

Tipo de artículo: Artículo original

Aplicación de minería de procesos para la verificación de la conformidad del workflow de producción en la compañía EuroFish

Process mining application for the verification of the conformity of the production workflow in the EuroFish company

Carlos Andrés Villacreses Parrales ^{1*} , <https://orcid.org/0000-0002-4048-4316>

Lenin Jonatan Pin García ² , <https://orcid.org/0000-0001-8272-3816>

¹ Maestrando del Programa de Maestría de Tecnologías de la Información y la Comunicación. Ingeniero en Tecnologías de la Información. Instituto de Posgrado. Universidad Estatal del Sur de Manabí. Ecuador. villacreses-carlos7200@unesum.edu.ec

² Doctor en Tecnologías de la Información y Comunicaciones, Ingeniero en Sistemas con maestrías en Docencia universitaria, Sistemas de Información Gerencial y Gestión Estratégica de Tecnologías de la Información. Profesor principal a tiempo completo y director de investigación de la Universidad Estatal del Sur de Manabí. Ecuador. jonatan.pin@unesum.edu.ec

* Autor para correspondencia: villacreses-carlos7200@unesum.edu.ec

Resumen

El presente trabajo investigativo parte desde la recopilación y limpieza de información, determinación de objetivos, preparación del log de eventos, validación de los resultados y la obtención del flujo ideal. La importancia del estudio se basa en marcar un punto de partida para la aplicación de la minería de proceso en diferentes ámbitos, para corroborar si los procesos se están cumpliendo de la mejor manera. Teniendo, como objetivo aplicar la minería de procesos dentro del workflow de producción en una compañía atunera, y conocer el estado real de su flujo empresarial. Se utilizaron métodos científicos predominando histórico-lógico, análisis-síntesis, inducción-deducción, los cuales permitieron tener las pautas necesarias para abordar el tema y obtener resultados validados con relación a la temática. Los resultados parten desde la aplicación del ETL con Pentaho hasta obtener el modelo en Celonis y compararlo con el modelo base de la empresa, y conocer los cuellos de botellas para darles una posible solución. Con los datos generados por los sistemas de información se logró procesar 17 millones de transacciones; la cual debido a las limitaciones de licencia de Celonis se redujo el análisis al año 2022; obteniendo 8016 eventos, los cuales forman el log de eventos para descubrir el flujo de producción de la compañía seleccionada. La presente investigación se asocia al proyecto metodología de peligros y puntos críticos de control aplicando minería de procesos. Finalmente, en la verificación de la conformidad se obtiene una desviación del 94% entre el modelo real con el modelo obtenido de los datos de la compañía.

Palabras clave: algoritmos; cuellos de botellas; flujo de trabajo; minería de procesos; verificación de la conformidad

Abstract

This research study commences with the collection and cleansing of data, goal determination, event log preparation, result validation, and the identification of the ideal process flow. The significance of this study lies in establishing a starting point for applying process mining in various contexts, to ascertain whether the processes are being executed optimally. The primary objective is to apply process mining within the production workflow of a tuna processing company, and to assess the actual state of its business flow. Scientific methods employed include historical-logical analysis, synthesis, and induction-deduction, which provided the necessary guidelines for addressing the topic and obtaining validated results related to the subject matter. The outcomes range from the application of ETL using Pentaho to the acquisition of the model in Celonis, comparing it to the



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional**
(CC BY 4.0)

company's baseline model and identifying bottlenecks to propose potential solutions. Utilizing the data generated by information systems, 17 million transactions were processed; however, due to Celonis licensing limitations, the analysis was narrowed down to the year 2022, yielding 8,016 events that form the event log for discovering the production flow of the selected company. This research is associated with the hazard analysis and critical control points methodology project, applying process mining. In conclusion, the verification of conformity revealed a 94% deviation between the actual model and the model obtained from the company's data.

Keywords: *algorithms; bottlenecks; conformity verification; process mining; workflow*

Recibido: 25/04/2023

Aceptado: 08/06/2023

En línea: 01/07/2023

Introducción

El mundo y las organizaciones están llenos de procesos, desde las compras hasta el mapeo del recorrido del cliente, la logística y la contabilidad; las entidades se enfrentan a procesos complejos, globales y, a veces, defectuosos en las actividades diarias. Los procesos sin fricción, por otro lado, aseguran que se pueda encontrar las provisiones adecuadas en la tienda de comestibles, que los aviones aterricen a tiempo y que los tiempos de espera de los pacientes en los hospitales se mantengan al mínimo sin dejar de funcionar, de forma que mejoren y automaticen todos los procedimientos que se efectúan en cualquier tipo de organización de comestibles.

La minería de procesos realiza la combinación de dos disciplinas como lo son la *Data Science* y el *Business Process Management*, principalmente empleando técnicas de ciencias de datos como el *Big Data* y la inteligencia artificial, abordando los problemas de la ciencia de procesos para mejorar y automatizar los procedimientos de algún servicio o elaboración de un producto. De la misma manera es importante que las empresas en la actualidad implementen o usen la minería de procesos como un recurso más dentro del workflow de elaboración de un producto o servicio, para dar solución a los cuellos de botella existentes o reducir los tiempos de ejecución de alguna actividad que no se esté cumpliendo de la mejor manera y poder evitar pérdidas tanto de tiempo como económicas para la empresa.

El objeto de investigación se basa en la aplicación de minería de procesos para la verificación de la conformidad del workflow de producción en una compañía atunera, recolectando y analizando los datos proporcionados en cada una de las actividades que se realizan dentro del proceso de producción, previamente registrados en los sistemas informáticos de la empresa.

Previamente teniendo la data a analizar se procede a realizar un proceso de ETL (*extraction, transform and loading*), para obtener el log de evento base del modelo a alcanzar, este log de evento o event log, está formada por los siguientes campos: ID de caso, la actividad realizada y la duración de ejecución para dicha actividad; luego de ello se procede a usar un software, como Disco, ProM o Celonis entre los más conocidos, que permiten previsualizar el



modelo obtenido en relación a los datos extraído de la data principal; detallándose los cuellos de botellas, retrasos e inconvenientes que existen dentro del flujo de procesos de las actividades que se realiza en la empresa y así poder analizarlos para darle una pronta solución, sea para ejecutarlo en menos tiempo o diferenciar si es necesaria o no la actividad; cada uno de estos software son esenciales y facilitan el previo análisis de la información.

Materiales y métodos

En el marco de la metodología de la investigación, el presente trabajo se realizó a partir del análisis de los acontecimientos y referentes históricos que se lograron obtener a través de varias fuentes de información; de acuerdo con la profundidad el estudio este se basa en la investigación descriptiva por lo que se detalla todo lo referente al estado del arte y los resultados que se han obtenido con la aplicación de la minería de procesos en diferentes empresas en la actualidad. Se utilizaron métodos de la investigación científica tales como:

Métodos teóricos

Histórico – lógico. – Se utilizó para la sustentación de la investigación y en particular para conocer las tres instancias como descubrir, monitorear y mejorar procesos reales, que se tiene al aplicar minería de procesos, permitiendo tener una validez que respalde esta información y así tener una idea clara para dar cumplimiento a los objetivos planteados.

Análisis – síntesis. – Se empleó para formar todo el contenido del estado del arte, detallando algunos casos de éxitos e importancia que se tiene el aplicar minería de procesos dentro de las organizaciones.

