



---

Universidad Autónoma del Estado de México  
Centro Universitario UAEM Ecatepec

**TESIS:**  
**Minería de Procesos para la detección de áreas de  
oportunidad en el proceso de atención de servicios básicos al cliente de la  
Comisión Federal de Electricidad.**

**P R E S E N T A:**  
Carmen Ibeth Martínez Aragón

Lic. en Informática Administrativa para obtener el grado de  
maestría en Ciencias de la Computación

**Comité de tutores:**

**Director de tesis:** Dra. en SC. Alejandra Morales Ramírez

**Co-director de tesis:** Dr. en SC. Cuauhtémoc Hidalgo Cortés

**Tutor de tesis:** Dr. en C.C Juan de Jesús Amador Reyes

**Ecatepec de Morelos, Estado de México, octubre 2025.**

## Contenido

Índice de tablas .....	12
Índice de figuras.....	12
Índice de graficas .....	16
I. Introducción .....	17
II. Objetivo general.....	18
III. Objetivos específicos .....	18
1. Minería de procesos .....	21
1.1. Propósitos y principales razones para aplicar Minería de Procesos .....	23
1.1.2. Objetivos y funciones de la Minería de Procesos .....	24
1.2. Tipos de Minería de Procesos .....	25
1.2.1. <i>Discovery</i> (Descubrimiento).....	26
1.2.2. Conformance (Conformidad).....	26
1.2.3. Enhancement (Mejora): .....	27
1.3. Tipos de proyectos de minería de procesos .....	28
1.4. Elementos de los registros de eventos. ....	29
1.4.1. Id de caso .....	31
1.4.2. Marca de tiempo .....	31
1.4.3. Otros elementos .....	32
1.4.4. Evento .....	32
1.4.5. Caso.....	32
1.4.6. Variantes .....	34
1.4.7. Actividad.....	34
1.4.8. Acción.....	35
1.4.9. Datos necesarios para realizar el análisis.....	35

1.5. Herramientas informáticas para la minería de procesos .....	37
2. Algoritmos y herramientas de modelado de la minería de procesos .....	41
2.1. Algoritmos de modelado .....	41
2.1.1. Algoritmo básico (Algoritmo “Alpha”) .....	41
2.1.2. Limitaciones del algoritmo Alpha .....	51
2.2. Algoritmos avanzados de descubrimiento.....	55
2.2.1. Algoritmo Alpha +.....	55
2.2.2. Algoritmo minado heurístico .....	55
2.2.2.1. Minado heurístico con parámetros por default .....	63
2.2.2.2. Validez del descubrimiento.....	66
2.2.3. Algoritmo minado inductivo básico e infrecuente.....	67
2.2.4. Algoritmo minado inductivo básico (Inductive miner IM).....	67
2.2.4.2 Arboles de proceso.....	68
2.2.4.3. Minado inductivo infrecuente.....	80
2.2.5. Algoritmos basados en región.....	82
3. Modelos de procesos.....	86
3.1. Modelo red Petri.....	87
3.1.1. Patrones que forman una red Petri .....	88
3.1.2. Desplazamiento del token a través del modelo.....	90
3.2. Modelo BPMN .....	92
3.3. Modelo EPC .....	93
3.4. Modelo red causal (Causal nets o C-Nets).....	94
3.5. Diferencia entre de las C-nets y las redes Petri.....	101
3.6. Tipos de procesos .....	102
4. Modelos o métodos utilizados para llevar a cabo la minería de procesos. ....	107

4.1. Modelo de ciclo de vida L (Life Cycle Model) .....	107
4.2. Método de minería de procesos de Velázquez.....	109
4.3. Metodología de Aguirre.....	110
4.4. Metodología para la aplicación de minería de procesos de Bozcaya.....	112
5. Aplicación de la metodología .....	114
5.1. Caso de estudio.....	114
5.1.1. Definición del caso .....	114
5.1.2. Proceso de atención de los trabajos programados.....	121
5.2. Etapa 1 “Preparación del registro”.....	122
5.3. Etapa 2 “Inspección de registros” .....	124
5.4. Etapa 3 “Control de análisis de flujo” .....	126
5.6. Etapa 4 “Análisis de rendimiento” .....	133
5.7. Etapa 5 “Análisis de roles”.....	138
Conclusiones.....	142
Glosario.....	144
Referencias.....	145

## Índice de tablas

Tabla 1. Objetivos de la Minería de Procesos.....	24
Tabla 2. Perspectivas Minería de Procesos.....	27
Tabla 3. Matriz de huella (Footprint).....	45
Tabla 4. Matriz de huella Ejercicio 1.....	48
Tabla 5. Matriz de huella ejercicio 2. ....	50
Tabla 6. Funciones.....	81
Tabla 8. Descripción de los trabajos programados. ....	117
Tabla 9. Datos seleccionados para conformar el registro de eventos. ....	123
Tabla 10. Extracto de datos obtenidos de las solicitudes.....	124
Tabla 11. Actividades descubiertas en el proceso. ....	125
Tabla 13. Variantes. ....	136
Tabla 14. Frecuencia de tiempos. ....	137
Tabla 15. Frecuencias de los recursos dentro del proceso. ....	138

## Índice de figuras

Figura 1. Registros de eventos utilizados por los tres tipos de minería de procesos. ....	23
Figura 2. Tipos de Minería de Procesos en términos de entradas y salidas.....	26
Figura 3. Datos de entrada de minería de procesos. ....	30
Figura 4. Elementos mínimos de un registro de eventos. ....	30
Figura 5. Ejemplo de id de caso.....	31
Figura 6. Ejemplo marca de tiempo de los registros de eventos.....	31
Figura 7. Atributos de un registro de eventos. ....	32
Figura 8. Ejemplo de evento. ....	32
Figura 9. Ejemplo de Casos diferentes. ....	33
Figura 10. Ejemplo de Variante. ....	34
Figura 11. Ejemplo de actividades de un registro de eventos.....	35
Figura 12. Acción.....	35
Figura 13. Registros de un tiempo específico. ....	36
Figura 14. Registro de eventos de un periodo determinado.....	37

