un diagrama de barras con información del cronograma donde las actividades se enumeran en el eje vertical, las fechas se muestran en el eje horizontal y las duraciones de las actividades se muestran cómo barras horizontales colocadas según las fechas de inicio y finalización. [41]

- **Dirección de proyectos.** La aplicación de conocimientos, habilidades, herramientas y técnicas a actividades del proyecto para cumplir con los requisitos de este. [41]
- **Duración.** El total de períodos de trabajo (sin incluir vacaciones u otros períodos no laborales) requeridos para terminar una actividad del cronograma o un componente de la estructura de desglose del trabajo. Generalmente, se expresa en jornadas o semanas laborales. A veces se equipara incorrectamente al tiempo transcurrido. [41]
- **Estimado.** Una evaluación cuantitativa del monto o resultado probable. Habitualmente se aplica a los costos, recursos, esfuerzo y duraciones de los proyectos y normalmente va seguido de un modificador (p.ej., preliminar, conceptual, de factibilidad, de orden de magnitud, definitivo). Siempre debería incluir alguna indicación de exactitud (p.ej., $\pm x$ por ciento). [41]
- **Estimar la Duración de las Actividades.** El proceso de establecer aproximadamente la cantidad de períodos de trabajo necesarios para finalizar actividades individuales con los recursos estimados. [41]
- **Estimar los Costos.** El proceso de desarrollar una aproximación de los recursos monetarios necesarios para completar las actividades del proyecto. [41]
- **Lecciones Aprendidas.** El conocimiento adquirido durante un proyecto el cual muestra cómo se abordaron o deberían abordarse en el futuro los eventos del proyecto, a fin de mejorar el desempeño futuro. [41]
- **Matriz de Trazabilidad de Requisitos.** Una cuadrícula que vincula los requisitos del producto desde su origen hasta los entregables que los satisfacen. [41]
- **Metodología.** Un sistema de prácticas, técnicas, procedimientos y normas utilizado por quienes trabajan en una disciplina. [41]
- **Objetivo.** Una meta hacia la cual se debe dirigir el trabajo, una posición estratégica que se quiere lograr, un fin que se desea alcanzar, un resultado a obtener, un producto a producir o un servicio a prestar. [41]
- **Presupuesto.** La estimación aprobada para el proyecto o cualquier componente de la estructura de desglose del trabajo o actividad del cronograma. [41]
- **Proceso.** Una serie sistemática de actividades dirigidas a producir un resultado final de forma tal que se actuará sobre una o más entradas para crear una o más salidas.
- **Proyecto.** Un esfuerzo temporal que se lleva a cabo para crear un producto, servicio o resultado único. [41]
- **Producto.** Un artículo producido, que es cuantificable y que puede ser un elemento terminado o un componente. Otras palabras para hacer referencia a los productos son materiales y bienes. [41]
- **Prototipos.** Un método para obtener una retroalimentación temprana respecto de los requisitos, proporcionando un modelo operativo del producto esperado antes de construirlo realmente. [41]
- **Requisito.** Una condición o capacidad que debe estar presente en un producto, servicio o resultado para satisfacer un contrato u otra especificación formalmente impuesta. También conocido como Requerimiento. [41]
- **Retraso.** La cantidad de tiempo en la que una actividad sucesora se deberá retrasar con respecto a una actividad predecesora. [41]

8.2 Bibliografía

- [1] Wikipedia, “Automatización robótica de procesos.” [Online]. Available: https://es.wikipedia.org/wiki/Automatización_robótica_de_procesos.

- [2] “Grupo ES3.” [Online]. Available: <https://es3.us.es/>. [Accessed: 04-Marzo-2024].
- [3] A. Jimenez-Ramirez, H. A. Reijers, I. Barba, and C. Del Valle, “A method to improve the early stages of the robotic process automation lifecycle,” in *International Conference on Advanced Information Systems Engineering*, 2019, pp. 446–461.
- [4] H. P. Fung, “Criteria, use cases and effects of information technology process automation (ITPA),” *Adv. Robot. Autom.*, vol. 3, 2014.
- [5] NASSCOM, “Cognitive RPA, The Future of Automation.” 2019.
- [6] C. Le Clair, G. O’Donnell, A. Lipson, and D. Lynch, “The Forrester Wave™: Robotic Process Automation, Q4 2019,” *Forrester Wave*, 2019.
- [7] “BluePrism.” [Online]. Available: <https://www.blueprism.com>.
- [8] “UiPath.” [Online]. Available: <https://www.uipath.com/>.
- [9] “Automation Anywhere.” [Online]. Available: <https://www.automationanywhere.com/>.
- [10] Samarpit, “RPA Developer Roles and Responsibilities.” [Online]. Available: <https://www.edureka.co/blog/rpa-developer-roles-and-responsibilities/>.
- [11] The Standish Group, “Chaos Report,” 2015.
- [12] Ministerio de empleo y seguridad social. Gobierno de España., “Convenio colectivo estatal de empresas de consultoría.” [Online]. Available: <https://www.boe.es/boe/dias/2023/07/26/pdfs/BOE-A-2023-17238.pdf>. [Accessed: 02-Marzo-2024].
- [13] M. de inclusión seguridad social y Migraciones, “Régimen General de la Seguridad Social.” [Online]. Available: <https://www.seg-social.es/wps/portal/wss/internet/Trabajadores/CotizacionRecaudacionTrabajadores/10721/10957/9932/4277>. [Accessed: 02-Marzo-2024].
- [14] Pc Componentes, “Portatil MSI Modern 15 I7 16gb 1TBssd 15.6.” [Online]. Available: <https://www.pccomponentes.com/portatil-msi-modern-15-b13m-281xes-intel-core-i7-1355u-16gb-1tb-ssd-156>. [Accessed: 03-Marzo-2024].
- [15] Agencia Tributaria, “Tabla de coeficientes de amortización lineal.” [Online]. Available: <https://sede.agenciatributaria.gob.es/Sede/ayuda/manuales-videos-folletos/manuales-practicos/irpf-2022/c07-rendimientos-actividades-economicas-estimacion-directa/fase-1-determinacion-rendimiento-neto/amortizaciones-dotaciones-ejercicio-fiscalmente-deducibles/requisitos-generales/coeficientes-amortizacion-lineal.html>. [Accessed: 04-Marzo-2024].
- [16] Pc Componentes, “Windows 11 Pro.” [Online]. Available: <https://www.pccomponentes.com/microsoft-windows-11-pro-64bit-espanol-licencia-permanente-fpp-descarga-digital>. [Accessed: 04-Marzo-2024].
- [17] Wikipedia, “Coste indirecto.” [Online]. Available: https://es.wikipedia.org/wiki/Coste_indirecto. [Accessed: 04-Marzo-2024].
- [18] Idealista, “Alquiler de oficina.” [Online]. Available:

- <https://www.idealista.com/inmueble/39723765/>. [Accessed: 06-Marzo-2024].
- [19] Selectra, “Precio diario de la luz (€/kWh): tarifas y comparativa.” [Online]. Available: <https://tarifasgasluz.com/comparador/precio-kwh>. [Accessed: 06-Marzo-2024].
- [20] OCU, “Precio del agua.” [Online]. Available: <https://www.ocu.org/alimentacion/agua/informe/el-precio-del-agua>. [Accessed: 06-Marzo-2024].
- [21] A. Rozinat & W. M. P. van der Aalst: "DecisionMining in ProM".
- [22] Felix Mannhardt, Massimiliano de Leoni, Hajo A. Reijers & Wil M. P. van der Aalst: "Decision Mining Revisited - Discovering Overlapping Rules".
- [23] Najah Mary El-Gharib, Daniel Amyot: "Robotic process automation using process mining — A systematic literature review".
- [24] Simone Agostinelli, Marco Lupia, Andrea Marrella, Massimo Macella: "Reactive synthesis of software robots in RPA from user interface logs", *Comput. Ind.* 142 (2022) 103721, <https://www.sciencedirect.com/science/article/pii/S016636152200118X>.
- [25] Volodymyr Leno, Adriano Augusto, Marlon Dumas, Marcello La Rosa, Fabrizio Maria Maggi, Artem Polyvyanyy: "Discovering data transfer routines from user interaction logs", *Inf. Syst* 107 (2022) 101916, <https://www.sciencedirect.com/science/article/pii/S0306437921001241>.
- [26] A. Martínez-Rojas, A. Jiménez-Ramírez, J.G. Enríquez, H.A. Reijers: "A screenshot-based task mining framework for disclosing the drivers behind variable human actions".
- [27] E. (Instituto N. de T. de la C. Inteco, “Ingeniería del Software: Metodologías y Ciclos de Vida,” *Inteco*, 2009.
- [28] K. Beck *et al.*, “Manifiesto for Agile Software Development,” *The Agile Alliance*, 2001. .
- [29] Wikipedia, “Manifiesto ágil.” [Online]. Available: https://es.wikipedia.org/wiki/Manifiesto_ágil#Principios_del_Manifiesto_Ágil. [Accessed: 05-Marzo-2024].
- [30] M. Trigás Gallego, “Metodología Scrum,” *Gest. Proy. informáticos*, 2012.
- [31] Wikipedia, “Scrum (rugby).” [Online]. Available: [https://en.wikipedia.org/wiki/Scrum_\(rugby\)](https://en.wikipedia.org/wiki/Scrum_(rugby)). [Accessed: 05-Marzo-2024].
- [32] J. O. Coplien, *A generative development process pattern language*. Cambridge University Press, New York, 1998.
- [33] F. P. J. Brooks and F. Mora, *Le mythe du mois-homme - Essai sur le génie logiciel*. 1996.
- [34] GV, “The Design Sprint.” [Online]. Available: <https://www.gv.com/sprint/>. [Accessed: 05-Marzo-2024].
- [35] Ntaskmanager, “Newbies’ Guide to Scrum Project Management 101.” [Online]. Available: <https://www.ntaskmanager.com/blog/newbies-guide-to-scrum-project-management-101/>. [Accessed: 05-Marzo-2024].

- [36] J. A. García-García, M. J. Escalona, F. J. Domínguez-Mayo, and A. Salido, “NDT-Suite: A Methodological Tool Solution in the Model-Driven Engineering Paradigm,” *J. Softw. Eng. Appl.*, 2014.
- [37] J. A. García-García, M. Alba Ortega, L. García-Borgoñon, and M. J. Escalona, “NDT-Suite: A model-based suite for the application of NDT,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2012.
- [38] The jQuery Project, “jQuery: The Write Less, Do More, JavaScript Library,” *Physical Review E*. 2011.
- [39] D. S. McFarland, *JavaScript & jQuery*. 2014.
- [40] Bootstrap, “Build fast, responsive sites with Bootstrap.”
- [41] PMI Project Management Institute, “Guía de los fundamentos para la dirección de proyectos (Guía del PMBOK),” 2013.
- [42] Scikit-learn. [Online]. Available: <https://scikit-learn.org/stable/>. [Accessed: 11-Mayo-2024].
- [43] Python-docx. [Online]. Available: <https://python-docx.readthedocs.io/en/latest/>. [Accessed: 11-Mayo-2024].
- [44] GraphViz. [Online]. Available: <https://graphviz.org/docs/library/>. [Accessed: 11-Mayo-2024].
- [45] PyDotPlus. [Online]. Available: <https://pydotplus.readthedocs.io/>. [Accessed: 11-Mayo-2024].