Inducción – deducción. – Se utilizó para la indagación de la problemática que se presenta dentro de la empresa EuroFish al no contar con la tecnología de minería de procesos, pueda identificar patrones, desviaciones y en última instancia eliminar los cuellos de botella presentes en cada una de las actividades que se realizan dentro de la organización.

Métodos empíricos

Bibliográfico – documental. – Permitió realizar una indagación profunda en diferentes fuentes bibliográficas como: revistas, artículos, trabajos de titulación y libros, para la construcción del marco teórico y darle un sustento verídico a la investigación propuesta.



Minería de Procesos

El *Process Mining* es la nueva tecnología líder cuando se trata de hablar de "negocios algorítmicos", es decir, empresas que utilizan algoritmos y grandes cantidades de datos en tiempo real para crear valor comercial, esto solo ha sido posible a través de la llegada de sistemas de información y herramientas administrativas.

Mientras que las personas tienden a diseñar y pensar en forma de flujos de procesos simples (imagen de la izquierda), la realidad tiende a ser más compleja con múltiples variantes (imagen de la derecha). En este sentido, se suele diferenciar entre procesos “*To-Be*” y “*As-Is*”:

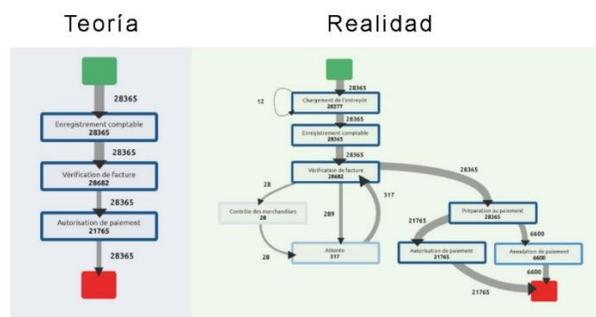


Figura 1. - “To Be” y “As Is” del process mining.

Adaptado. Procesos Objetivos 2023, <https://procesosobjetivos.com/process-mining/>

To-Be presenta el flujo de proceso ideal y perfecto sin fricción, tal como suele diseñarse en teoría. *As Is*, por otro lado, presenta el flujo del proceso real con todas las desviaciones y complejidades que ocurren en los procesos operativos de la vida real (Reinkemeyer, 2020).

La minería de procesos está estrechamente relacionada con el *Business Process Management (BPM)* y el *Workflow Management (WFM)*. El BPM es la disciplina que combina el conocimiento de la tecnología de la información y el conocimiento de las ciencias de la gestión y lo aplica a los procesos operativos del negocio; mientras que el WFM se enfoca más en la automatización de los procesos comerciales (W. M. P. van der Aalst, 2011).

Es importante entender el ciclo que tiene el BPM, el cual se lo identifica con las siguientes fases y descubrimiento del proceso, produciendo como resultado una arquitectura o mapa de procesos; las fases del BPM se compone de identificación de procesos, descubrimiento de procesos, análisis de procesos, diseño de procesos, configuración e implementación, ejecución y monitoreo, y evaluación. (Cocconi et al., 2019, p. 2)



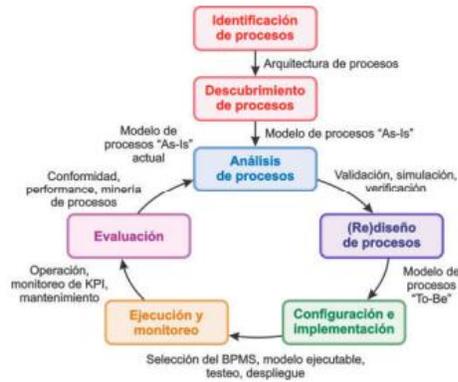


Figura 2. - Ciclo de vida del BPM.

Adaptado. Sedici [Fotografía], por (Cocconi et al., 2019), <http://sedici.unlp.edu.ar/handle/10915/77070>

Process Mining logra esta unión tomando las huellas digitales que se crean en los sistemas de TI y usándolas para reconstruir y visualizar los flujos de procesos. A partir de aquí, la tecnología de minería de procesos puede identificar patrones y desviaciones y, en última instancia, eliminar los cuellos de botella.

¿Qué son las huellas digitales?

Los registros de eventos son el formato en el que podemos recuperar nuestras huellas digitales de los sistemas de Tecnologías de la Información (TI) subyacentes. Son esencialmente los libros de registro que los sistemas de TI mantienen para registrar qué eventos tienen lugar para cada ID de caso y en qué momento.

Las técnicas de minería de procesos extraen conocimiento de los registros de eventos con el fin de descubrir, monitorear, y mejorar los procesos (Wil Van der et al., 2011, p. 3). Lo que venimos llamando “registro de evento” (*event log*), está formado por 4 piezas claves, las cuales son:

- un **id de caso**: un identificador único, como un artículo de pedido de compra, un número de factura o un número de pedido
- una **actividad**: la descripción de lo que ha sucedido, por ejemplo, la creación de una orden de compra o la recepción de bienes
- un **recurso**: la persona que ejecuta la actividad, por ejemplo, el usuario que registra la compra
- una **marca de tiempo**: la fecha y hora en que tuvo lugar la actividad (W. Van der Aalst, 2016, pp. 128–130)



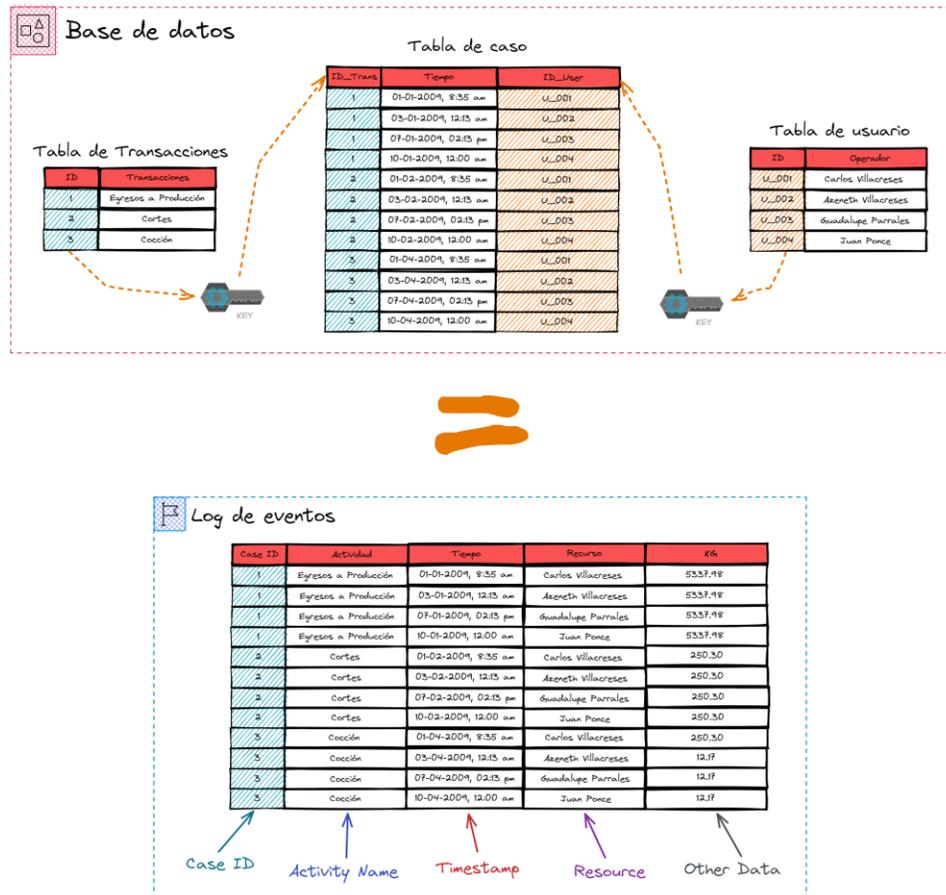


Figura 3. - Ejemplo de estructura del registro de eventos (event logs).