Figura 15. Resultado: red Petri. ....	48
Figura 16. Red Petri ejercicio 1. ....	49
Figura 17. Red Petri ejercicio 2. ....	51
Figura 18. Lugares implícitos. ....	52
Figura 19. Bucle longitud uno. ....	53
Figura 20. Bucle longitud dos. ....	53
Figura 21. Dependencias no locales. ....	54
Figura 22. Red Petri con etiqueta de actividad única. ....	54
Figura 23. Modelo en flor. ....	55
Figura 24. Sucesión directa. ....	56
Figura 25. Fórmula medida de dependencia. ....	57
Figura 26. Gráfico de dependencia. ....	57
Figura 27. Patrón de secuencia. ....	58
Figura 28. Frecuencia de actividades. ....	59
Figura 29. Medidas de dependencia. ....	60
Figura 30. Grafica de dependencia usando umbrales de 2 y 0.7. ....	60
Figura 31. Gráfico de dependencia con umbrales de 5 y .9. ....	61
Figura 32. Modelo de procesos usado para la generación de registros de eventos con y sin ruido. ....	62
Figura 33. Modelo con 1000 registros y con parámetros de umbral predeterminados. ....	64
Figura 34. Modelo con umbrales umbral relativo al mejor =0.20, umbral de dependencia= 0.85. ....	65
Figura 35. Elementos del diagrama de flujo de trabajo. ....	68
Figura 36. Operador secuencial. ....	69
Figura 37. Operador XOR. ....	69
Figura 38. Operador paralelo. ....	70
Figura 39. Operador bucle. ....	70
Figura 40. Árbol de procesos. ....	71
Figura 41. Arbol de proceso de siete actividades. ....	73
Figura 42. Conversion del árbol de procesos de 7 actividades a red Petri. ....	73
Figura 43. Árbol de procesos de cinco actividades. ....	74

Figura 44. Conversión del árbol de procesos a red Petri. ....	74
Figura 45. Pasos del minado inductivo. ....	75
Figura 46. División de elección exclusiva. ....	77
Figura 47. División de secuencia. ....	78
Figura 48. División simultánea. ....	78
Figura 49. División de bucle. ....	79
Figura 50. Ejemplo de modelo en flor. ....	80
Figura 51. Minado poco infrecuente. ....	81
Figura 52. Diagrama por regiones. ....	83
Figura 53. Diagrama resultante de las abstracciones. ....	83
Figura 54. Diagrama final. ....	83
Figura 55. Red Petri equivalente al diagrama final. ....	84
Figura 56. Red Petri. ....	87
Figura 57. Elementos de los modelos BPMN. ....	88
Figura 58. Patrón XOR separación. ....	88
Figura 59. Patrón XOR unión. ....	89
Figura 60. Patrón AND separación. ....	89
Figura 61. Patrón AND unión. ....	89
Figura 62. Paso 1 desplazamiento del token. ....	90
Figura 63. Paso 2 desplazamiento del token. ....	90
Figura 64. Paso 3 desplazamiento del token. ....	91
Figura 65. Paso 4 desplazamiento del token. ....	91
Figura 66. Finalización del desplazamiento del token. ....	91
Figura 67. Notación modelos BPMN. ....	92
Figura 68. Elección diferida. ....	93
Figura 69. Modelo BPMN. ....	93
Figura 70. Notación EPC. ....	94
Figura 71. Modelo EPC. ....	94
Figura 72. Notación Redes Causales. ....	95
Figura 73. Compuertas. ....	95
Figura 74. Ejemplo de red causal. ....	96

Figura 75. Ejemplo de red causal.....	98
Figura 76. Entrada a.....	98
Figura 77. Elección de enlace y creación de obligaciones.....	98
Figura 78. Ejecución de b y creación de una nueva obligación.....	99
Figura 79. Ejecución de d y creación de obligaciones.....	99
Figura 80. Ejecución de e y creación de obligación. ....	100
Figura 81. Ejecución de g y eliminación de obligación.....	100
Figura 82. Actividad final. ....	101
Figura 83. Ejemplo de proceso tipo lasaña. ....	103
Figura 84. Ejemplo de proceso tipo espagueti. ....	105
Figura 85. Modelo de ciclo de vida L (Life Cycle Model) propuesta por Van der Aalst, 2011. ....	108
Figura 86. Diagrama flujo del Método de minería de procesos propuesta por Velázquez, 2017. .....	109
Figura 87. Metodología de Aguirre, 2015. ....	111
Figura 89. Procesos de CFE.....	114
Figura 90. Sistemas para el registro de solicitudes.....	115
Figura 91. Tipos de solicitudes. ....	116
Figura 92. División de los trabajos programados. ....	116
Figura 93. Proceso de una solicitud. ....	121
Figura 94. Solicitud de suministro de servicios básicos al cliente de CFE.....	123
Figura 95. Modelo red Petri del proceso.....	126
Figura 96. Modelo EPC .....	127
Figura 97. Modelo BPMN. ....	128
Figura 98. Modelo red causal.....	129
Figura 99. Modelo descubierto. ....	131
Figura 100. Modelo de la empresa.      Figura 101. Modelo descubierto ajustado. ....	132
Figura 102. Tiempos de rendimiento del proceso principal. ....	133
Figura 103. Tiempos de rendimiento del proceso general.....	134
Figura 104. Cuellos de botella dentro del proceso.....	135

## **Índice de graficas**

Grafica 1. Frecuencia relativa de las actividades.....	126
Grafica 2. Tiempo de atención.....	137
Grafica 3. Frecuencias de participación de los recursos. ....	139

## **I. Introducción**

Cognizant (2024) señala que el uso de tecnología, como software, hardware, plataformas digitales y la automatización inteligente de procesos (Intelligence Process Automation, IPA), contribuye a la creación y entrega de productos y servicios, mejorando la velocidad y agilidad de los procesos en toda la empresa. Además, permite automatizar y simplificar los flujos de trabajo, acelerar los procesos de negocio y optimizar la precisión, calidad y fiabilidad de estos, facilitando que las compañías operen con mayor rapidez para adaptarse a los cambios del mercado y reducir costos operativos.