ANEXO

A. Manual de instalación

Antes de comenzar la instalación se recomienda el uso de un sistema con un mínimo de 16GB de memoria RAM. A pesar de que la plataforma se puede ejecutar directamente en un sistema Windows, se recomienda el uso de Docker para facilitar las instalaciones de ciertas librerías necesarias y garantizar el correcto funcionamiento de las librerías de machine learning.

Por lo tanto, antes de la instalación de la plataforma, es necesario instalar Docker en su sistema. Si su computador usa un sistema operativo Windows, es necesario habilitar la funcionalidad WSL (Windows Subsystem for Linux). La instalación de Docker es compleja y variará en función del sistema operativo base donde se implemente (<https://docs.docker.com/engine/install/>). En el caso de Windows y Mac es necesario instalar Docker Desktop, cuyas instrucciones de instalación se encuentran en el enlace anterior.

a. Despliegue

Una vez realizados los pasos anteriores, procedemos a descomprimir el zip “screenrpa.zip” donde se encuentra el código del proyecto.

A continuación, deberemos copiar el archivo “.env.sample” ubicado en la carpeta Docker (/screenrpa/docker/.env.sample), pegarlo en (/screenrpa/), renombrarlo a “.env” (/screenrpa/.env), y completarlo con los datos correspondientes, sustituyendo los valores indicados entre < > por los valores deseados.

```
POSTGRES_USER=<user>
DB_USER=<user>
POSTGRES_PASSWORD=<password>
DB_PASSWORD=<password>
POSTGRES_DB=<screenrpa>
DB_NAME=<screenrpa>
DB_HOST=db
DB_PORT=5432
DJANGO_SETTINGS_MODULE=core.settings
METADATA_PATH=/screenrpa/resources/metadata
API_VERSION=api/v1/
GUI_COMPONENTS_DETECTION_CROPPING_THRESHOLD=2
GAZE_MINIMUM_TIME_STARING=10
RESULTS_TIMES_FORMAT=seconds
DECISION_TREE_TRAINING_FOLDERNAME=decision-tree
GUI_QUANTITY_DIFFERENCE=1
PLOT_DECISION_TREES=FALSE
SCENARIO_NESTED_FOLDER=FALSE
SEVERAL_ITERATIONS=2
DEBUG=False
FIXATION_DURATION_THRESHOLD=10
```

```
FLATTENED_DATASET_NAME=flattened_dataset  
DECISION_TREE_TRAINING_ITERATIONS=10  
EMAIL_HOST_USER=<email for validation>  
EMAIL_HOST_PASSWORD=<password>  
DISABLE_MULTITHREADING=False
```

Tras esto ejecutaremos el comando en la consola:

```
docker-compose -f docker/docker-compose-prod.yml up -d
```

Y una vez finalizado este comando, podremos acceder a la plataforma en <http://localhost:8080>.

B. Manual de usuario

En esta sección se explicará cómo utilizar la plataforma ScreenRPA para configurar y ejecutar un caso de estudio desde la perspectiva de un nuevo usuario. Por lo tanto, en primer lugar será necesario registrarse e iniciar sesión.

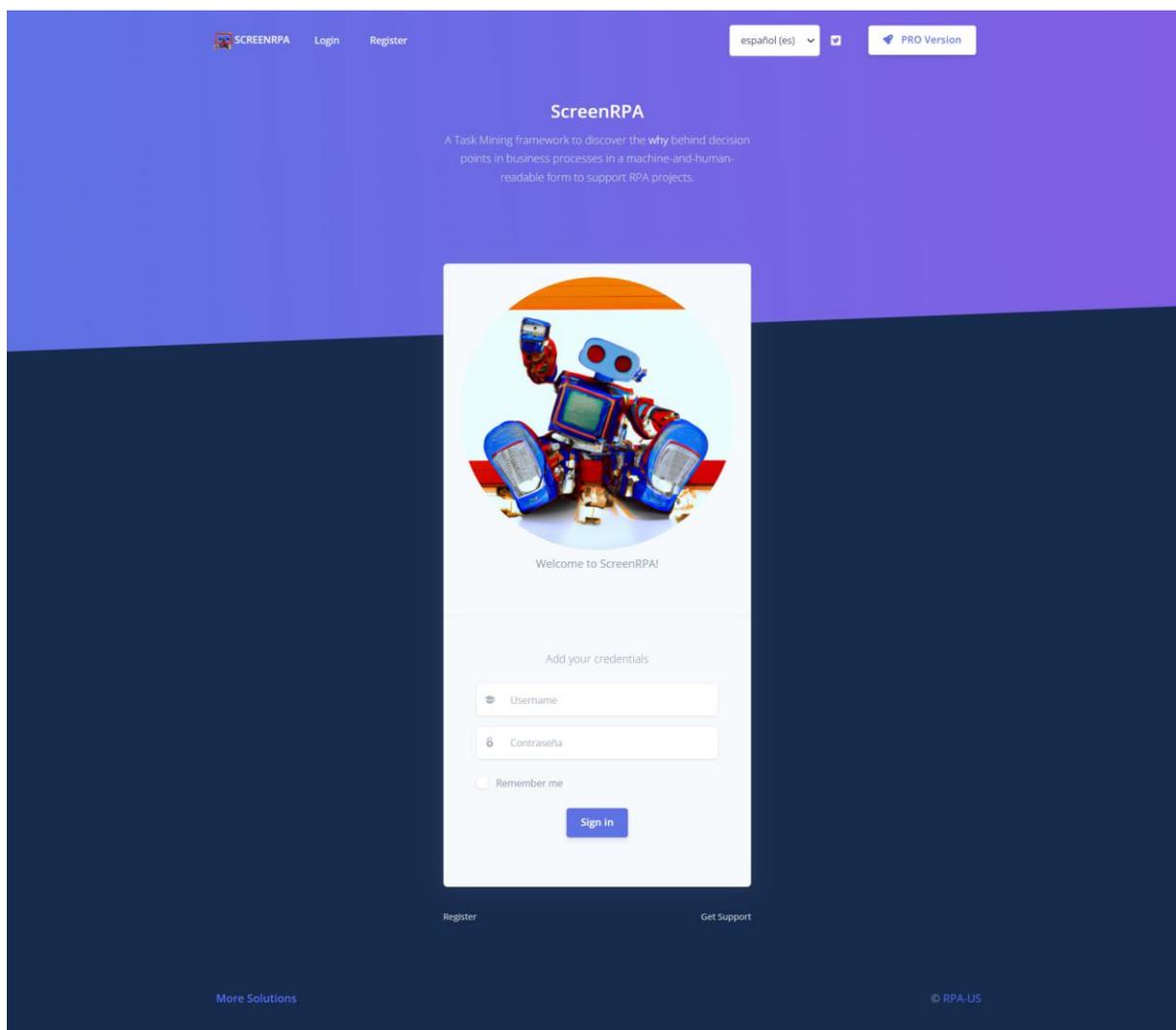
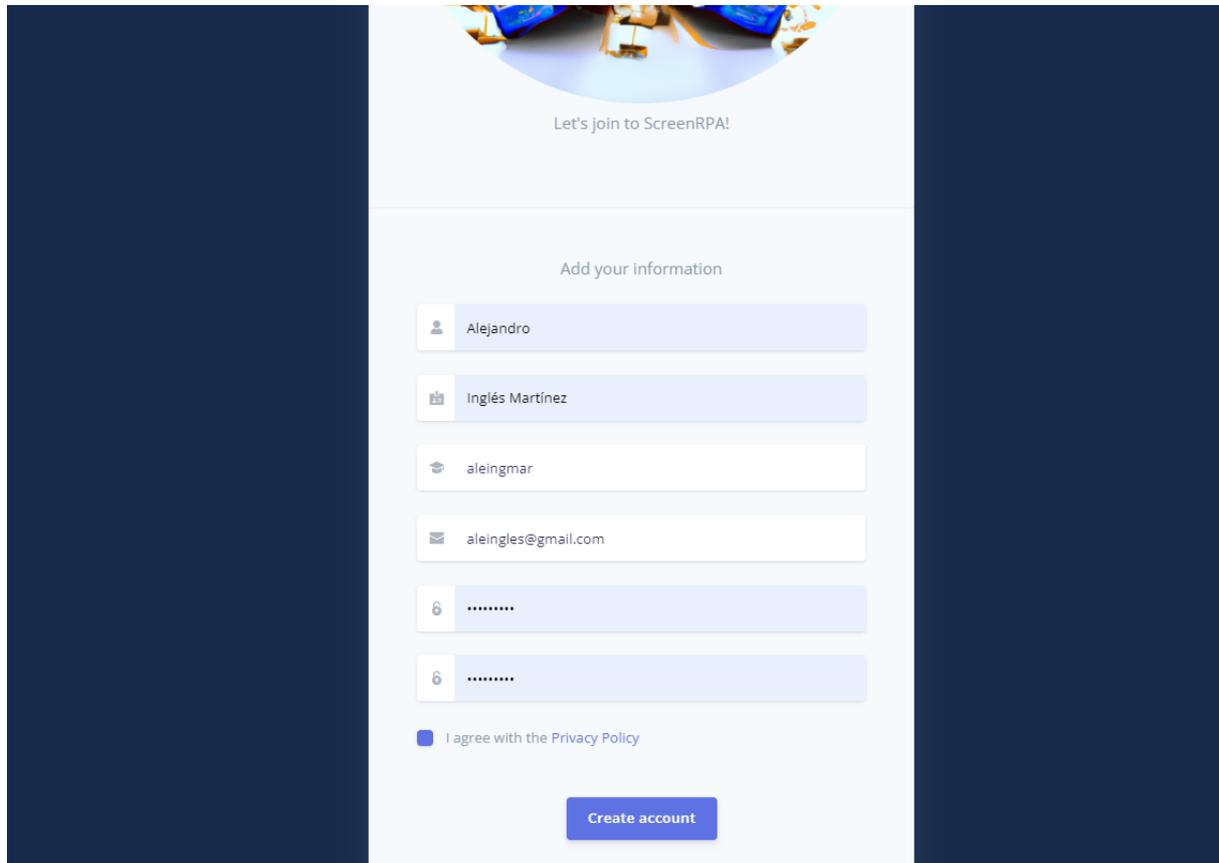


Figura 52. Página de inicio

Desde la página de inicio de la plataforma, haremos click en el botón “Registrarse”, lo que nos llevará a un formulario de registro de usuario en el que podremos rellenar nuestros datos para crear una nueva cuenta.



Let's join to ScreenRPA!

Add your information

Alejandro

Inglés Martínez

aleingmar

aleingles@gmail.com

6

6

I agree with the Privacy Policy

Create account

Figura 53. Formulario de registro de usuario relleno.

Una vez hecho esto podremos iniciar sesión y acceder a la plataforma.