Adaptado. Celonis [Basado], por autor de la investigación, 2023

El punto de datos de ID de caso se incluirá tanto en la tabla de casos como en el registro de eventos y, por lo tanto, actúa como una clave común para ambas tablas (W. Van der Aalst, 2016).

Beneficios

Con la aplicación de la minería de procesos las entidades u organizaciones pueden obtener los siguientes beneficios:

- Descubrir el modelo de ejecución real del proceso
- Determinar si el proceso cumple con la reglamentación y procedimientos documentados
- Analizar la interacción del personal que ejecuta el proceso
- Descubrir cuellos de botella
- Monitorear la productividad del personal



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

- Predecir el tiempo de ciclo de un caso
- Determinar la relación entre las variables de un caso (Aguirre Mayorga & Rincón García, 2015)

Tipos de minería

El eje central de la minería de procesos es el mejoramiento de los procesos de cada actividad, a través de herramientas que permitan el análisis de la gran cantidad de datos que provienen desde varios sistemas informáticos. En la **Figura 4** se esquematiza los tres tipos principales (W. M. P. van der Aalst, 2011): descubrir (*Discovery*), cumplimiento (*Conformance*) y mejora (*Enhancement*) (Aguirre Mayorga & Rincón García, 2015).

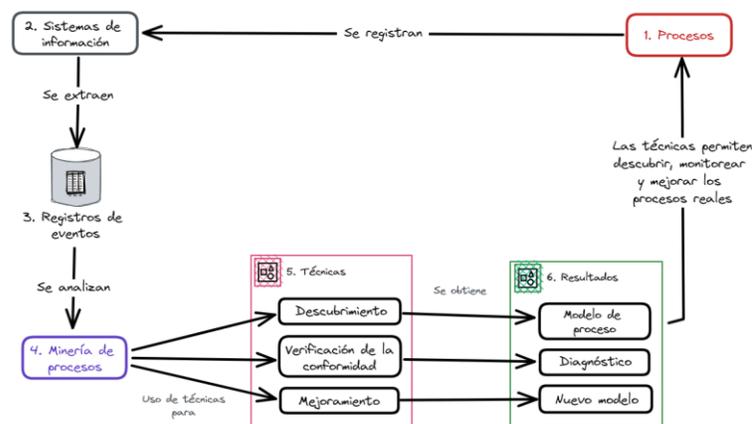


Figura 4. - Tipos de minería de procesos.

Adaptado. IIESCA [Fotografía], por (Céspedes González et al., 2018), <https://www.uv.mx/iiesca/files/2018/11/12CA201801.pdf>

- **Descubrir (*Discovery*).** – es la técnica de descubrimiento toma un registro de eventos y produce un modelo sin usar ninguna información a priori (Cocconi et al., 2019).
- **Cumplimiento (*Conformance*).** – en este modelo existe un proceso y se compara con el *log* de eventos previamente obtenido; además de ello se verifica la conformidad del workflow para saber si los registros son registrados de manera correcta y conocer si se ajusta al modelo real o viceversa (Contreras et al., 2016).
- **Mejora (*Enhancement*).** - aquí, la idea es ampliar o mejorar un modelo de proceso existente utilizando información sobre el proceso real registrado en algún registro de eventos. Mientras que la verificación de conformidad mide la alineación entre el modelo y la realidad, este tercer tipo de minería de procesos tiene como objetivo cambiar o ampliar el modelo a priori; es decir, modificar el modelo para reflejar mejor la realidad. (W. Van der Aalst, 2016)



En la Figura 4 se logra apreciar que los resultados de la aplicación de la minera depende hasta donde se quiera llegar (González, 2016): si se aplica el descubrimiento, esta comienza desde el log de eventos y el resultado final será la obtención del modelo de acuerdo a los datos; en cambio la conformidad parte desde el log de eventos junto al modelo obtenido, para lograr obtener un diagnóstico de la situación teórica (*To Be*) con la realidad (*As Is*); mientras que el mejoramiento parte desde el log de eventos, previamente pasado por el descubrimiento y la conformidad para luego tener como resultado mejorar el modelo y obtener uno nuevo, el cual coadyuve a mejorar y este se adapte de la mejor manera a cada actividad de los procesos de la empresa.

Tipos de representación de modelos

PETRI NETS

Las redes de Petri son el lenguaje de modelado de procesos más antiguo y mejor investigado que permite el modelado de concurrencia. Aunque la notación gráfica es intuitiva y simple, las redes de Petri son ejecutables y se pueden usar muchas técnicas de análisis para analizarlas (W. Van der Aalst, 2016, p. 59). Son un sistema de modelado para procesos concurrentes formado por tres elementos: transiciones, lugares y arcos (González, 2016, p. 18).

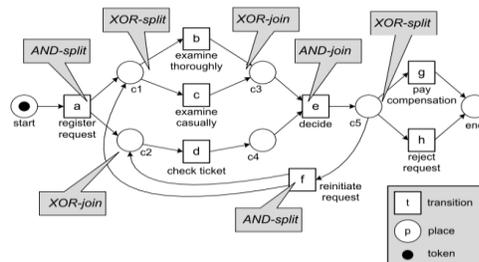


Figura 5. - Red de Petri marcada

Adaptado. Springer [Fotografiado], por (W. Van der Aalst, 2016), <http://link.springer.com/10.1007/978-3-662-49851-4>

SISTEMA DE TRANSICIONES

Este sistema es el más básico de representar, por lo que consiste en una serie de estados y transiciones (Díez Arias, 2016, p. 30).



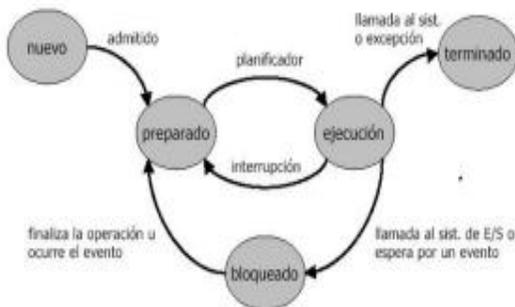


Figura 6. - Representación del sistema de transiciones.

Adaptado. Tesis [Fotografiado], por (Díez Arias, 2016), https://oa.upm.es/43904/1/TFG_FELIX_DIEZ_ARIAS.pdf

BPMN

Se ha convertido en uno de los lenguajes más utilizados para modelar procesos de negocio (W. Van der Aalst, 2016). Es un lenguaje que permite comunicar los procesos de manera clara, completa y eficiente; especialmente diseñada para coordinar secuencias de los procesos y el flujo que se recorre en las diferentes actividades (Velásquez L., 2017).