En la actualidad, las organizaciones no solo buscan subsistir, sino también destacar en un entorno altamente competitivo, dinámico e innovador. La innovación tecnológica se ha convertido en una herramienta clave para alcanzar este objetivo, posibilitando el desarrollo de sistemas de información más robustos, con capacidad suficiente para almacenar y gestionar grandes volúmenes de datos provenientes de clientes, proveedores, procedimientos y fuentes externas. Con el tiempo, los datos han pasado de ser un simple producto para convertirse en una materia prima fundamental que debe ser explotada para generar conocimiento.

Sin embargo, a medida que aumenta la cantidad de datos almacenados, la capacidad de las organizaciones para asimilarlos disminuye. Esto hace necesario el uso de herramientas avanzadas de análisis de datos que permitan extraer información útil de grandes volúmenes de datos. En este contexto, la minería de procesos y la minería de datos juegan un papel fundamental. La minería de procesos es una disciplina de investigación relativamente joven, situada en la intersección entre la inteligencia computacional, la minería de datos, la modelación y el análisis de procesos de negocio (González, Leal, Martínez y Morales, 2019; Van der Aalst, 2011). Su objetivo es analizar los procesos organizacionales a partir de los datos generados por las actividades que los conforman (Silva, 2017).

La minería de procesos busca descubrir, monitorear y mejorar los procesos organizacionales mediante la extracción de conocimiento de los registros de eventos disponibles en los sistemas de información (Aguirre y Rincón, 2015). Según Van der Aalst (2011), esta disciplina comprende tres técnicas principales:

- Descubrimiento: Genera un modelo del proceso a partir del registro de eventos sin necesidad de información previa.
- Verificación de conformidad: Compara el modelo del proceso existente con el registro de eventos para identificar desviaciones.
- Mejoramiento: Extiende o optimiza un modelo de proceso existente utilizando la información almacenada en los registros de eventos.

A pesar del gran potencial de la minería de procesos para generar conocimiento que optimice los negocios y aporte valor a los clientes, muchas organizaciones aún no han adoptado una cultura que fomente su uso. Esto conlleva diversos problemas, como falta de claridad en los procesos, imprecisión en los objetivos, sobrecarga de trabajo, desconocimiento del comportamiento de los clientes, asignación ineficiente de responsabilidades, ausencia de políticas y lineamientos, falta de seguimiento al personal, deficiencias en la capacitación, ejecución manual de tareas, uso inadecuado de los sistemas, entre otros (Windgassen, 2024).

En este contexto, el presente trabajo de investigación tiene como objetivo general y específicos los siguientes:

## **II. Objetivo general**

Demostrar la utilidad de las técnicas y herramientas de minería de procesos para la extracción de conocimiento y su aplicación en la identificación de áreas de oportunidad en el desempeño del proceso de atención de servicios básicos al cliente de la Comisión Federal de Electricidad (CFE).

## **III. Objetivos específicos**

- Obtener la información necesaria para el análisis
- Limpiar y preprocesar los datos
- Identificar técnicas (algoritmos) y herramientas de modelado que se utilizaran para el descubrimiento de modelos de procesos.
- Analizar las herramientas informáticas disponibles para la minería de procesos, evaluando su usabilidad y aplicabilidad.
- Examinar el proceso actual de atención de servicios básicos al cliente en la CFE.
- Generar un modelo de procesos a partir del registro de eventos.

- Comparar el modelo de proceso existente con el obtenido a partir del análisis del registro de eventos.
- Identificar trayectorias ocultas en el proceso.
- Determinar las actividades que requieren mayor tiempo dentro del proceso.
- Detectar cuellos de botella.
- Identificar tareas repetitivas.
- Analizar desviaciones en los procesos.
- Descubrir la productividad real del personal encargado de ejecutar el proceso.

# **CAPÍTULO I**

**Marco**

**Teórico**

La minería de procesos fue definida en el año 2011 por el Manifiesto de Minería de Procesos el cual fue escrito por un grupo de expertos respaldados por IEEE Task Force on Process Mining (Fuerza de Trabajo de la IEEE sobre Minería de Procesos), con el propósito de promover la investigación, desarrollo, implementación, evolución y entendimiento.

Esta disciplina es respectivamente joven y se coloca entre la inteligencia computacional y la minería de datos. Existen dos motivos primordiales para el creciente interés en esta disciplina: en primer lugar, cada vez más se registran eventos que proporcionan información detallada de los procesos y en segundo lugar existe una necesidad de mejorar y ayudar los procesos de negocio en ambientes competitivos que cambian precipitadamente (EcuRed, 2018; Van der Aalst, 2011 y Windgassen, 2024).

Por ejemplo, el modelo Auditoría 2.0 propone utilizar la mayoría de las técnicas de minería de procesos para el agrupamiento de datos, con el objetivo de detectar desviaciones, obtener varios modelos que se relacionen en vez de tener un único modelo y analizar los procesos desde una perspectiva en línea.

La predicción de comportamientos y el soporte operacional son un área relativamente nueva dentro del análisis de procesos y una de las áreas de investigación menos explotada por la minería de procesos, esta última se divide en tres etapas: la primera es detección, la segunda predicción y la tercera recomendación de comportamientos; es utilizada para el análisis de trazas finalizadas o casos históricos y no para casos actuales (Acosta, López, Vázquez y Velastegui, 2018).

## **1. Minería de procesos**

Con el propósito de entender un poco más la Minería de Procesos, se proponen algunas definiciones tomadas de diferentes fuentes para tener un conocimiento más amplio sobre el tema a desarrollar.

Martínez (2013) citado en Pérez (2015) y Hernández (2012) coinciden que la Minería de Procesos es una disciplina que tiene por objeto la detección y extracción automática de información valiosa de los registros de eventos para captar el proceso de negocio a medida que se ejecuta.