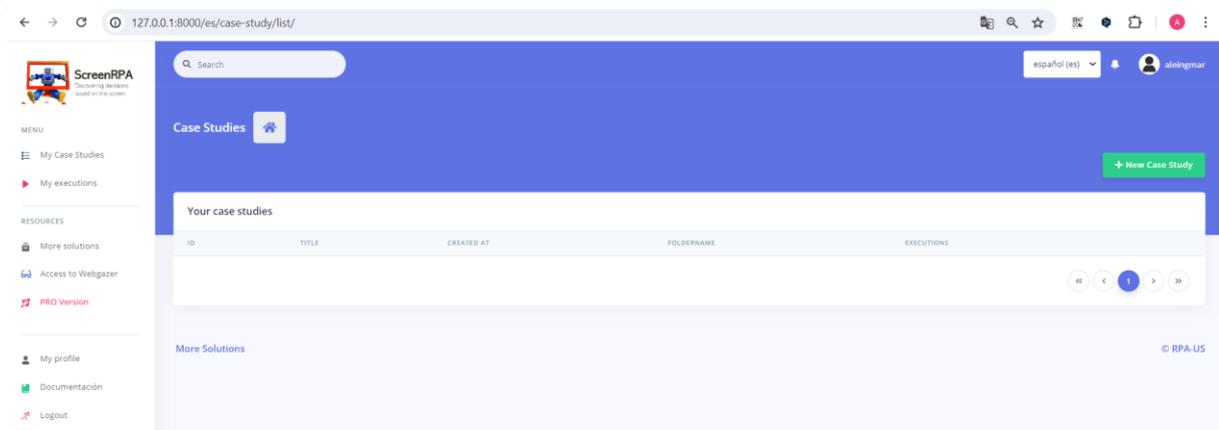


Figura 54. Listado de casos de estudio vacío

Al ingresar a la plataforma por primera vez, veremos una lista vacía de casos de estudio. Para continuar, será necesario crear uno nuevo.

Es importante señalar que la creación de casos de estudio debe realizarse externamente mediante una herramienta de monitoreo. Por lo tanto, para este manual, asumiremos que ya disponemos de un caso de estudio, el cual se adjunta con el nombre "SCR_TFG.zip". Los nombres de los escenarios a analizar en este caso de estudio son "sc_0_size100_Balanced" y "sc_0_size50_Balanced".

The screenshot shows the 'Create a new Case Study' form in the ScreenRPA interface. The form is divided into two main sections: 'GENERAL INFORMATION' and 'CONFIGURATION'.
GENERAL INFORMATION:
- **Title:** A text input field containing 'New'.
- **Description:** A larger text area containing 'This is a nice experiment...'.
- **Log file (zip):** A section with a note '(all UI logs has to be named as 'log.csv' to be loaded without log processing)'. Below it is a file selection button 'Seleccionar archivo' and a text input field containing 'SCR_TFG.zip'.
CONFIGURATION:
- **Map Log Columns to process Info*:** A grid of input fields for mapping log columns to case study fields.
- **Case*:** trace_id
- **Activity*:** activity_id
- **Screenshot*:** Screenshot
- **Variant*:** auto_variant
- **Timestamp*:** Timestamp
- **eyetracking_recording_time stamp*:** Recording timestamp
- **eyetracking_gaze_point_x*:** Gaze point X
- **eyetracking_gaze_point_y*:** Gaze point Y
- **NameApp*:** NameApp
- **EventType*:** MorKeyb
- **CoorX*:** Coor_X
- **CoorY*:** Coor_Y
- **Header*:** header
- **Scenarios to study*:** A text input field containing 'sc_0_size100_Balanced,sc_0_size50_Balanced'.
A green 'Guardar' button is located at the bottom right of the form.

Figura 55. Formulario de creación de un caso de estudio

Una vez creado el Caso de Estudio podremos acceder a él desde el listado anterior, y encontraremos en la parte inferior las configuraciones disponibles que se pueden aplicar actualmente.



Figura 56. Acceso a las configuraciones de las fases de un caso de estudio.

De manera similar al caso de estudio, podremos acceder al listado de configuraciones de una fase clickando en ella y, desde ahí, crear una nueva.

Las fases a ejecutar para analizar y procesar el caso de estudio de ejemplo son: “UI Element Detection”, “Process Discovery”, “Extract Training Dataset” y, por último, “Decision Tree Training”. Para llevar a cabo correctamente la configuración, estas fases se deben configurar con las opciones mostradas en las figuras incluidas en este manual de usuario (figuras 57-62).

En el caso de la configuración de la fase “UI Element Detection” se precargan los resultados, ya que en esta fase, a grandes rasgos, se reconocen elementos en las imágenes mediante técnicas de IA y este proceso puede tardar varias horas. El comprimido con la precarga a adjuntar se llama “precarga-ui-screen2som-SCR_TFG.zip”.

Figura 57. Formulario de creación de configuración de la fase “UI Element Detection”

Una vez creada cada configuración, deberemos activarla. Esto aplica para todas las configuraciones a aplicar al ejecutar el caso de estudio.

ID	TITLE	CREATED AT	PRELOADED RESULTS	TECHNIQUE	CLASSIFICATION	EXECUTION STATUS	ACTIONS
64	New	17 de junio de 2024 a las 10:09	SI precarga_ui_element_resultados_BgdubEp.zip	rpa-us	SI	Not Executed	Edit, Eliminar, Activate

Figura 58. Activación de configuración de una fase

The screenshot displays the 'Create a new Feature Extraction Technique Configuration' form in the ScreenRPA application. The interface is in Spanish and shows a user named 'aleingmar'. The form is divided into several sections:

- GENERAL INFORMATION:** Contains four fields: 'Title' (value: 'New'), 'Identifier' (value: 'sta'), 'Type' (value: 'Single: Feature extraction just after UI Elem. Detection'), and 'Technique' (value: 'centroid_ui_element_class').
- PRELOAD EXECUTION RESULTS:** Includes a checkbox for 'Preload Execution Results .zip' and a file selection button labeled 'Seleccionar archivo' with the text 'Ningún archivo seleccionado' and a help icon.
- CONFIGURATIONS:** Includes a checkbox for 'Apply filtering (Relevant Component Selection)'. Below it, a text area contains the condition: 'compo["relevant"]=="True" or compo["relevant"]=="Nested'.

At the bottom right of the form is a green 'Generate' button. The footer of the page includes 'More Solutions' and '© RPA-US'.

Figura 59. Formulario de creación de configuración de la fase "Feature Extraction"



- MENU
- My Case Studies
- My executions
- RESOURCES
- More solutions
- Access to Webgazer
- PRO Version
- My profile
- Documentación
- Logout

Search: No search available | Language: español (es) | User: aleingmar

← Go Back to Process Discovery Configurations | **New Process Discovery Configuration** | Home / Case Study 63 / Process Discovery Configs / New Process Discovery | Restart fields

Create a new Process Discovery Configuration

GENERAL CONFIG

Title *
New

Type*
rpa-us

MODEL CONFIG

Model Type
VGG

Labeling
Automatic

Remove Loops

CLUSTERING CONFIG

Clustering Type
Agglomerative

N Components
0,95

Use PCA

Show Dendrogram

PRELOAD EXECUTION RESULTS

Preload Execution Results .zip

Seleccionar archivo | Ningún archivo seleccionado ?

CONFIGURATIONS

Configurations

```
{ "model_type": "vgg", "clustering_type": "agglomerative", "labeling": "automatic", "use_pca": false, "n_components": 0.95, "show_dendrogram": false, "remove_loops": true }
```

Guardar

More Solutions | © RPA-US

Figura 60. Formulario de creación de configuración de la fase “Process Discovery”.

The screenshot displays the 'ScreenRPA' web application interface. On the left, there is a navigation menu with options like 'My Case Studies', 'My executions', 'More solutions', 'Access to Webgazer', 'PRO Version', 'My profile', 'Documentación', and 'Logout'. The main content area is titled 'New Extract Training Dataset Configuration' and contains a form with the following sections:

- GENERAL INFORMATION:** A text input field for 'Title' with the value 'New'.
- PRELOAD EXECUTION RESULTS:** A checkbox for 'Preload Execution Results .zip' which is unchecked. Below it is a file selection button 'Seleccionar archivo' and the text 'Ningún archivo seleccionado'.
- CONFIGURATION:** A text input field for 'User Actions Columns' with the value 'Coor_X,Coor_Y,Click'. Below it is a text area for 'Additional Configuration (JSON)' containing the value '[]'.

At the bottom right of the form is a green 'Generate' button. The footer of the page includes 'More Solutions' and '© RPA-US'.

Figura 61. Formulario de creación de configuración de la fase “Extract Training Dataset”.

En la configuración de la fase “Decision Tree Training”, es crucial especificar la técnica de minería de decisiones a utilizar en el apartado de “library”. Para este caso, se debe escribir la técnica de “overlapping”, que aplicará el método propuesto en este trabajo para mejorar la minería de decisiones y descubrir reglas no deterministas.

Create a new Decision Tree Training Configuration

Title *
Overlapping

Library:
overlapping

One hot columns:
NameApp, User

PRELOAD EXECUTION RESULTS
 Preload Execution Results .zip

Seleccionar archivo Ningún archivo seleccionado ?

Configuration:
[{"feature_values": {"1": {"or_cond": {"numeric_qua2_TextView_2_B": 11.0, "numeric_qua2_ImageView_2_B": 3.0, "numeric_qua2_ImageButton_2_B": 63.0, "status_categorial_sta_enabled_1044.0-775.0_2_B": 1.0, "sta_enabled_717.5-606.5_2_B": 1.0}, "or_cond_2": {"sta_checked_649.0-1110.5_4_D": 1.0, "status_categorial_sta_checked_649.0-1110.5_4_D": 1.0, "numeric_qua2_Checkbox_checked_4_D": 1.0, "numeric_qua2_Checkbox_unchecked_4_D": 0.0}}, "2": {"or_cond": {"numeric_qua2_TextView_2_B": 11.0, "numeric_qua2_ImageView_2_B": 3.0, "numeric_qua2_ImageButton_2_B": 63.0, "sta_enabled_717.5-606.5_2_B": 1.0, "status_categorial_sta_enabled_717.5-606.5_2_B": 1.0}, "or_cond_2": {"sta_checked_649.0-1110.5_4_D": 0.0, "status_categorial_sta_checked_649.0-1110.5_4_D": 0.0, "numeric_qua2_Checkbox_checked_4_D": 0.0, "numeric_qua2_Checkbox_unchecked_4_D": 1.0}}, "3": {"or_cond": {"numeric_qua2_TextView_2_B": 0.0, "numeric_qua2_ImageView_2_B": 0.0, "numeric_qua2_ImageButton_2_B": 5.0, "status_categorial_sta_enabled_717.5-606.5_2_B": 0.0, "sta_enabled_717.5-606.5_2_B": 0.0}, "or_cond_2": {"sta_checked_649.0-1110.5_4_D": 1.0, "status_categorial_sta_checked_649.0-1110.5_4_D": 1.0, "numeric_qua2_Checkbox_checked_4_D": 1.0, "numeric_qua2_Checkbox_unchecked_4_D": 0.0}}, "4": {"or_cond": {"numeric_qua2_TextView_2_B": 0.0, "numeric_qua2_ImageView_2_B": 0.0, "numeric_qua2_ImageButton_2_B": 5.0, "sta_enabled_717.5-606.5_2_B": 0.0, "status_categorial_sta_enabled_717.5-606.5_2_B": 0.0}, "or_cond_2": {"sta_checked_649.0-1110.5_4_D": 0.0, "status_categorial_sta_checked_649.0-1110.5_4_D": 0.0, "numeric_qua2_Checkbox_checked_4_D": 0.0, "numeric_qua2_Checkbox_unchecked_4_D": 1.0}}}], "centroid_threshold": 10000}

Columns to drop before decision point:
Timestamp_start, Timestamp_end

Generate

Figura 62. Formulario de creación de configuración de la fase “Decision Tree Training”.