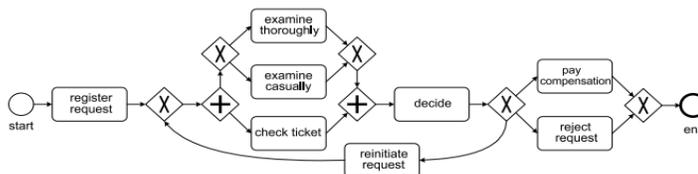


Figura 7. - Modelo de proceso utilizando la notación BPMN

Adaptado. Springer [Fotografiado], por (W. Van der Aalst, 2016), <http://link.springer.com/10.1007/978-3-662-49851-4>

FUZZY NET

Simplifica el modelo completo conservando eventos o bordes altamente significativos, agregando bordes y nodos menos significativos, pero altamente correlacionados mediante el agrupamiento y abstrayendo bordes y nodos menos significativos y poco correlacionados al eliminarlos del modelo simplificado. (Bogarín et al., 2018, p. 7)

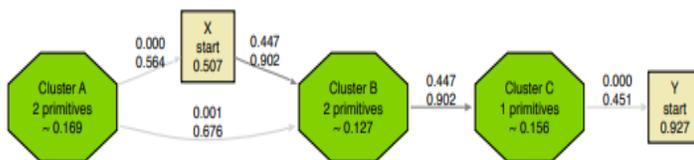


Figura 8. - Modelo de proceso utilizando Fuzzy Miner

Adaptado. Springer [Fotografiado], por (Günther & van der Aalst, 2007), http://link.springer.com/10.1007/978-3-540-75183-0_24



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

HEURISTIC NETS

Un gráfico de ciclo dirigido que representa los comportamientos más frecuentes de las actividades del conjunto de datos utilizado. En las redes heurísticas las cajas cuadradas representan las acciones y los arcos/enlaces representan dependencias/relaciones entre acciones. (Bogarín et al., 2018, p. 7)

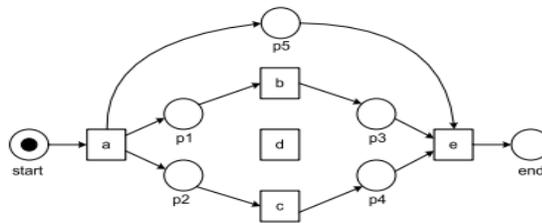


Figura 9. - Modelo de proceso utilizando Heuristic Net

Adaptado. Springer [Fotografiado], por (W. Van der Aalst, 2016), <http://link.springer.com/10.1007/978-3-662-49851-4>

IMPORTANCIA DEL PROCESS MINING

La minería de proceso permite descubrir información valiosa que mejora la productividad de la empresa, teniendo un impacto positivo dentro del workflow empresarial. Existen tres principales indicadores claves que permiten a la empresa mejorar su rendimiento:

- **KPIs de tiempo:** ¿Cuánto tiempo se tarda en completar un proceso concreto?
- **KPIs de costes:** ¿Cuánto cuesta completar un proceso concreto?
- **KPIs de calidad:** ¿El resultado de este proceso cumple los criterios establecidos? (Suárez, 2021)

PROCESO ETL

El proceso ETL es una etapa fundamental en la gestión y tratamiento de grandes cantidades de datos. Esta técnica permite extraer datos de una fuente, transformarlos y cargarlos en su destino correspondiente (Gutiérrez C, 2020, p. 70). Los entornos de origen y destino pueden ser bases de datos, archivos, colas de mensajes, entre otros. La transformación de los datos puede incluir cambios en su estructura, formatos o limpieza de los mismos para que sean compatibles con su destino. Además, el proceso ETL se basa en técnicas de consolidación que permiten agrupar datos de diferentes fuentes en un solo lugar (Curto Diaz, 2019).

Es importante destacar que el diseño del proceso de ETL debe ser cuidadosamente planificado y ejecutado para asegurar la calidad y consistencia de los datos en el almacén. Además, la implementación de un proceso de ETL eficiente es fundamental para el éxito de cualquier proyecto de integración de datos.



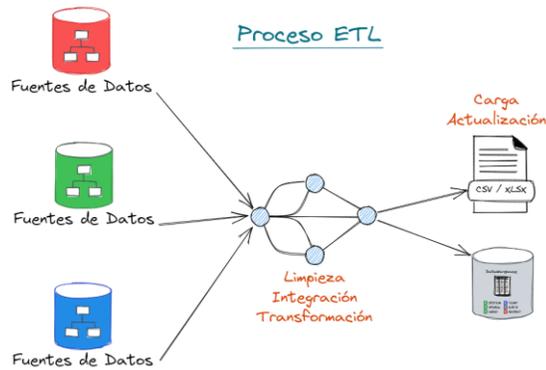


Figura 10. – Marco de trabajo para formar el log de eventos
Fuente: Fotografiada por Autor de la investigación, 2023

Resultados y discusión

El objetivo principal de la investigación es verificar la conformidad entre el modelo obtenido y el modelo de la empresa; pero para poder lograr cumplir ese objetivo se cumplieron varios pasos, tomando como referencia la metodología de Aguirre Mayorga & Rincón García (2015) propuesta en su tesis doctoral, y se detalla a continuación:

Definición del proyecto

La etapa de definición del proyecto es crucial para comprender el proceso en cuestión y sus principales dificultades. A través de esta comprensión detallada, es posible identificar los objetivos de mejora necesarios o las preguntas clave que deben ser abordadas mediante la aplicación de técnicas de minería de procesos. De esta manera, se puede planificar de manera efectiva el enfoque y el alcance de la minería de procesos en el proyecto, para obtener resultados óptimos. En esencia, esta fase sienta las bases para el éxito del proyecto de minería de procesos, permitiendo que se identifiquen claramente los problemas y las oportunidades de mejora en el proceso y se establezcan objetivos concretos y alcanzables.

Preparación de los datos

La etapa en cuestión se enfoca en el proceso ETL, el cual es fundamental en el proyecto. Este proceso se compone de tres fases clave: extracción, transformación y carga de datos. Para llevar a cabo esta tarea, se utiliza el software Pentaho, el cual facilita la extracción automática de información a través del uso de sentencias SQL. Cabe destacar que esta etapa es crítica para el éxito del proyecto, por lo que aquí es donde se trabaja con los datos crudos y sin limpiar, lo que significa que cualquier error o inconsistencia en esta fase podría tener un impacto significativo en los resultados finales. Por lo tanto, es crucial prestar atención a cada detalle durante esta etapa y asegurarse de que los



datos se estén extrayendo, transformando y cargando de manera precisa y eficiente; solo así se puede garantizar la calidad y la validez de los resultados obtenidos en las etapas posteriores del proyecto.

El proceso ETL comienza con el análisis y la localización de los datos necesarios para el proyecto, una vez identificados, se procede al proceso de extracción. En el caso particular de este proyecto, los datos proporcionados por la empresa se encuentran almacenados en una base de datos SQL Server de los 2010 - 2022. Esta base de datos se utiliza como fuente de datos principal para el proceso de limpieza.

Después de analizar detalladamente la información contenida en cada una de las tablas de la base de datos, se identifican las cuatro piezas clave que deben estar presentes en el log de eventos. Estas son:

- **Case ID:** se requiere un registro del objeto sobre el cual se realizó la acción para poder analizar cómo se están utilizando los diferentes objetos y detectar posibles problemas o errores en el proceso.
- **Actividad:** es importante registrar la actividad realizada por el usuario para poder entender cómo se está utilizando el sistema y detectar posibles errores o problemas en el proceso.
- **Recurso:** es necesario identificar el usuario que realizó cada evento para poder analizar el comportamiento de cada usuario individualmente y detectar posibles patrones.
- **Tiempo:** se requiere un registro preciso de la fecha y hora de cada evento registrado para poder realizar un análisis temporal adecuado.