Por su parte Bose y Van der Aalst (2010) y Van der Aalst y otros (2012) y coinciden en que las técnicas de Minería de Procesos son capaces de extraer conocimientos de los registros de eventos comúnmente disponibles en los sistemas de información. Estas técnicas proporcionan un medio para descubrir procesos, analizar datos, monitorear, mejorar y analizar los procesos.

En tanto que Aguirre y Rincón (2015) y Van der Aalst (2011) aluden que la minería de procesos es una disciplina que tiene como objetivo utilizar técnicas capaces de extraer o generar conocimientos a partir de los registros que se encuentran en un sistema de información y son idóneos para descubrir, monitorear y mejorar los procesos, para encontrar desviaciones, construir modelos de simulación, predecir casos y dar recomendaciones.

En conclusión, se puede decir que la Minería de Procesos es una disciplina que tiene por objeto la extracción de información y generación de conocimientos a través de los registros de eventos, para descubrir, monitorear, mejorar los procesos y dar recomendaciones.

Turdibayeva (2024) indica que los registros de eventos son el modelo de datos central y la fuente de datos principal que es utilizada en la minería de procesos, son datos de transacciones en bruto que se recopilan y ajustan de diferentes sistemas para el análisis de los procesos.

Por lo que los registros de eventos (*event log*) son el fundamento principal de la minería de procesos y estos son posible registrarlos, tal que cada evento se refiera a una actividad específica en un proceso (ej. compra de un producto) y se relacione a un caso particular (ej. solicitud de proceso de registro de un producto, envió de orden de producto).

Los registros también pueden almacenar información adicional acerca de los eventos, de hecho, siempre que sea posible, se debe utilizar información extra con las técnicas de minería de procesos, tales como el recurso (usuario o dispositivo ej. Pedro Colín, Analista) que ejecuta o inicia la actividad, el registro de tiempo del evento (ej. 14:51:00 hrs.), inicio y fin de la actividad (ej. Inicio: 13/02/2021 y fin 16/02/2021), los elementos de datos registrados con el evento (ej. producto a comprar, proveedor, departamento, etc.) (Aguirre y Rincón, 2015).

Estos registros de eventos se pueden encontrar disponibles en los sistemas PAIS (Process Aware Information Systems), como lo son:

- Sistemas de workflow

- BPMS (Business Process Management Suite)
- ERP (Enterprise Resource Management)
- CRM (Customer Relationship Management)
- Otros.

En la figura 1 se muestran los registros de eventos del proceso de atención de servicios básicos al cliente de suministro eléctrico de la Comisión Federal de Electricidad que pueden ser utilizados por los tres tipos de Minería de Procesos.

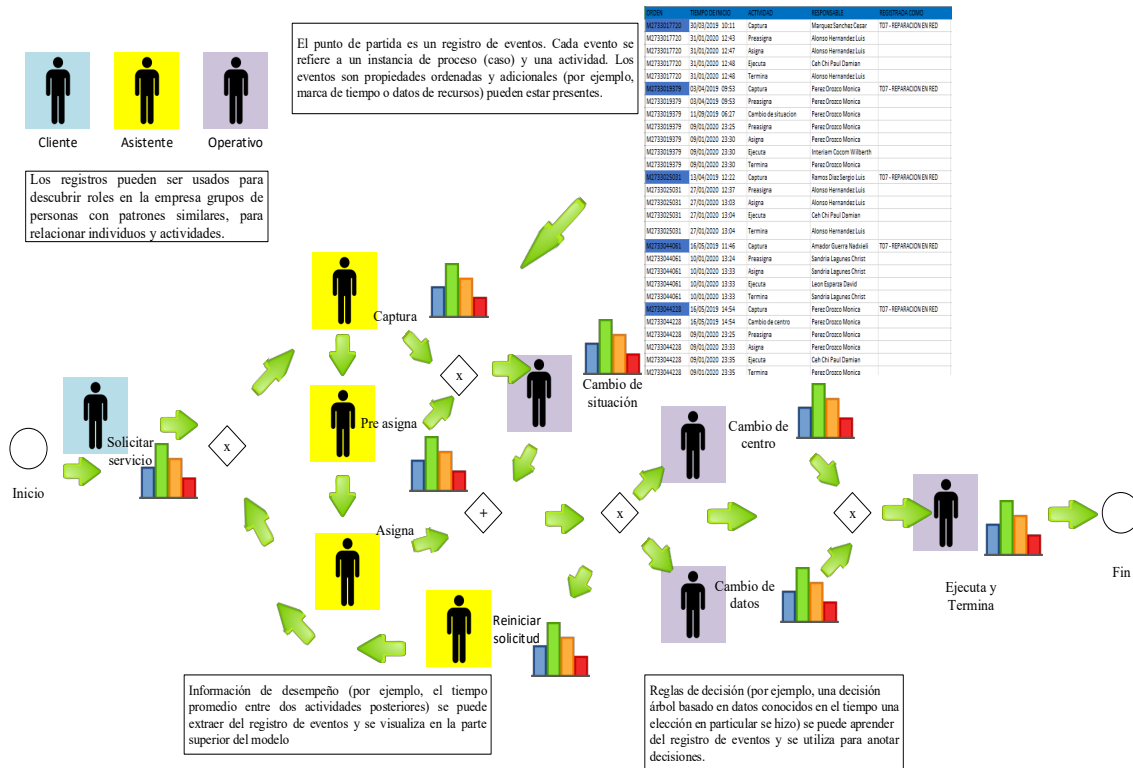


Figura 1. Registros de eventos utilizados por los tres tipos de minería de procesos.

### 1.1. Propósitos y principales razones para aplicar Minería de Procesos

La Minería de Procesos es particularmente útil en situaciones donde los eventos se registran, pero no existe un sistema que indique a las personas que deben realizar una actividad de una manera particular. Por lo anterior, la necesidad de las empresas para obtener más información acerca de cómo funcionan sus procesos, lo cual es un factor fundamental para las técnicas de minería de procesos (Hernández, 2012).