Una vez configuradas todas las fases se procede a ejecutar con ellas el caso de estudio.

127.0.0.1:8000/es/case-study/list/

Search

Case Studies

Your case studies

ID	TITLE	CREATED AT	FOLDERNAME	Details	Eliminar	Download .zip	Execute
63	Prueba 100 casos	17 de junio de 2024 a las 10:05	CASO_ESTUDIO_100CASOS_1_escenario_WZ-HARY_1718627142				

Figura 63. Ejecución de un caso de estudio.

Si todo sale bien, se podrá acceder a la ejecución realizada desde el listado de ejecuciones, disponible en el menú de la derecha bajo "My executions". En esta vista, se muestra la configuración del caso de estudio ejecutado, las fases ejecutadas (a las que se puede acceder para ver los detalles de la configuración utilizada y los resultados generados), un listado de reportes creados a partir de la ejecución y el botón para acceder al formulario de creación de un reporte.

The screenshot displays the 'Execution 291 details' page in the ScreenRPA interface. The page is divided into several sections:

- GENERAL INFORMATION:** Includes fields for Title (New), Author (aleingmarr), Description (This is a nice experiment...), Folder path (media/unzipped/SCR_FTG_1719044723/executions/exec_291), and Decision Oracle (CheckBox_4_D and ImageView_2_B).
- CONFIGURATION:** Includes Special columns names (a JSON array of column names) and Scenarios to study (['sc_0_size100_Balanced', 'sc_0_size50_Balanced']).
- SET UP PHASES:** A grid of buttons representing different execution phases. 'Log Processing', 'Pre-Filtering', and 'Post-filtering' are in grey, while 'UI Elements Detection', 'Feature extraction', 'Process Discovery', 'Extract training dataset', and 'Decision tree training' are in blue.
- REPORTS:** A section indicating 'There are no reports of this execution.' with a 'Create' button.

Figura 64. Vista en detalle de una ejecución.

Las fases ejecutadas son accesibles al pulsar sobre ellas y se muestran en un tono azulado. Por otro lado, las fases que no han sido ejecutadas aparecen en un tono grisáceo y no se pueden manipular.

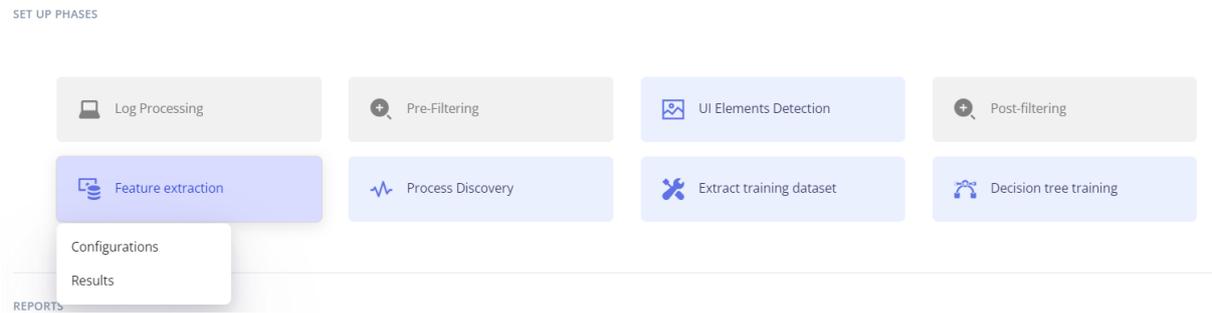


Figura 65. Fases realizadas durante la ejecución.

En todas las fases ejecutadas, es posible visualizar tanto la configuración utilizada como sus resultados. Para el manual de usuario, presentaremos una por una las vistas de resultados desarrolladas en el TFG. Sin embargo, solo mostraremos una vista de detalle de configuración, ya que se sobreentiende que el resto siguen un formato similar y no aporta mayor interés mostrarlas todas.

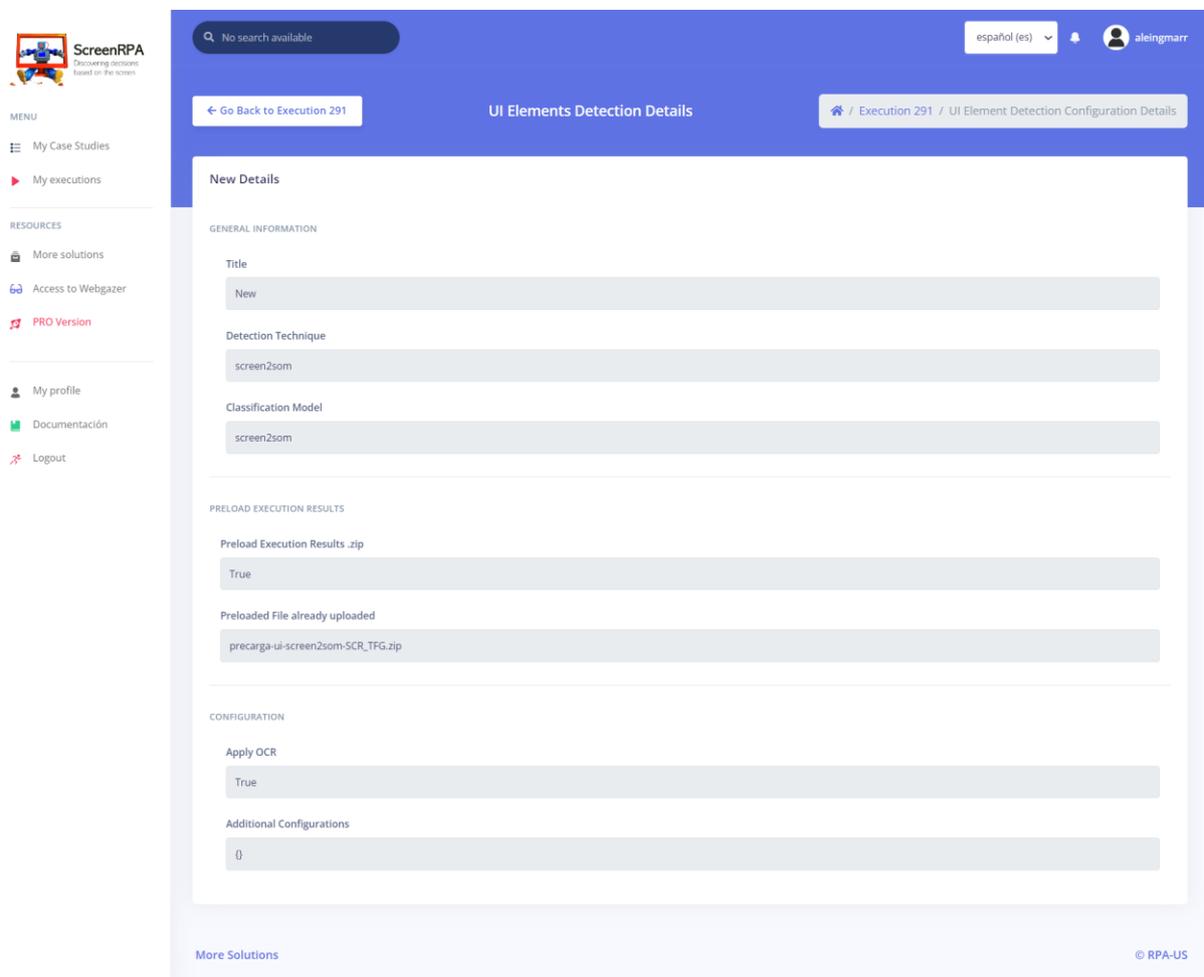
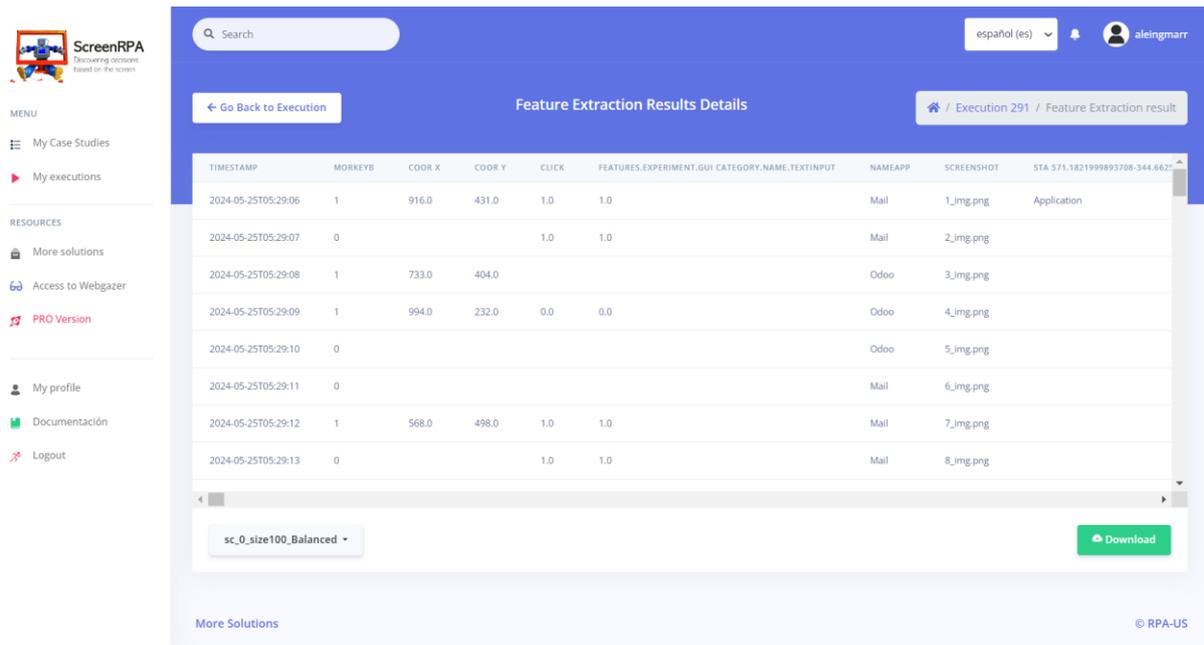


Figura 66. Vista en detalle de configuración de la fase “UI Element Detection”.

En la vista de resultado de la fase “Feature Extraction” se puede visualizar el dataset enriquecido con las características extraídas a partir del reconocimiento de elementos de las imágenes durante la fase

de "UI Element Detection". El desplegable de la parte inferior permiten visualizar y seleccionar los datasets enriquecidos de los distintos escenarios del caso de estudio, y el botón a la derecha permite descargar el dataset enriquecido visualizado en la interfaz en formato ".csv".

La interfaz permite desplazar el contenido del dataset tanto vertical como horizontalmente, posibilitando así la visualización completa del mismo. Como ejemplo de una característica añadida en un log enriquecido, se puede observar la última columna visible en la figura, donde el nombre de la columna indica el centro del elemento reconocido en las imágenes durante la fase de "UI Element Detection".