Estas cuatro piezas clave permiten registrar de manera precisa y completa cada evento en el *log*, lo que facilita su posterior análisis y permite detectar patrones y tendencias que puedan ser útiles para la optimización del proceso.

Tabla 1. - Piezas claves para formar el log de eventos.

Piezas	Atributos
Case ID	Día Juliano
Actividad	Puntos críticos de control
Recurso	Usuario
Tiempo	Día de producción

Fuente: Autor de la investigación, 2023

Como CASE ID se toma en consideración el día juliano, debido a que es un dato registrado durante todo el proceso productivo de la compañía, las actividades se fueron localizando considerando los puntos críticos de control de su proceso HACCP y de acuerdo con el diagrama de la empresa se tiene (Figura 11):



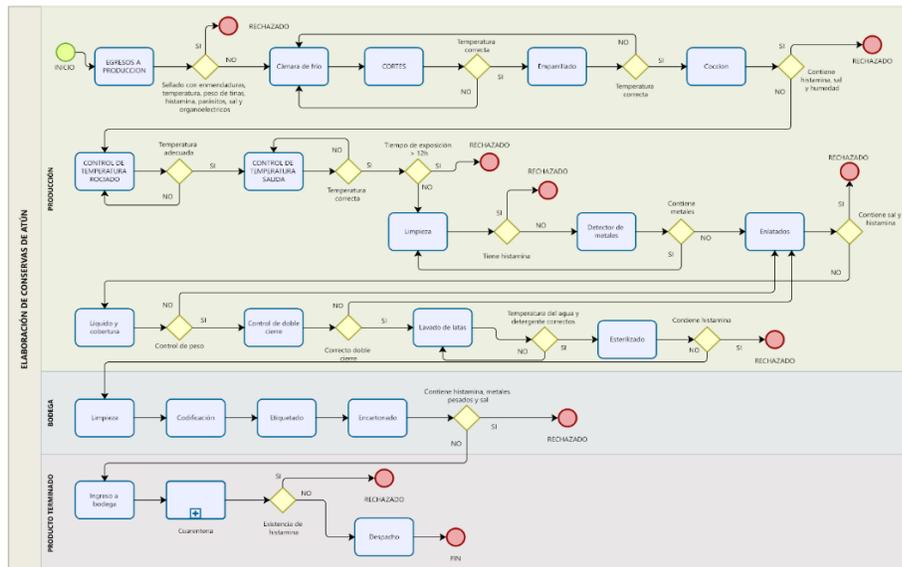


Figura 11. - Diagrama BPM del proceso de producción.

Fuente: Empresa atunera, 2023

Al identificar los puntos críticos de control basados en el diagrama BPM del proceso de producción de la empresa atunera se procede a identificar cada proceso dentro de la data proporcionada; una vez culminado el análisis y localización de la información se obtiene que dentro de la base de datos no se encuentran registrados los procesos; esto es debido a que aún no se tiene automatizado ese punto de control o no se ha registrado los eventos de dichas actividades; llegando al resultado siguiente:

Tabla 2. - Puntos críticos de control encontrados en la base de datos

Puntos críticos de control
Egresos de producción
Cortes
Cocción
Control de temperatura - Rociado
Control de temperatura - Salida
Limpieza
Control de doble cierre
Encartonado
Ingreso a bodega
Despacho

Fuente: Autor de la investigación, 2023



Después de identificar los puntos críticos de control dentro del workflow empresarial mediante el análisis de la base de datos, se procedió a llevar a cabo el proceso de limpieza de los datos. El proceso de limpieza se dividió en varios pasos desde el procesamiento de 17701251 transacciones comprendida entre los años 2010 al 2022; luego se filtra solo el año 2022 y debido a las limitaciones de los softwares utilizados en cuanto a la cantidad de datos aceptados, se limitó la selección a los primeros tres meses (enero a marzo), obteniendo 8016 eventos, comprendidas en 83 casos y repartidas en las 10 actividades o puntos críticos de control del flujo empresarial. Además, se agruparon los datos para evitar la duplicidad de la ejecución de las actividades realizadas por el usuario.

Tabla 3. - Descripción del total de datos obtenidos por cada actividad de acuerdo con el filtrado

Actividades	2010 - 2022	2022	Enero - Marzo	Agrupado
Egresos de producción	460073	18323	12610	72
Cortes	1226334	40498	28263	205
Cocción	297516	2945	2042	137
Control de temperatura - Rociado	5419	158	112	72
Control de temperatura - Salida	5419	158	112	72
Limpieza	14660506	2394962	1375948	131
Control de doble cierre	360861	78621	46770	256
Encartonado	111072	15397	9983	5067
Ingreso a bodega	43150	3889	3852	162
Despacho	530901	22555	13110	1842

Fuente: Autor de la investigación, 2023

Con todo este proceso de identificación y limpieza de la data, se obtuvieron un total de 8016 eventos de todas las actividades. Una vez analizada y limpiada toda la información se debe de obtener el log de evento en un formato que las aplicaciones de minería de procesos puedan aceptar; en este caso todos los datos obtenidos se exportan a formato CSV a través de la herramienta de Pentaho, como se denota a continuación:



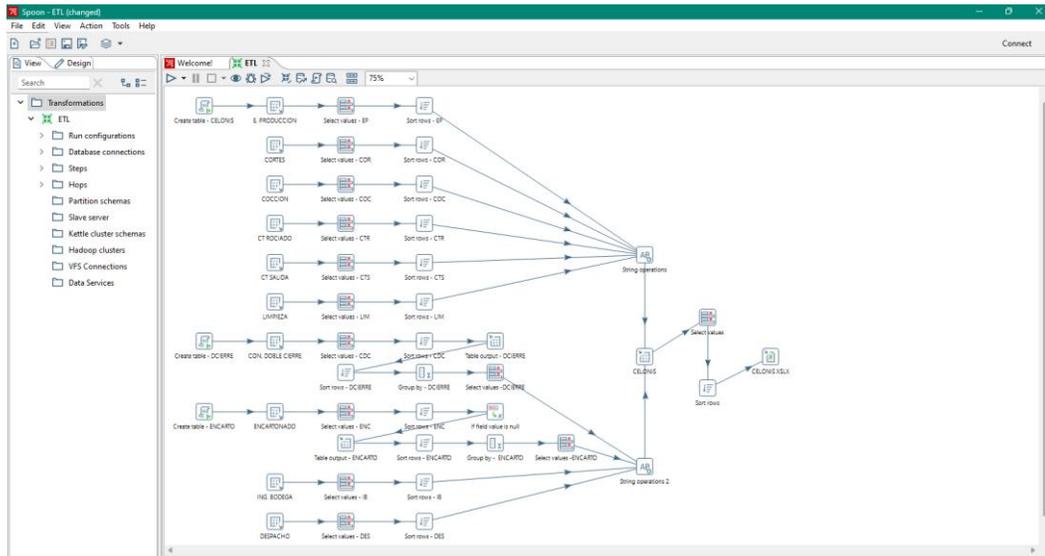


Figura 12. - Diagrama ETL en Pentaho.

Fuente: Autor de la investigación, 2023

Una vez formada el esquema para extraer los datos desde varias fuentes de información se obtiene un archivo en formato CSV, con los datos esenciales para conocer el modelo de acuerdo a la data obtenida.