Los principales propósitos que destaca Pérez (2015) para aplicar la minería de procesos son los siguientes: a) Esquematisar los procesos que se desarrollan en la organización soportados en sistemas transaccionales, para compararlos con los procesos definidos; b) lograr la optimización del uso de recursos y la excelencia en las operaciones por medio del análisis, ajuste y mejora de los procesos; c) Monitoreo constante de los procesos críticos de la organización; d) utilizar los tres tipos básicos de minería de procesos: descubrimiento, la conformidad y la mejora; e) dar seguimiento a los registros de eventos de tal forma que puedan ser fácilmente exploradas; f) apoyar en el registro real de cumplimiento que las aplicaciones disponen como soporte a sus transacciones; g) facilitar la evaluación de procesos; i) proporcionar la evaluación del comportamiento de los procesos, identificación de cuellos de botella, flujos alternativos, actividades y/o tareas no realizadas.

### 1.1.2. Objetivos y funciones de la Minería de Procesos

En la tabla 1 se muestran los diversos objetivos de la Minería de Procesos según Aguirre y Rincón (2015) y Santiago (2015).

Tabla 1. Objetivos de la Minería de Procesos.

Objetivo	Descripción
<b>Descubrir el modelo de ejecución real del proceso.</b>	A través de algoritmos de Minería de Procesos aplicados al análisis de los registros de eventos, se puede descubrir y llegar al modelo real de un proceso.
<b>Determinar si el proceso cumple con la reglamentación y procedimientos documentados.</b>	Al tener el modelo real de un proceso se puede comparar con los procedimientos documentados para determinar si se está cumpliendo con los estándares, protocolos, reglamentación y políticas de ejecución de un proceso.
<b>Analizar la interacción del personal que ejecuta el proceso.</b>	A través de la aplicación de técnicas de minería de procesos se puede construir la red social del proceso (social network) para analizar la interacción entre los individuos y descubrir bucles (loops) que pueden demorar la ejecución de un proceso.

<b>Descubrir cuellos de botella.</b>	A través de la animación del flujo de los casos reales del proceso se pueden determinar los cuellos de botella para actuar sobre estos y mejorar el nivel de desempeño del proceso.
<b>Monitorear la productividad del personal.</b>	A través del análisis de los registros de ejecución del proceso se pueden descubrir los datos de productividad real de las personas que ejecutan el proceso, así como los tiempos de ciclo por actividad.
<b>Predecir el tiempo de ciclo de un caso.</b>	A través de la aplicación de técnicas de la minería como lo son los árboles de decisión se puede predecir el tiempo de ejecución restante de un proceso.
<b>Determinar la relación entre las variables de un caso.</b>	A través de la aplicación de técnicas de clasificación se puede determinar cómo diferentes variables asociadas a un caso (por ejemplo: tipo de producto a comprar, comprador, proveedor, departamento solicitante), pueden influir en los tiempos de ciclo de un proceso.

Nota: Elaboración propia tomado de Aguirre y Rincón, 2015.

## 1.2. Tipos de Minería de Procesos

La minería de procesos tiene como propósito extraer información clave sobre los procesos dentro de una organización, utilizando como fuente los datos almacenados en los registros de eventos. A partir de esta información, es posible llevar a cabo tres enfoques principales en minería de procesos (figura 2) (Gastón, 2007 y Loaiza, 2018;):

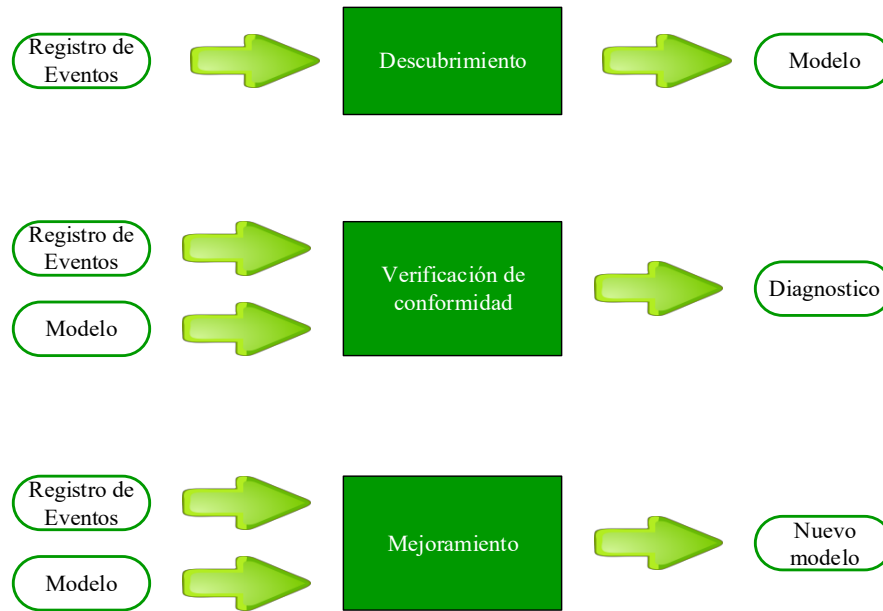


Figura 2. Tipos de Minería de Procesos en términos de entradas y salidas.

### 1.2.1. *Discovery* (Descubrimiento)

El primer tipo de minería de procesos es una técnica que utiliza el *event log* de un modelo específico para descubrir el modelo real, toma todos los registros y chequea donde existen más retrabajos; además de que el descubrimiento de procesos es el más destacado de los tres tipos de minería de procesos. Para muchas organizaciones es sorprendente ver que las técnicas existentes son realmente capaces de descubrir los procesos reales meramente basado en las muestras de ejecución de los registros de eventos.

Los modelos descubiertos según Loaiza (2018) pueden presentar la información desde las siguientes perspectivas:

Control de flujo: está orientada en el orden de ejecución de las actividades.

Perspectiva organizacional: se identifican las responsabilidades y la relación de los recursos involucrados en la ejecución de los procesos.