TIMESTAMP	MORKEYB	COORD X	COORD Y	CLICK	FEATURES.EXPERIMENT.GUI	CATEGORY.NAME.TEXTINPUT	NAMEAPP	SCREENSHOT	STA.571.1821999893708-344.662
2024-05-25T05:29:06	1	916.0	431.0	1.0	1.0		Mail	1_img.png	Application
2024-05-25T05:29:07	0			1.0	1.0		Mail	2_img.png	
2024-05-25T05:29:08	1	733.0	404.0				Odoo	3_img.png	
2024-05-25T05:29:09	1	994.0	232.0	0.0	0.0		Odoo	4_img.png	
2024-05-25T05:29:10	0						Odoo	5_img.png	
2024-05-25T05:29:11	0						Mail	6_img.png	
2024-05-25T05:29:12	1	568.0	498.0	1.0	1.0		Mail	7_img.png	
2024-05-25T05:29:13	0			1.0	1.0		Mail	8_img.png	

Figura 67. Vista de resultado de la fase "Feature Extraction".

En la vista en detalle de los resultados de la fase "Process Discovery" podemos observar el diagrama de proceso descubierto con todas las actividades que se llevan a cabo y las decisiones tomadas por el usuario que determinan el orden de realización de estas. En el desplegable situado abajo a la izquierda, podemos seleccionar cualquiera de los escenarios del caso de estudio ejecutado. En el desplegable del centro, podemos elegir las distintas variantes de proceso descubiertas en los datos. Al seleccionar una variante, el diagrama de proceso resalta las actividades que sigue el usuario en ese camino. El botón de la derecha permite descargar el diagrama de procesos en formato ".bpmn" y ".dot" para el escenario seleccionado.

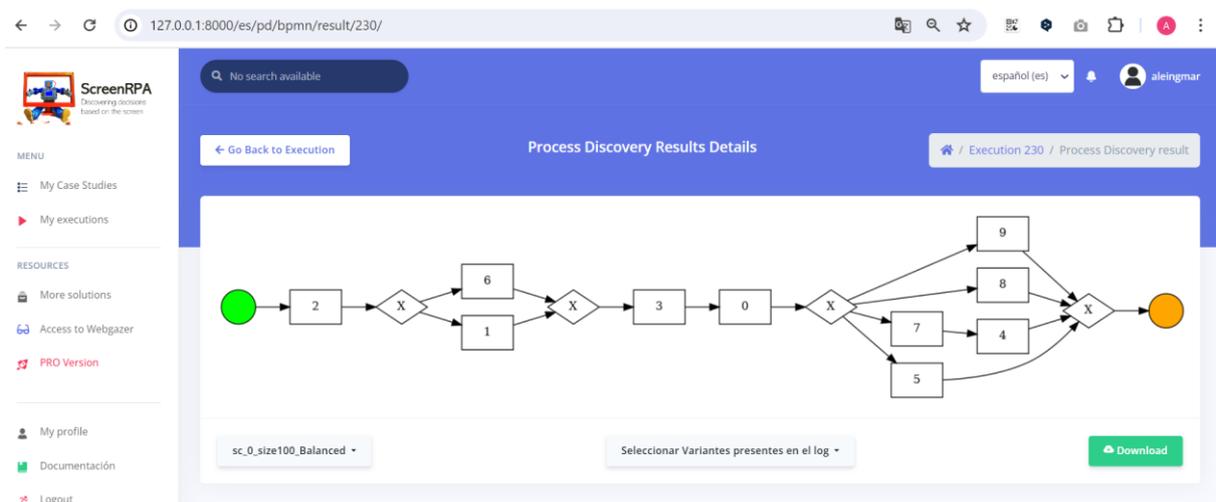


Figura 68. Vista del resultado de la fase “Process Discovery”.

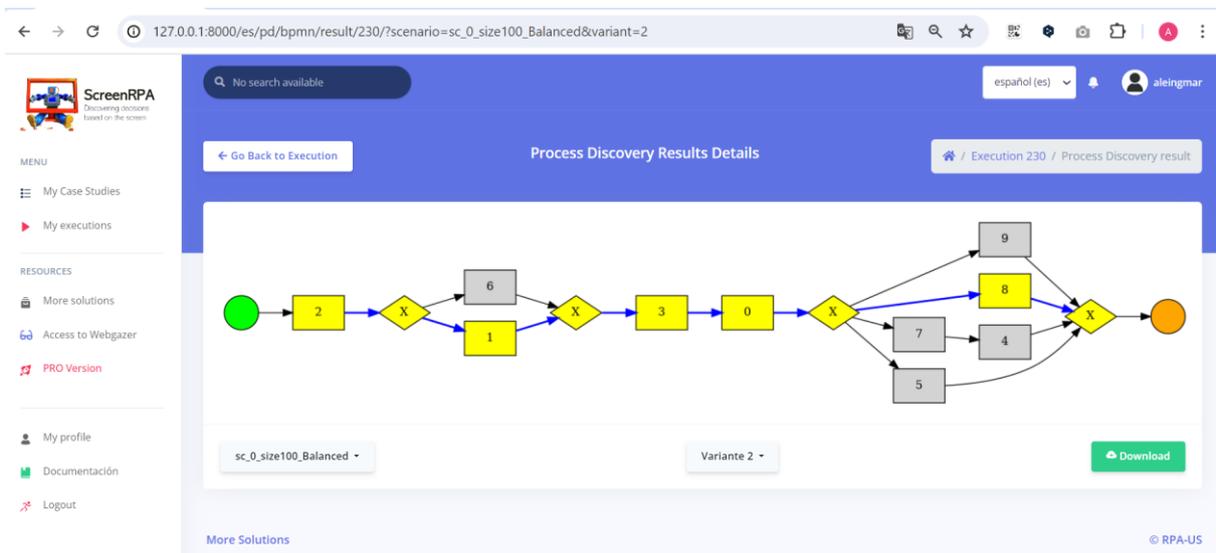


Figura 69. Vista de una variante específica del proceso descubierto.

En la vista de resultados de la fase "Extract Training Dataset", podemos ver el dataset aplanado para la aplicación de técnicas de decision mining. Habrá un "flattened dataset" para cada uno de los puntos de decisión del proceso. Con estos datasets, entrenaremos los árboles en la fase posterior (Decision Tree Training) para descubrir las reglas que determinan el curso de ejecución de las actividades del proceso.

Como se muestra en la imagen, podemos ver el dataset con sus respectivas filas y columnas, las cuales se pueden desplazar hacia abajo y hacia la derecha para ver la totalidad del dataset. Cada fila corresponde a un caso distinto de ejecución del proceso y a las características medidas hasta la decisión a predecir. El sufijo "_x" al final del nombre de las columnas indica la actividad en la que se recogió esa característica. En el desplegable de la izquierda, como anteriormente, se pueden seleccionar distintos escenarios del caso de estudio, y en el desplegable del medio se puede seleccionar visualizar el dataset aplanado correspondiente a cada una de las decisiones del proceso. El botón de la derecha permite descargar el dataset que se visualiza en formato ".csv", específico para el escenario y punto de decisión seleccionados.

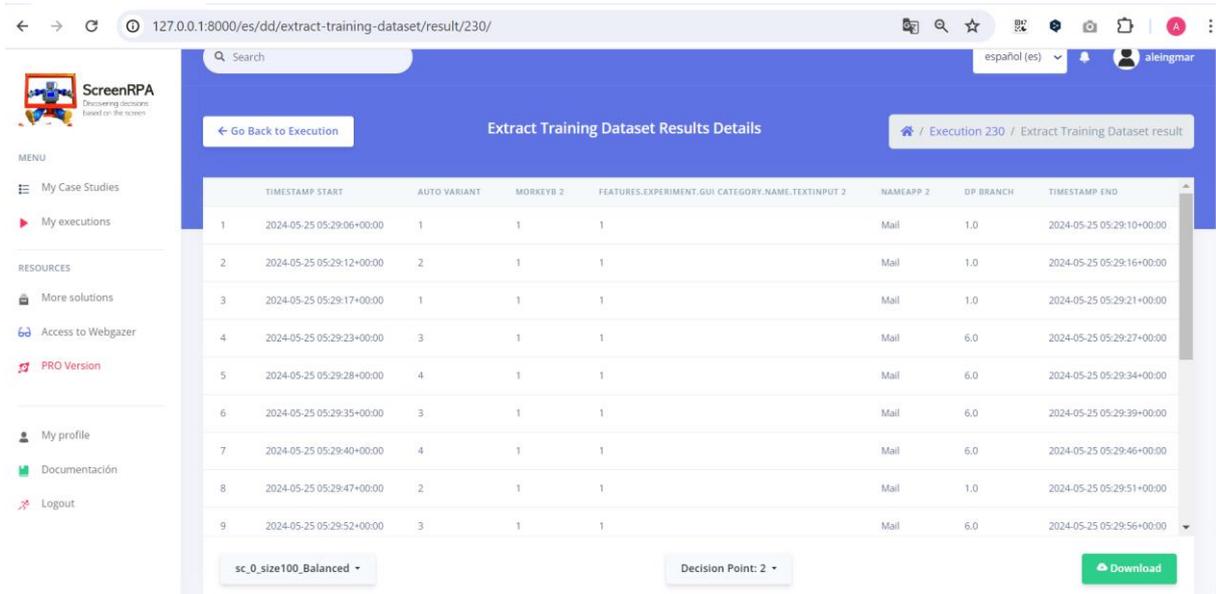


Figura 70. Vista del resultado de la fase "Extract Training Dataset".

En la vista de resultados de "Decision Tree Training", se puede observar para cada punto de decisión lo siguiente: a la izquierda, la representación visual del primer árbol de clasificación entrenado durante la primera capa del proceso de minería de decisiones "overlapping"; y a la derecha, las reglas deterministas y no deterministas encontradas para cada opción de decisión. Al hacer clic en cualquiera de las dos partes, se puede hacer zoom para ver los detalles más claramente. Los desplegables de selección de escenario, punto de decisión y descarga se encuentran en la parte superior derecha. Además, se puede descargar el árbol específico que se visualiza, correspondiente al escenario y punto de decisión seleccionados.