Análisis de los procesos

Los resultados obtenidos del modelo de minería de proceso utilizando las herramientas Celonis, ProM y Disco han demostrado ser altamente efectivos para analizar y optimizar los procesos empresariales. Celonis, por ejemplo, ha permitido identificar cuellos de botella en los procesos de producción de la empresa, lo que permite mejorar significativamente la eficiencia y reducción de tiempos entre las actividades. Al utilizar la herramienta de Celonis se obtiene el modelo (*Discovery*) como se identifica a continuación:



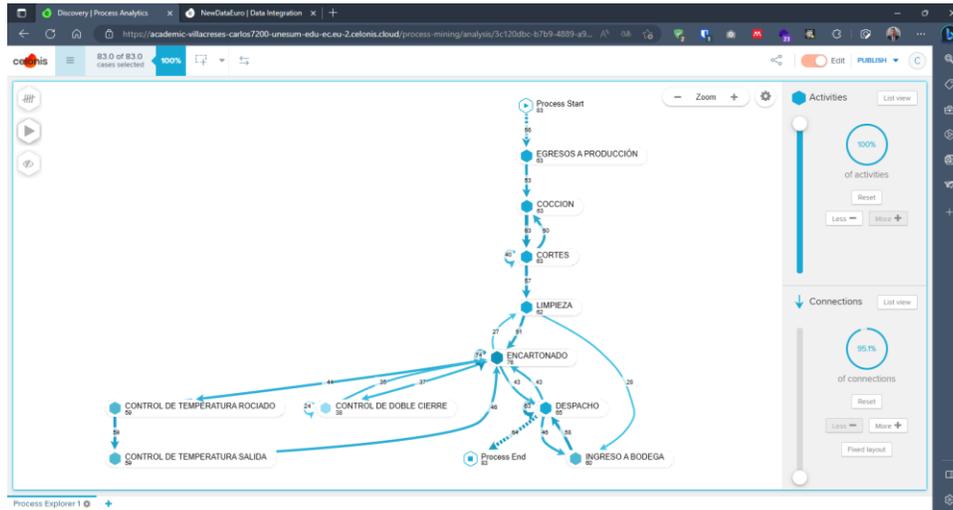


Figura 13. – Descubrimiento del modelo de acuerdo a la data obtenida.

Fuente: Fotografiada por Autor de la investigación, 2023

Una vez que se ha obtenido el modelo base a partir de la aplicación de la minería de procesos, se pueden realizar múltiples descripciones detalladas del mismo. Para ello, se pueden utilizar diferentes herramientas y visualizaciones que ofrece Celonis, permitiendo explorar el modelo y analizarlo en profundidad. En las siguientes imágenes se pueden observar algunas de estas descripciones, que incluyen desde la visualización del flujo de actividades y sus relaciones, hasta la identificación de cuellos de botella y la evaluación del desempeño del proceso en términos de tiempos. Estas descripciones son fundamentales para comprender en detalle el proceso en cuestión y poder tomar decisiones informadas sobre cómo mejorarlo y optimizar su desempeño.

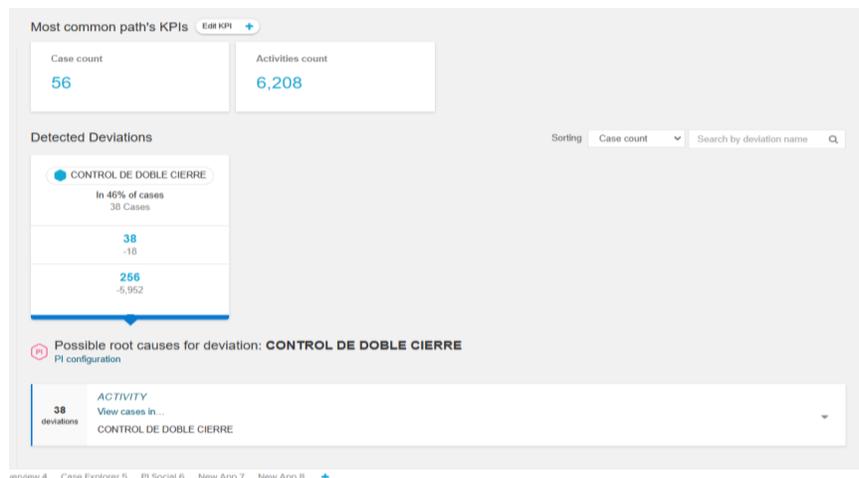


Figura 14. – Desviación en la actividad **Control de doble cierre** obtenida en Celonis.

Fuente: Fotografiada por Autor de la investigación, 2023.



Al realizar una revisión del proceso general dentro de Celonis se obtiene los siguientes resultados: 1 caso ejecutado por día, 91 eventos totales por día y la duración media de todos los casos desde el inicio hasta el fin, da un total de 40 días.

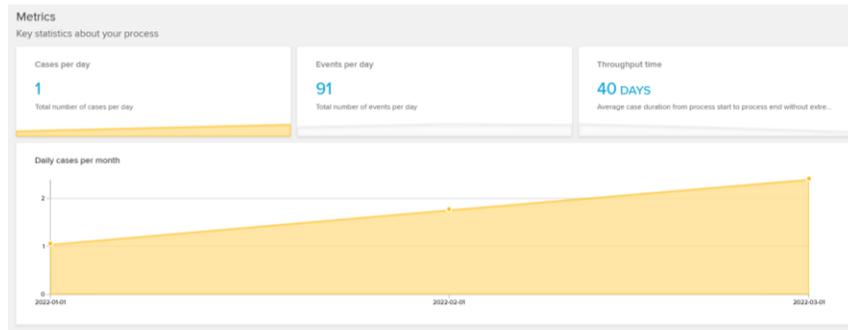


Figura 15. – Descripción general de los procesos en Celonis

Fuente: Fotografiada por Autor de la investigación, 2023.

Por consiguiente, el tiempo promedio de rendimiento por los casos desde inicio a fin se ejecutan en 40 días (Figura 16), en cambio el tiempo máximo da un total de 87 días.

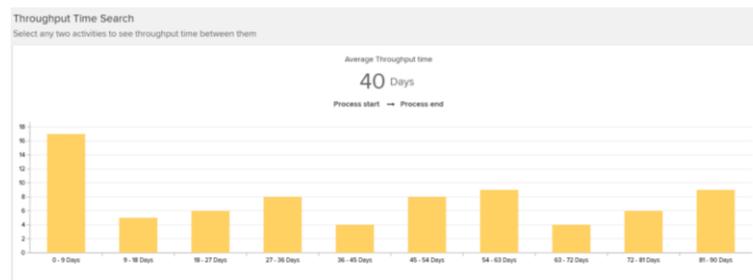


Figura 16. – Tiempo promedio por ejecución de todas las actividades

Fuente: Fotografiada por Autor de la investigación, 2023

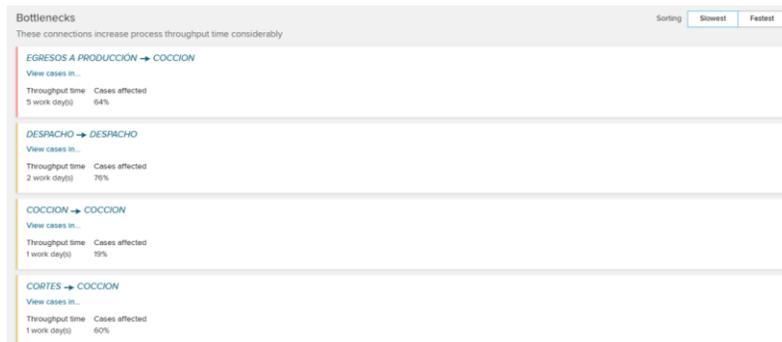


Figura 17. – Cuellos de botellas encontrados dentro del workflow de producción.