Datos: muestra información sobre el flujo de datos en la ejecución del proceso.

### 1.2.2. *Conformance* (Conformidad)

El segundo tipo de Minería de Procesos es la conformidad, es una técnica que se utiliza para verificar la conformidad del proceso virtual con el procedimiento real, es decir; se toma el modelo obtenido por el sistema y es comparado con el procedimiento real para encontrar desviaciones. La verificación de conformidad puede ser usada para chequear la realidad, tal como está almacenada en el registro de eventos, es equivalente al modelo y viceversa, mide el alineamiento entre el modelo y la realidad.

### 1.2.3. Enhancement (Mejora):

El tercer tipo de minería de procesos es la mejora que tiene como principal objetivo extender o mejorar el proceso, en él se pueden ver los tiempos de operación de cada actividad, los cuellos de botella que retrasan el proceso, ubicar los retrabajos para conocer donde se invierte más tiempo en una actividad, los responsables, niveles de servicio, tiempos de procesamiento y frecuencias, tomar el modelo descubierto y animarlo en tiempo real.

Los tres tipos de minería de procesos (Discovery, Conformance y Enhancement) cuenta con diferentes perspectivas las cuales se describen la tabla 2.

Tabla 2. Perspectivas Minería de Procesos.

Nombre	Característica	Objetivo	Resultado	Actores
<b>Perspectiva de control de flujo</b>	Se enfoca en el orden de ejecución de las actividades, puede ser por su ruta en el proceso, por los actores que trabajan en él o por los valores de los correspondientes elementos de datos.	Encontrar una buena caracterización de todos los caminos posibles. Estructurar la organización clasificando a las personas en términos de roles y unidades organizacionales, mostrar la red social.	Se expresa en una red de Petri o en alguna notación de procesos (BPMN, diagramas UML)	Personas Sistemas Departamentos
<b>La perspectiva</b>	Se relaciona con la ocurrencia y frecuencia de los	Descubrir cuellos de botella, medir niveles de servicio,	Predecir el tiempo de procesamiento	Procesos

<b>de tiempo</b>	eventos	monitorear utilización recursos.	la restante de casos ejecución.	de Recursos en
<b>Perspectiva de caso</b>	Se centra en las propiedades de los casos, teniendo en cuenta características como lo son: la ruta en el proceso, los actores que trabajan en él o los valores de los correspondientes elementos de datos.	¿Saber cuál es el rendimiento promedio para los casos de un determinado proceso?	Se debe contar con información detallada de las actividades como puede ser: la edad y el sexo del usuario involucrado.	Proceso Personas Datos Valores de los datos
<b>Perspectiva organizacional</b>	¿Saber quién realiza un trabajo determinado? trabajando sobre la información referente a los usuarios del sistema que ejecutaron cada actividad específico.	El objetivo es estructurar la organización al clasificar a las personas en términos de funciones y roles y a si mismo mostrar la relación entre los distintos usuarios (construir una red social).	Conocer la relación que hay entre las personas de los departamentos.	Personas Departamentos Funciones

Nota: Elaboración propia tomado de Loaiza, 2018.

Una correcta elección de las perspectivas de Minería de Procesos repercute directamente en que se obtenga la información requerida para mejorar o descubrir un mal funcionamiento del sistema. Además, muchas técnicas solo se pueden aplicar en determinadas perspectivas, de ahí la importancia de seleccionar la o las técnicas correctas en un proyecto.

### 1.3. Tipos de proyectos de minería de procesos

La minería de procesos cuenta con tres tipos de proyectos según Santiago (2015) y Loaiza (2018).

- Proyectos orientados a preguntas: en estos casos, se plantean hipótesis o interrogantes que se buscan responder mediante el análisis de minería de procesos, con el objetivo de confirmar o descartar determinadas suposiciones.
- Proyectos basados en objetivos: consisten en mejorar el desempeño en los indicadores de gestión.
- Proyectos basados en datos: son proyectos donde no hay objetivos establecidos, por lo que se pretende descubrir aspectos importantes con respecto a la ejecución real de los procesos, no tienen una fase de definición del proyecto, dado que por su naturaleza exploratoria se parte de unos datos y se debe realizar un análisis de minería de procesos para encontrar patrones y hallazgos que se usan para el diagnóstico del proceso.

#### **1.4.Elementos de los registros de eventos.**

Una de las ventajas de la minería de procesos es que trabaja con datos que se encuentran almacenados en los sistemas de información, para que puedan ser analizados esos datos se debe identificar y tener un panorama general de cada uno de los elementos que están presentes en los registros de eventos como lo es el id de caso, la marca de tiempo, las actividades, los atributos, las instancias y los eventos (figura 3) y tener en cuenta cuales son los requisitos mínimos necesarios para su análisis.

Por lo que en el siguiente apartado se definen los elementos de los registros de eventos con los que trabaja la minería de procesos para su mejor entendimiento (FLUXICON, 2020).