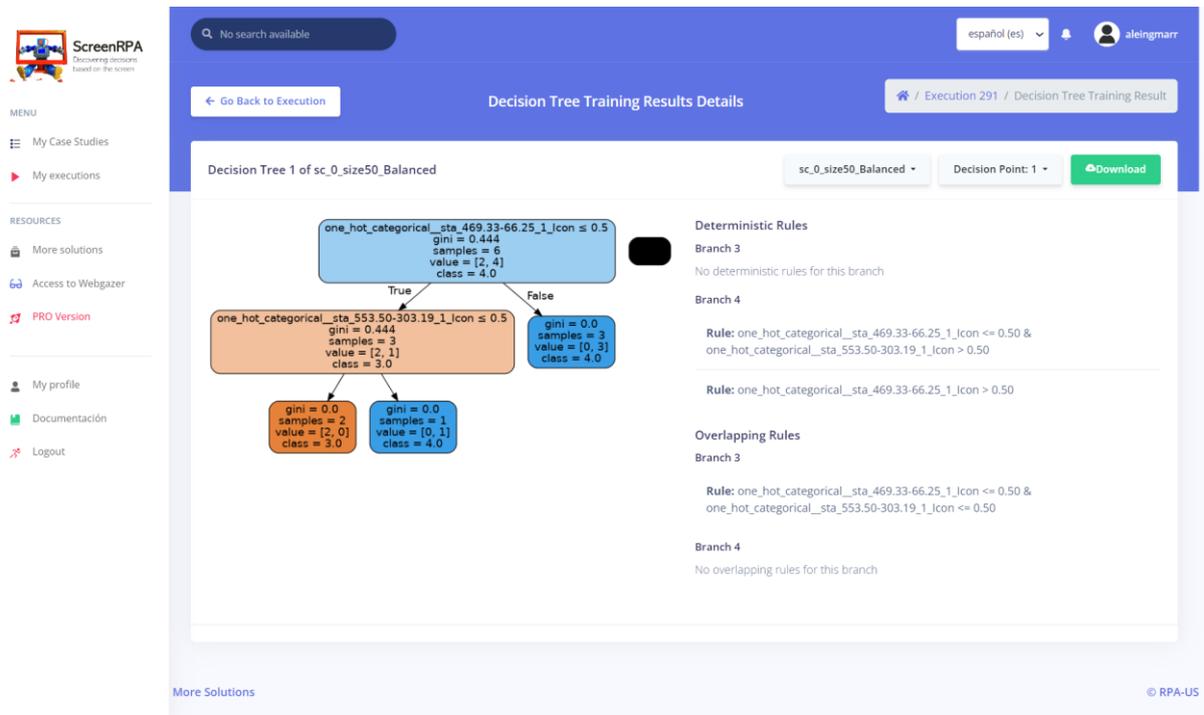


Figura 71. Vista de resultado de la fase “Decision Tree Training”.

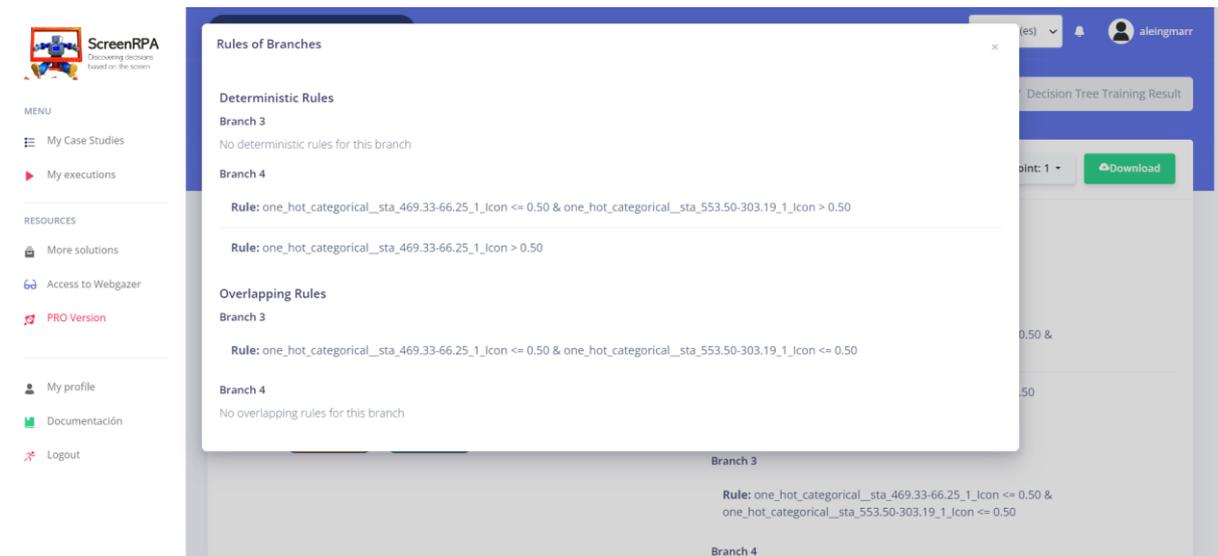


Figura 72. Zoom de las reglas descubiertas por el árbol de decisión.

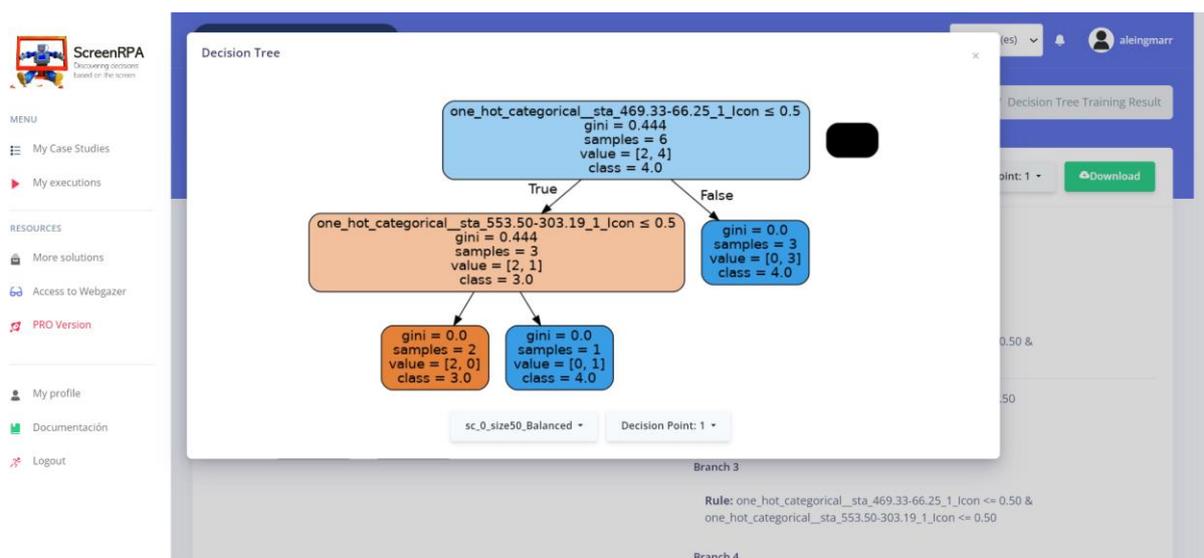


Figura 73. Zoom del árbol de decisión entrenado en la fase de “Decision Tree training”.

Pulsando el botón de creación que está en la parte inferior derecha de la vista en detalle de una ejecución, accedemos al formulario de creación del reporte donde indicamos el objetivo y propósito del documento (que se plasmaran después en este) y por último en la checklist de la parte inferior, indicamos la selección de contenido a añadir en el reporte. Como aclaratorio, en cada opción se resalta entre paréntesis la dependencia de esos contenido a la ejecución de determinadas fases. Aunque si una fase no fuera ejecutada directamente el contenido dependiente de esta no se mostraría en la checklist. Solo se muestra el contenido extraíble para el reporte.

Pulsando el botón de creación situado en la parte inferior derecha de la vista de detalle de una ejecución, accedemos al formulario de creación del reporte. En este formulario, indicamos el objetivo y propósito del documento (que se reflejarán posteriormente en el reporte). Además, en el listado de la parte inferior, seleccionamos el contenido a añadir en el informe, con una aclaración entre paréntesis sobre la dependencia de esos contenidos respecto a la ejecución de las respectivas fases de una ejecución. Si una fase no ha sido ejecutada, el contenido dependiente de esta no se mostrará siquiera en la checklist. Es decir, solo se muestra el contenido extraíble para el reporte. En nuestro caso, se muestra desbloqueado todo el contenido posible a plasmar en el informe.

The screenshot shows the 'Generate Report Configuration' interface. On the left is a sidebar menu with options like 'My Case Studies', 'My executions', 'More solutions', 'Access to Webgazer', 'PRO Version', 'My profile', 'Documentación', and 'Logout'. The main area has a search bar and a user profile 'aleingmarr'. The form itself is titled 'Generate a new Report' and contains the following sections:

- INTRODUCTION**
 - Purpose:** A text area with the placeholder 'Purpose of the PDD...'
 - Objective:** A text area with the placeholder 'Objective of the PDD'
- AS IS PROCESS DESCRIPTION**
 - Process Overview: (Locked: Process Discovery)
 - Applications Used: (No Locked)
 - AS IS Process Map: (Locked: Process Discovery)
 - Detailed AS IS Process Actions: (Locked: Decision Tree Training)
 - Input Data Description: (No Locked)

A green 'Generate' button is located at the bottom right of the form. At the bottom of the page, there are links for 'More Solutions' and '© RPA-US'.

Figura 74. Formulario de creación de reporte.

Una vez creado, aparecerá en el listado de reportes generados a partir de una ejecución en la vista en detalle de la misma. Al hacer clic en uno de ellos, podemos previsualizarlo en formato PDF, ver la configuración con la que se generó, descargar un archivo “.zip” con un reporte para cada escenario ejecutado en el caso de estudio y eliminar el reporte si es necesario.

The screenshot displays the 'Execution 291 details' page in the ScreenRPA application. The interface includes a sidebar with navigation options like 'My Case Studies', 'My executions', and 'Resources'. The main content area is divided into 'GENERAL INFORMATION' and 'CONFIGURATION' sections. A modal window titled 'Special columns names' is open, showing a list of columns and a 'Reporte 341' button. The configuration section contains several process steps: Pre-Filtering, UI Elements Detection, Post-filtering, Process Discovery, Extract training dataset, and Decision tree training. A 'Create' button is visible at the bottom right of the main content area.

Figura 75. Listado de reportes generado a partir de una ejecución.

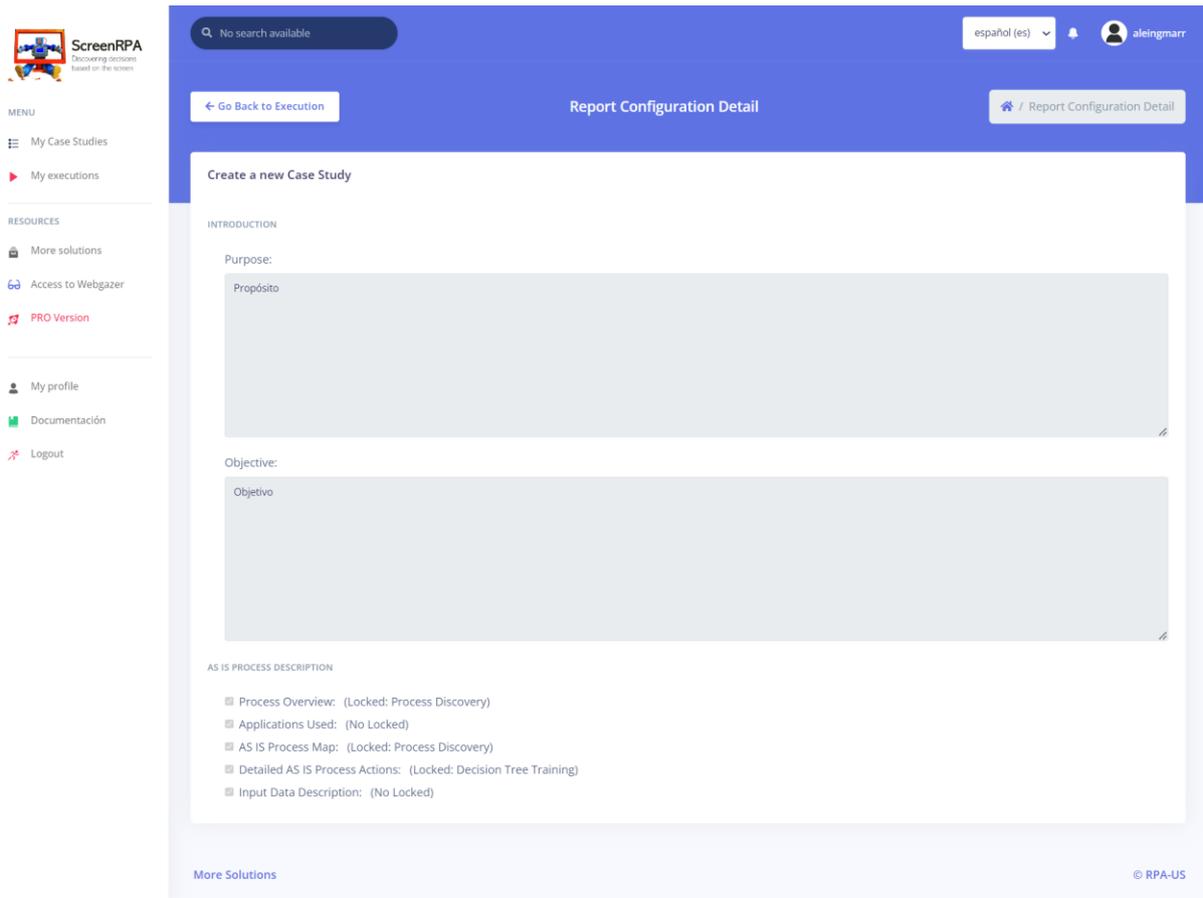


Figura 76. Detalle de configuración de un reporte.