Fuente: Fotografiada por Autor de la investigación, 2023





Figura 18. – Cuello de botella con mayor impacto
Fuente: Fotografiada por Autor de la investigación, 2023

En la Figura 18 se muestra que existe un cuello de botella entre la actividad **EGRESOS A PRODUCCIÓN** → **COCCIÓN**. Esta actividad está tardando significativamente más tiempo en completarse en comparación con las demás actividades en el proceso. Esto puede deberse a diversas razones como una alta demanda de esta actividad, falta de recursos suficientes para realizarla o posibles problemas en su ejecución. La identificación de este cuello de botella es importante por lo que puede afectar el rendimiento y la eficiencia de todo el workflow. Es importante que se tomen medidas para solucionar los cuellos de botellas encontrados y optimizar los procesos.

Se ha identificado que la cantidad de casos ejecutados por cada actividad y la más frecuente es **Encartonado**, con una representación del 63% del total de ejecuciones. El segundo evento por actividad es **Despacho** con un 23%; mientras que con un 3% está **Control de doble cierre**; y el resto de actividades tienen menos del 3% de ejecuciones.

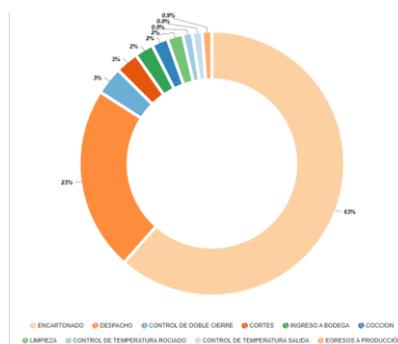


Figura 19. – Cantidad de casos ejecutados por actividad
Fuente: Fotografiada por Autor de la investigación, 2023

En la verificación de la conformidad en Celonis, se debe tener en cuenta que el modelo base debe estar en formato BPMN y ser lo más detallado posible, como se muestra en la Figura 20. Una vez que se tiene este modelo base, se procede a cargar la data principal en Celonis para obtener un modelo de proceso específico de la organización. A partir de aquí, se puede comparar ambos modelos y realizar la verificación de la conformidad. Para ello, se utiliza el



Conformance Checking, el cual permite identificar discrepancias entre el modelo base y el modelo específico de la organización, y de esta manera detectar posibles problemas en el proceso.

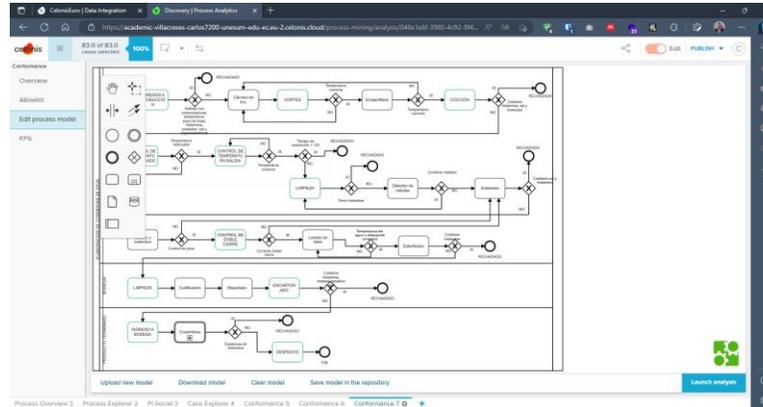


Figura 20. – Modelo base en formato BPMN

Fuente: Fotografiada por Autor de la investigación, 2023

Identificado los puntos críticos de control se procede a ejecutar el análisis; teniendo como resultado un porcentaje de similitud de 6%, dando como representación un indicador muy bajo con respecto al conformidad entre ambos modelos. Sin embargo, el 94% de los casos no respetan el proceso real que se tiene dentro de la empresa y pertenecen al grupo con un tiempo de duración mayor que los demás, alterando el tiempo promedio del workflow.

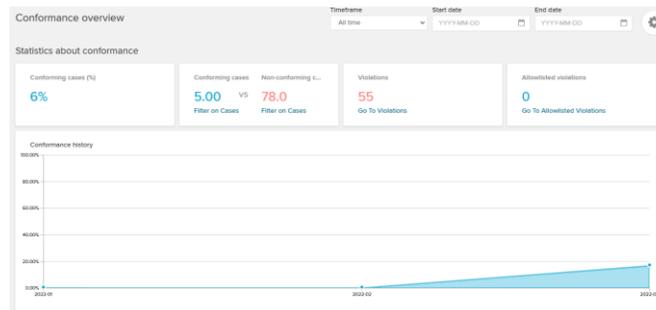


Figura 21. – Descripción de la fase de verificación de la conformidad

Fuente: Fotografiada por Autor de la investigación, 2023

En este caso quiere decir que no se sigue el flujo o el tiempo de pasar una actividad a otra es demasiado alto la transición.



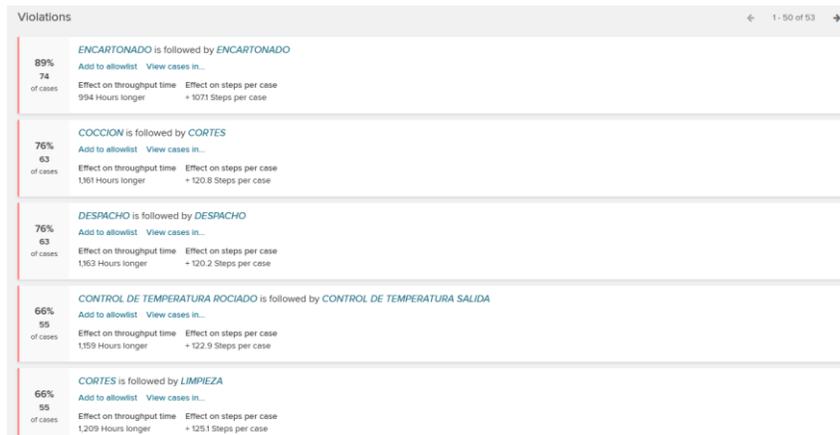


Figura 22. – Secuencia de actividades no coincidentes con el proceso documentado

Fuente: Fotografiada por Autor de la investigación, 2023

Rediseño del proceso

La revisión de los diagramas de proceso permitió detectar discrepancias entre el proceso teórico y el proceso real. En algunos casos, se identificó que se estaban infringiendo las reglas de negocio, como la falta de registro de algunos procesos o el exceso de tiempo en su ejecución. Es importante considerar la mejora de estos procesos para cumplir con las reglas de negocio en el menor tiempo posible y evitar que afecten el flujo de trabajo de procesos posteriores. La optimización de los procesos puede llevar a una mayor eficiencia y reducción de costos en la empresa.

Discusión

En base a los criterios de los autores empleando en la investigación como lo menciona Jans et al. en el (2021), que los modelos son procesos que tienen un enfoque mixto, donde los procedimientos, se orientan en la minería de procesos, de forma que permiten capturar los comportamientos que pueden ser complejos como no; mencionando que los modelos son los nuevos paradigmas que se integran en la verificación de conformidad; por otro lado, los autores Dunzer et al. (2019), mencionan que los modelos que se emplean en la minería de procesos, son instancias que se usan para la comprobación de la conformidad, además de ser instancias de un modelo que estudia a futuro los comportamientos y cumplimiento de las organizaciones en los procesos que a menudo usan.