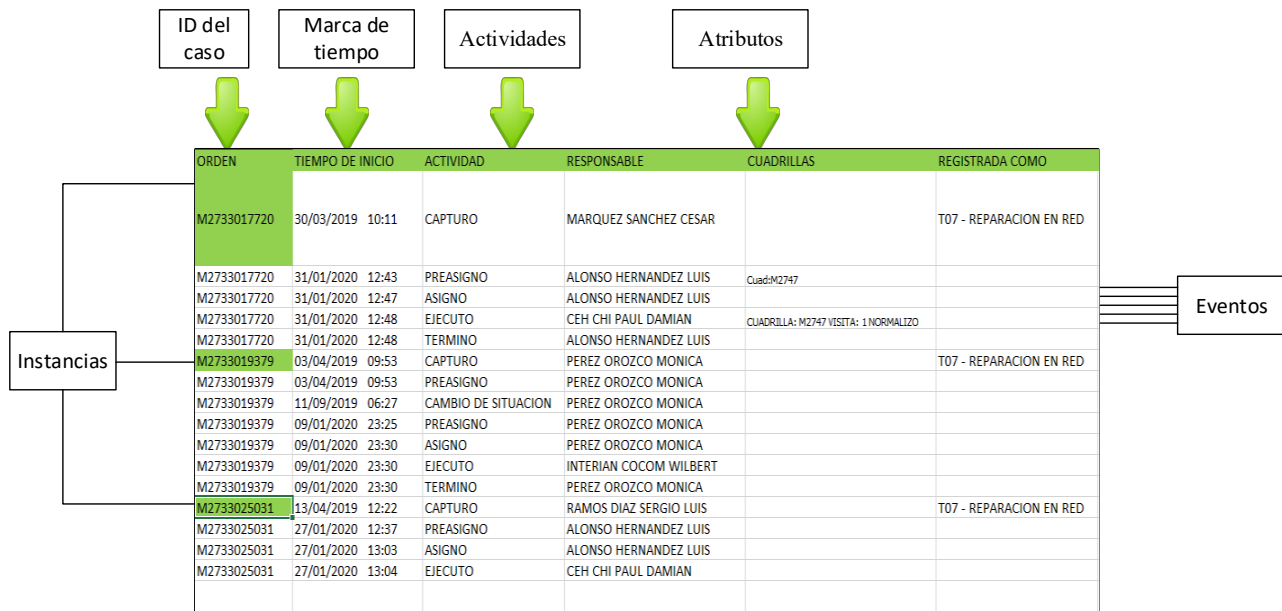


Figura 3. Datos de entrada de minería de procesos.

La base principal de la minería de procesos son los datos históricos del proceso y la observación de los mismos por tanto se deben identificar al menos tres elementos en el *event log*: el primero es el id del caso, el segundo la marca de tiempo y por último la actividad (figura 4) los cuales se definen a continuación.



Figura 4. Elementos mínimos de un registro de eventos.

### 1.4.1. Id de caso

Son las acciones específicas que identifican de manera única la ejecución de un proceso formando el identificador de caso (ej. El proceso de compra, la captura de orden de compra es un caso).

El id de caso influye en el alcance del proceso puesto que indica donde empieza y donde termina el proceso (figura 5).



ORDEN	TIEMPO DE INICIO	ACTIVIDAD
M2733017720	30/03/2019 10:11	CAPTURO
M2733017720	31/01/2020 12:43	PREASIGNO
M2733017720	31/01/2020 12:47	ASIGNO
M2733017720	31/01/2020 12:48	EJECUTO
M2733017720	31/01/2020 12:48	TERMINO

Figura 5. Ejemplo de id de caso.

### 1.4.2. Marca de tiempo

Indica cuando tuvo lugar cada actividad (inicio y fin de la actividad) también llamado tiempo de ejecución, sirve para conocer el comportamiento temporal de cada actividad, saber el orden de las actividades (figura 6); la marca de tiempo es el único requisito que se puede omitir al analizar los datos.



ORDEN	TIEMPO DE INICIO	ACTIVIDAD
M2733017720	30/03/2019 10:11	CAPTURO
M2733017720	31/01/2020 12:43	PREASIGNO
M2733017720	31/01/2020 12:47	ASIGNO
M2733017720	31/01/2020 12:48	EJECUTO
M2733017720	31/01/2020 12:48	TERMINO

Figura 6. Ejemplo marca de tiempo de los registros de eventos.

### 1.4.3. Otros elementos

Pueden existir atributos adicionales que se consideren relevantes en el proceso y deben ser incluidos para mejorar la importancia y valor del análisis (figura 7).

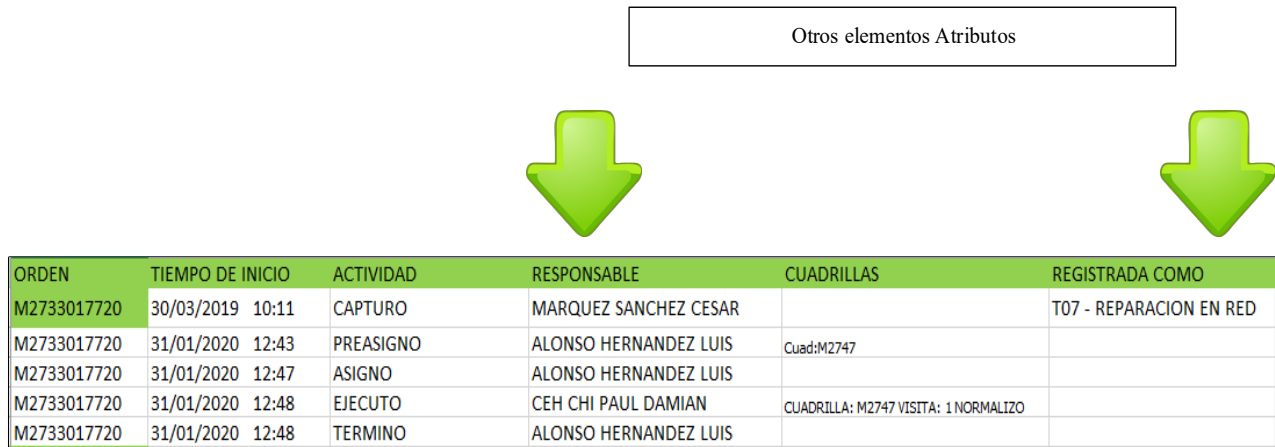


Figura 7. Atributos de un registro de eventos.

### 1.4.4. Evento

Son datos almacenados en el sistema que registran de manera digital las acciones o actividades realizadas, dejando una traza de cada operación efectuada. (figura 8).