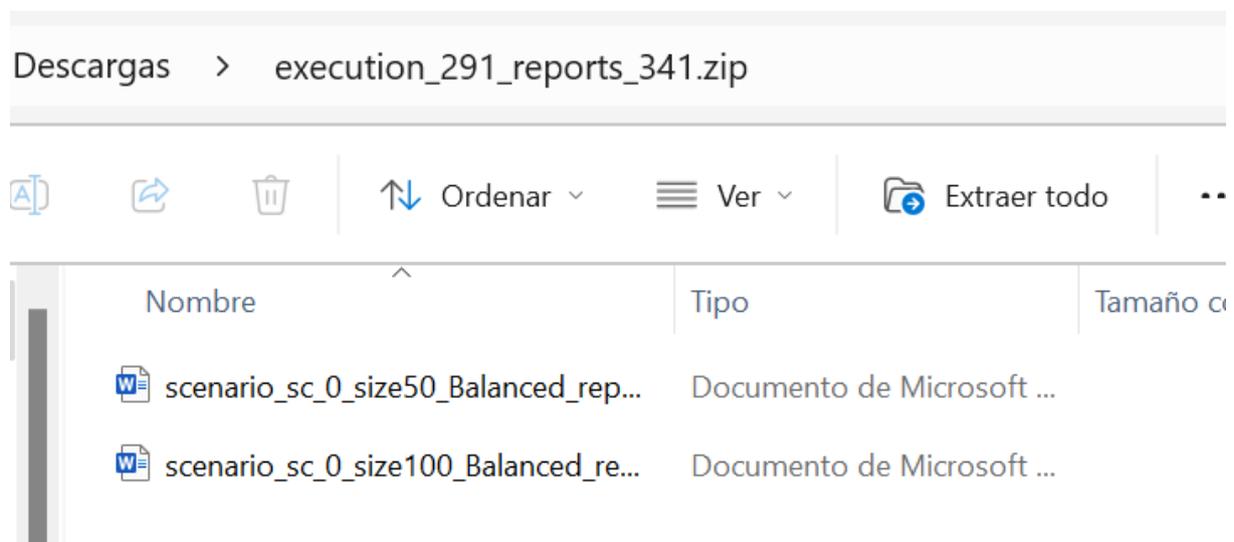


Figura 77. Comprimido con la descarga de un reporte.

El contenido del reporte se divide en dos secciones principales: “Introducción” y “Descripción del proceso As-Is”. En la introducción se detallan el propósito del estudio y los objetivos que se buscan alcanzar en el análisis del proceso.

I. INTRODUCTION

Purpose

The aim of the report is to provide human-understandable information on the current process to be improved.

Objectives

Discover the diagram that follows the process, the activities that take place in the process and the reasons for user decisions

Figura 78. Introducción del reporte.

En la sección “Descripción del Proceso As-Is”, como su nombre indica, se describe de forma ilustrativa y detallada el proceso actual en estudio. A grandes rasgos, esta sección incluye una representación gráfica del diagrama de proceso descubierto, la identificación de las posibles variantes dentro del mismo, la explicación de las actividades realizadas durante la ejecución de cada una de estas variantes y un análisis de las decisiones tomadas por los usuarios que influyen en el curso del proceso. Todo esto se desarrolla en varios apartados.

Para facilitar su interpretación, en esta sección del manual de usuario incluiremos imágenes de fragmentos clave del reporte. Estas imágenes nos ayudarán a referenciarlas y de manera ilustrativa describir el conocimiento que aportan.

En la sección "Applications Used" del reporte, se enumeran las diferentes aplicaciones empleadas por los usuarios durante el proceso, acompañadas de capturas de pantalla para su identificación.

Applications Used

The applications used by the user during the execution of the process are:

• Mail

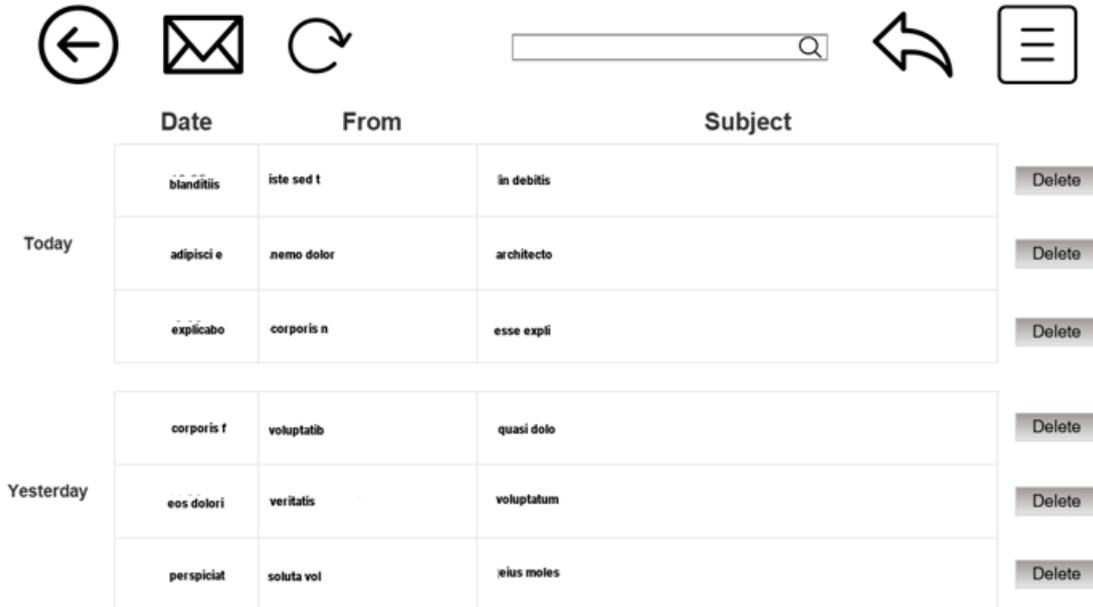


Figura 79. Aplicaciones utilizadas por el usuario durante el proceso.

En el apartado “As-Is Process Map” se presenta el diagrama del proceso descubierto tras el análisis realizado en la plataforma. Este diagrama muestra de manera gráfica las actividades que lleva a cabo el usuario y los puntos del proceso en los que el usuario toma decisiones que provocan la divergencia del proceso en distintos caminos y define las distintas variantes de proceso.

AS IS Process Map

This section contains various process maps contributing to a better understanding of how the process is performed pre-automation.

Below is the process model in BPMN format associated with the case study with a total of 10 activities and 2 decision points.

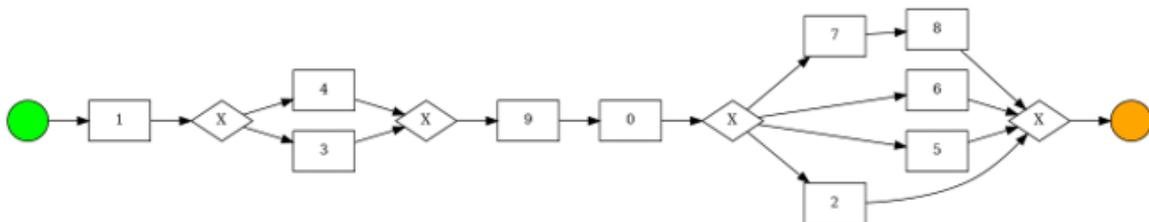


Figura 80. Descripción gráfica del proceso descubierto.

El apartado central de la Descripción del Proceso As-Is es "Detailed as is process actions". En esta sección, se describen minuciosamente y de manera comprensible para el lector las variantes que puede seguir el proceso. Se explican las actividades realizadas, los puntos de decisión encontrados y las reglas que rigen dichas decisiones por parte del usuario.

Al comienzo de la descripción de cada variante, se visualiza en el diagrama de proceso el flujo de actividades que sigue la variante a describir.

Detailed As Is Process Actions

Variant 1

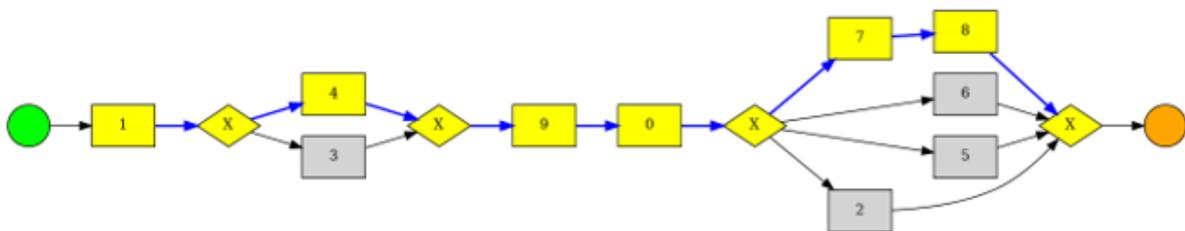


Figura 81. Flujo de actividades descrito en una variante.

Para describir las distintas variantes del proceso, se detallan de forma ordenada y secuencial las actividades que se llevan a cabo en cada una de ellas. La descripción de las actividades en el reporte se refleja de tres maneras:

En la figura 82, se muestra la descripción de una actividad de evento de clic, donde se representa en la captura de pantalla asociada a dicho evento las coordenadas promedio donde el usuario hace clic durante la ejecución de esta actividad.

En la figura 83, se describe una actividad de clic cuyas coordenadas están fuera de la resolución de la pantalla, probablemente debido a un error en la grabación del caso de estudio. Para mayor comprensión, se dibuja el borde de la imagen por donde el clic sobresale de la pantalla.