Conclusiones

Una etapa clave dentro del proceso de indagación es la revisión literaria sistemática, la cual se lleva a cabo mediante una metodología de investigación rigurosa. Esto implica seleccionar y analizar exhaustivamente libros, tesis, artículos académicos en relación a la temática, en particular aquellos que contengan modelos y metodologías útiles que puedan



agregar mayor valor al modelo que se está diseñando. Las fuentes bibliográficas se utilizan como referencia y guía en el proceso de diseño, permitiendo incorporar buenas prácticas y estrategias eficaces para lograr un modelo robusto y eficiente. Por tanto, se puede afirmar que la revisión literaria es fundamental en el diseño de un modelo de calidad y relevante para el ámbito de aplicación en el que se utilizará.

Asimismo, es importante destacar que la limpieza y el procesamiento de la data es una etapa fundamental en la investigación y análisis de los datos. Si la data no se encuentra en un formato adecuado o contiene errores, el análisis no podrá ser confiable y podría llevar a conclusiones erróneas. Por lo tanto, se debe prestar atención a esta etapa y asegurarse de que se realice correctamente para garantizar que el modelo obtenido sea preciso y tenga una base sólida para su posterior estudio. En resumen, la limpieza y procesamiento de la data son una parte crucial en el proceso de análisis de datos y deben ser realizados de manera rigurosa para garantizar la validez de los resultados obtenidos.

Finalmente, la aplicación de técnicas de minería de procesos, como la creación del modelo base y el análisis del Conformance, permitió identificar posibles cuellos de botella y desviaciones en el proceso que afectan el rendimiento y la eficiencia. A través de la comparación del proceso teórico con el proceso real, se pudieron identificar las principales áreas de mejora y oportunidades para optimizar los procesos. Como resultado, se encontró que no todos los puntos críticos de control se encuentran automatizados y que el flujo obtenido con el proceso real no cumple con la conformidad, obteniendo así un 6% de los casos que sí siguen el flujo. Mientras tanto, el porcentaje restante no tiene registros de los procesos o mantiene un exceso de tiempo en su ejecución.

Conflictos de intereses

Los autores no poseen conflictos de intereses.

Contribución de los autores

1. Conceptualización: Carlos Andrés Villacreses Parrales, Lenin Jonatan Pin García
2. Curación de datos: Carlos Andrés Villacreses Parrales, Lenin Jonatan Pin García
3. Análisis formal: Carlos Andrés Villacreses Parrales
4. Investigación: Carlos Andrés Villacreses Parrales, Lenin Jonatan Pin García
5. Metodología: Lenin Jonatan Pin García
6. Administración del proyecto: Lenin Jonatan Pin García
7. Software: Carlos Andrés Villacreses Parrales, Lenin Jonatan Pin García
8. Validación: Carlos Andrés Villacreses Parrales, Lenin Jonatan Pin García



9. Visualización: Carlos Andrés Villacreses Parrales, Lenin Jonatan Pin García
10. Redacción – borrador original: Carlos Andrés Villacreses Parrales, Lenin Jonatan Pin García
11. Redacción – revisión y edición: Carlos Andrés Villacreses Parrales, Lenin Jonatan Pin García

Financiamiento

La investigación no requirió fuente de financiamiento.

Referencias

- Aguirre Mayorga, H. S., & Rincón García, N. (2015). Minería de procesos: desarrollo, aplicaciones y factores críticos. *Cuadernos de Administración*, 28(50). <https://doi.org/10.11144/Javeriana.cao28-50.mpda>
- Bogarín, A., Cerezo, R., & Romero, C. (2018). A survey on educational process mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(1), e1230. <https://doi.org/10.1002/widm.1230>
- Céspedes González, Y., Molero Castillo, G., & Arieta Melgarejo, P. (2018). Diseño De Una Guía De Análisis Para La Minería De Procesos Enfocada En El Usuario. *Revista Ciencia Administrativa*, 114–124. <http://ezproxy.ecci.edu.co:2111/ehost/detail/detail?vid=0&sid=bd9d8c98-77bd-43aa-b2d7-16b39d3af1d5%40sessionmgr4008&bdata=Jmxhbm9ZXMmc2l0ZT1laG9zdC1saXZI#AN=132947744&db=bth>
- Cocconi, D., Pérez, M., Ferreyra, J. P., & Verino, C. (2019). Aplicación de la minería de procesos para determinar recursos computacionales en la ejecución de procesos de negocio. *XXI Workshop de Investigadores En Ciencias de La Computación*. <http://sedici.unlp.edu.ar/handle/10915/77070>
- Contreras, N., Vergara, C., & Montenegro, Y. (2016). Process Mining para Gestión de Incidencias en Telefónica Chile. *Revista Ingeniería de Sistemas*, 30, 20.
- Curto Diaz, J. (2019). Diseño de procesos ETL (p. 48).
- Díez Arias, F. (2016). Aplicación de la minería de procesos al análisis y modelado de misiones multi-robot.
- Dunzer, S., Stierle, M., Matzner, M., & Baier, S. (2019). Conformance checking. *Proceedings of the 11th International Conference on Subject-Oriented Business Process Management*, 1–10. <https://doi.org/10.1145/3329007.3329014>
- González, C. M. (2016). Minería de Procesos: en ambientes sensorizados. <https://e->



archivo.uc3m.es/bitstream/handle/10016/23451/TFG_CarlosMiguel_Gonzalez_Escobosa.pdf

- Günther, C. W., & van der Aalst, W. M. P. (2007). Fuzzy Mining – Adaptive Process Simplification Based on Multi-perspective Metrics (pp. 328–343). https://doi.org/10.1007/978-3-540-75183-0_24
- Gutiérrez C, G. (2020). Modelo de datos y automatización del proceso ETL para dar apoyo a la metodología NAUTIA desarrollada por la Plataforma UPM sobre Refugiados [Politécnica de Madrid]. https://oa.upm.es/64075/1/TFG_GUILLERMO_SANCHEZ_GUTIERREZ_CABELLO.pdf
- Jans, M., Weerdt, J. De, Depaire, B., Dumas, M., & Janssenswillen, G. (2021). Conformance Checking in Process Mining. *Information Systems*, 102, 101851. <https://doi.org/10.1016/j.is.2021.101851>
- Reinkemeyer, L. (2020). Process Mining in a Nutshell. In *Process Mining in Action* (pp. 3–10). Springer International Publishing. https://doi.org/10.1007/978-3-030-40172-6_1
- Suárez, R. (2021). Minería de procesos: Qué es, ventajas, importancia y usos. Tudashboard.Com. <https://tudashboard.com/mineria-de-procesos/>
- Van der Aalst, W. (2016). Process mining: Data science in action. In *Process Mining: Data Science in Action*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-49851-4>
- van der Aalst, W. M. P. (2011). Process Mining. In *Process Mining* (Vol. 136, Issue 2). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-19345-3>
- Velásquez L., E. E. (2017). Método de Minería de Procesos para incrementar la Productividad para proyectos de investigación de las cajas municipales. <http://repositorio.unsa.edu.pe/bitstream/handle/UNSA/4504/IIMveloe.pdf?sequence=1&isAllowed=y>
- Wil Van der, A., Adriansyah, A., & Alves De Medeiros. (2011). Manifiesto sobre Minería de Procesos. Manifiesto Sobre Minería de Procesos, 0(0), 1–21. <http://www.win.tue.nl/ieeetfpm>