En la figura 84, se presenta la descripción de una actividad de escritura en teclado, adjuntando una captura de pantalla tomada durante el proceso de escritura e imprimiendo el texto escrito por el usuario. En el caso particular de nuestros datos y de la figura 84, se muestra una situación en la que, debido a un error en la obtención de los datos de los casos de estudio, no se logra capturar los textos que el usuario escribe.

Activity 1

Action 1

The user clicks at point (1029.0, 401.0)

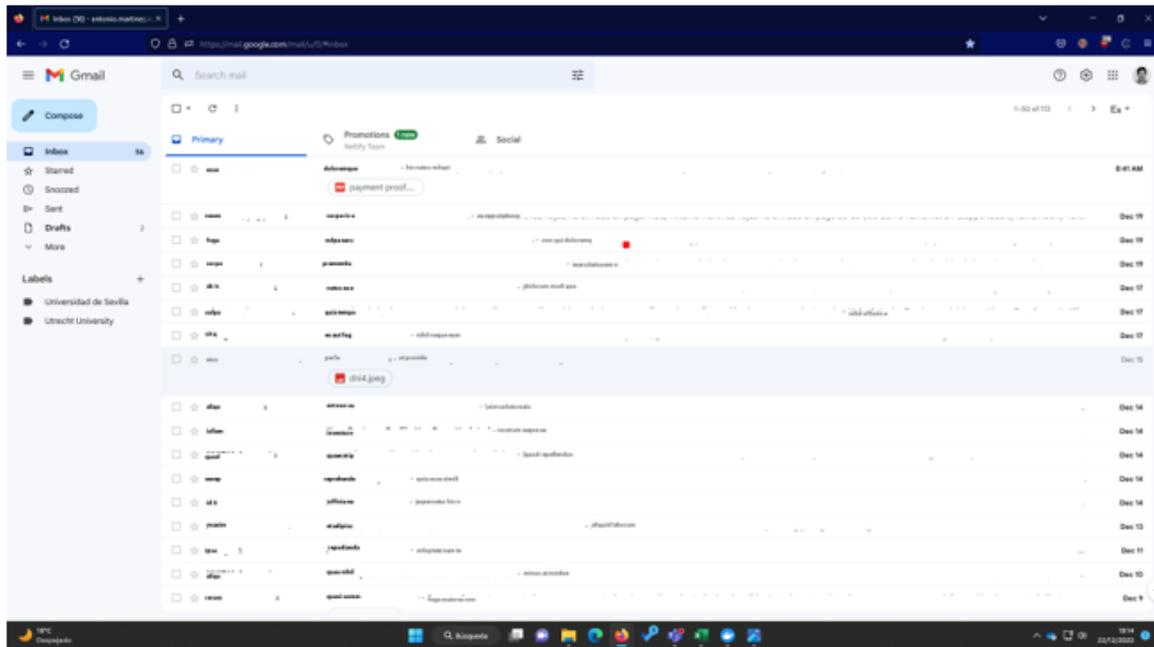


Figura 82. Descripción de actividad de click.

Activity 2

Action 1

ERROR: coordinates recorded incorrectly, out of screen resolution by right border (highlighted in red in the picture). User clicks on 1341.25, 410.0 and the screen resolution is 1130, 707.

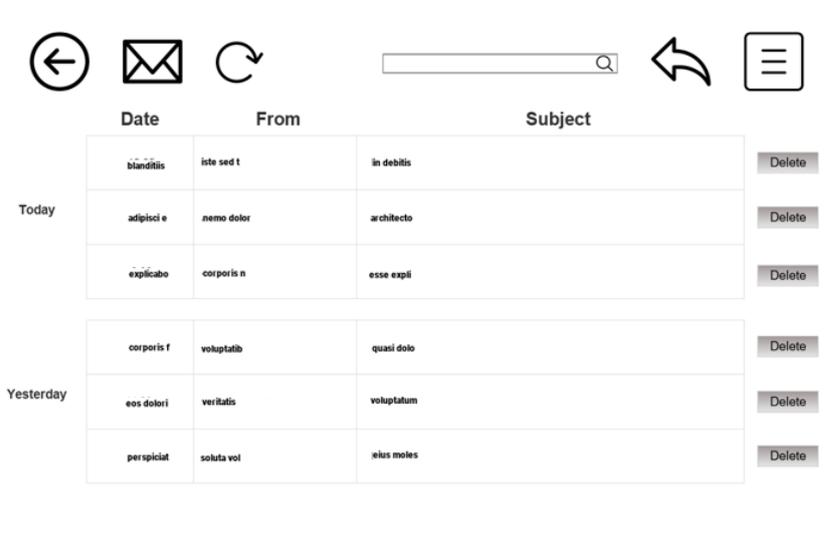


Figura 83. Descripción de actividad de click fuera de resolución de pantalla.

Activity 9

Action 1

The user writes "No TextInput"

CRM platform: Generate invoice

Customer: John Doe (johndoe@mail.com)	Invoice Date: 12/09/2020	
Description: no stability in the payment of bills	Due Date: 12/09/2020	
Product	Quantity	Price
[9853513342] Online infoproduct class 3	1	\$ 45.00
	Tax:	\$ 0.00
	Total:	\$ 45.00

[Save and download](#)

Figura 84. Descripción de actividad de escritura por teclado donde no se posee el texto escrito.

Además de la descripción de las actividades, se detallan en cada punto de decisión las reglas que determinan el flujo de la variante descrita. Estas reglas se identifican como deterministas o no deterministas.

En el caso de las reglas deterministas (figura 86), cada regla que provoca que la variante siga su flujo de proceso se desglosa con sus características, imprimiéndose en el documento la captura de pantalla asociada a la actividad en la que se recogen dichas características. Además, sobre dichas imágenes se dibuja la posición de los elementos referenciados en las reglas.

Así, de forma gráfica se consigue entender las reglas que dictan las decisiones tomadas por el usuario. Se puede apreciar que la presencia o ausencia de determinados elementos en las capturas de pantalla asociadas a las actividades provoca que los usuarios decidan realizar unas actividades u otras.

Para las reglas no deterministas (figura 87), para cada solapamiento entre ramas se describen las reglas superpuestas de forma gráfica. Las reglas solapadas se representan de manera similar a las reglas deterministas, con capturas de pantalla que ilustran las características relevantes. En el caso de la figura 87, por ejemplo, se describe el solapamiento de una regla entre la rama de la variante descrita y la rama 6.

Decision Point

After this activity there is a decision point where the user decides between taking 2 different branches.

In the case of this variant (variant 3) the user chooses to take branch 6.

In order for the user to decide this branch, one of these deterministic rules must be given:

- `one_hot_categorical__sta_571.18-344.66_2_Application > 0.50 & one_hot_categorical__sta_1048.18-276.89_2_NaN > 0.50 & one_hot_categorical__sta_1048.18-276.89_2_BtnSq <= 0.50 & one_hot_categorical__sta_1048.19-592.80_2_BtnSq > 0.50 & one_hot_categorical__sta_377.24-592.58_2_NaN <= 0.50`

In this decision we find rules that overlap with the rest of the decision branches. These rules are:

- Rule overlapping with branch 1: `one_hot_categorical__sta_571.18-344.66_2_Application <= 0.50.`
- Rule overlapping with branch 1: `one_hot_categorical__sta_571.18-344.66_2_Application > 0.50 & one_hot_categorical__sta_1048.18-276.89_2_NaN > 0.50 & one_hot_categorical__sta_1048.18-276.89_2_BtnSq <= 0.50 & one_hot_categorical__sta_1048.19-592.80_2_BtnSq <= 0.50.`

Figura 85. Descripción de reglas descubiertas en un punto de decisión.

Determinist Rules

one_hot_categorical_sta_469.33-66.25_1_Icon <= 0.50 & one_hot_categorical_sta_553.50-303.19_1_Icon > 0.50

Activity 1: one_hot_categorical_sta_469.33-66.25_1_Icon <= 0.50

Activity 1: one_hot_categorical_sta_553.50-303.19_1_Icon > 0.50

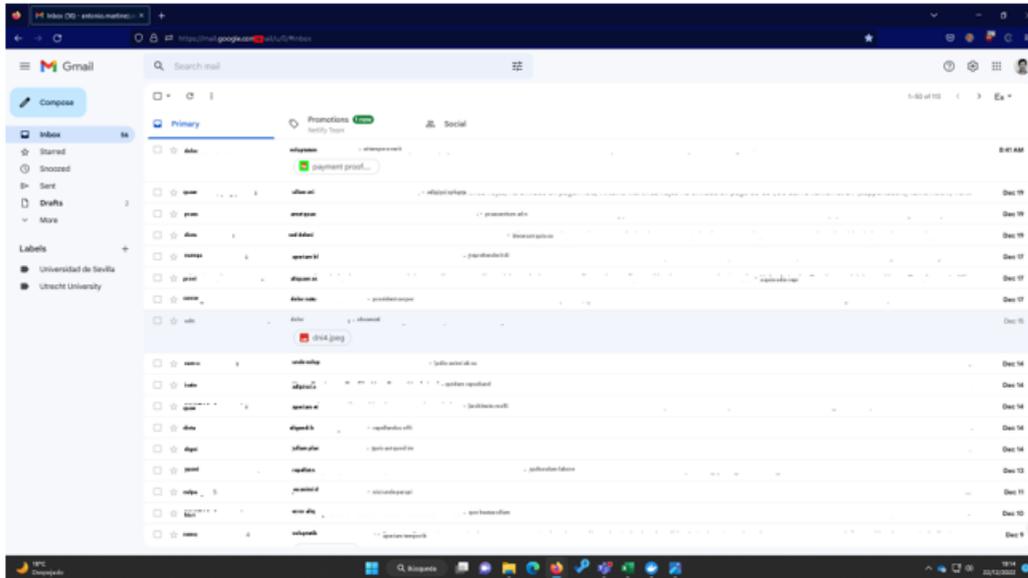


Figura 86. Descripción gráfica de reglas deterministas.

Overlapping Rules

Branch 6:

one_hot_categorical_sta_571.18-344.66_2_Application <= 0.50

Activity 2: one_hot_categorical_sta_571.18-344.66_2_Application <= 0.50

	Date	From	Subject	
Today	quidem rer	nulla quib	explicabo	Delete
	sit beatae	consequunt	nobis aspe	Delete
	optio labo	non nulla	eligendi d	Delete
Yesterday	sit neque	quo dignis	ipsam odio	Delete
	recusandae	eligendi f	laboriosam	Delete
	repudianda	iusto ipsa	nihil faci	Delete

Figura 87. Descripción gráfica de reglas solapadas entre ramas.

Finalmente, en el último apartado del reporte se presentan los datos de interacción de usuario de un escenario que recibe como entrada inicial la plataforma. Estos datos junto a las capturas de pantallas asociadas, permiten a la plataforma realizar todas las fases del estudio, extrayendo el conocimiento descrito en el reporte.

Input Data Description

The following table should contain details regarding the inputs that every action of the process takes.

Timestamp	MorKeyb	Coord_X	Coord_Y	Click	features.experiment.GUI_category.name.TextInput	Name App	Screenshot
2024-05-25T05:29:06	1	916.0	431.0	1.0	1.0	Mail	1_img.png
2024-05-25T05:29:07	0	nan	nan	1.0	1.0	Mail	2_img.png

Figura 88. Datos iniciales de entrada en la plataforma.