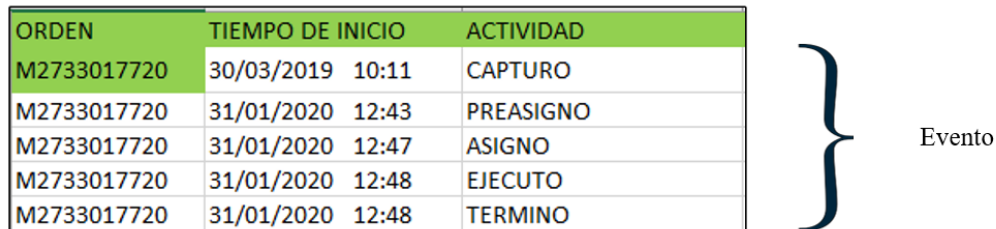


Figura 8. Ejemplo de evento.

### 1.4.5. Caso

Serie de eventos que se realizan consecutivamente, a continuación, se muestran tres ejemplos de casos diferentes (figura 9).

1 caso



ORDEN	TIEMPO DE INICIO	ACTIVIDAD
M2733017720	30/03/2019 10:11	CAPTURO
M2733017720	31/01/2020 12:43	PREASIGNO
M2733017720	31/01/2020 12:47	ASIGNO
M2733017720	31/01/2020 12:48	EJECUTO
M2733017720	31/01/2020 12:48	TERMINO

2 caso



ORDEN	TIEMPO DE INICIO	ACTIVIDAD
M2733019379	03/04/2019 09:53	CAPTURO
M2733019379	03/04/2019 09:53	PREASIGNO
M2733019379	11/09/2019 06:27	CAMBIO DE SITUACION
M2733019379	09/01/2020 23:25	PREASIGNO
M2733019379	09/01/2020 23:30	ASIGNO
M2733019379	09/01/2020 23:30	EJECUTO
M2733019379	09/01/2020 23:30	TERMINO

3 caso



ORDEN	TIEMPO DE INICIO	ACTIVIDAD
M2733053160	29/05/2019 12:22	CAPTURO
M2733053160	29/05/2019 12:22	CAMBIO DE CENTRO
M2733053160	11/02/2020 13:22	PREASIGNO
M2733053160	11/02/2020 13:30	ASIGNO
M2733053160	11/02/2020 13:31	EJECUTO
M2733053160	11/02/2020 13:31	TERMINO

Figura 9. Ejemplo de Casos diferentes.

### 1.4.6. Variantes

Así como se agrupan los eventos en casos, los casos que se parecen también se agrupan para tener una variante (figura 10).

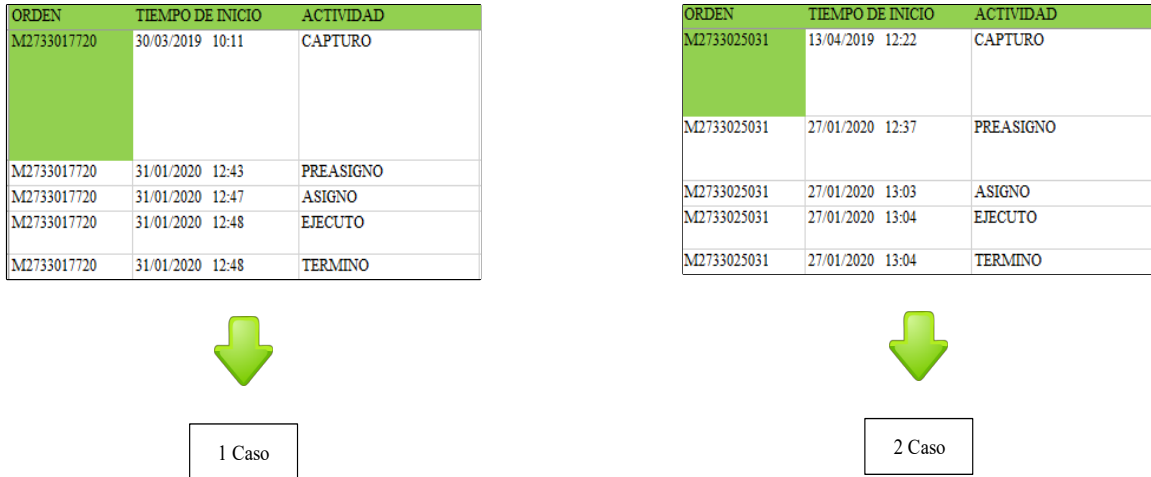


Figura 10. Ejemplo de Variante.

Ahora bien, un proceso consta de dos elementos clave: las actividades y las acciones.

### 1.4.7. Actividad

Una actividad es un paso en el proceso y este puede constar de varios pasos, los pasos se pueden repetir para un solo caso, mientras que todos los pasos no pueden suceder siempre. Para cada paso o cambio de estado que realiza el proceso debe haber diferentes nombres.

Ejemplo: actividad registro de contratación de servicios de corriente eléctrica, los pasos de esa actividad son: “capturo”, “preasigno”, “asigno”, “ejecuto”, “termino”, “cambio de centro”, “cambio de datos”, “reclasifico”; son pasos que pueden o no ser utilizados y a su vez pueden repetirse según sea el caso (figura 11).

Actividades



ORDEN	TIEMPO DE INICIO	ACTIVIDAD
M2733017720	30/03/2019 10:11	CAPTUREO
M2733017720	31/01/2020 12:43	PREASIGNO
M2733017720	31/01/2020 12:47	ASIGNO
M2733017720	31/01/2020 12:48	EJECUTO
M2733017720	31/01/2020 12:48	TERMINO

Figura 11. Ejemplo de actividades de un registro de eventos.

#### 1.4.8. Acción

La acción es cuando se inicia (ej. La actividad “capturo” inicia el 30/03/2019 10:11) y termina una actividad concreta (cuando empieza la otra actividad que es “preasigno”, figura 12).

Acciones



ORDEN	TIEMPO DE INICIO	ACTIVIDAD
M2733017720	30/03/2019 10:11	CAPTUREO
M2733017720	31/01/2020 12:43	PREASIGNO
M2733017720	31/01/2020 12:47	ASIGNO
M2733017720	31/01/2020 12:48	EJECUTO
M2733017720	31/01/2020 12:48	TERMINO

Inicio
Fin

Figura 12. Acción.

#### 1.4.9. Datos necesarios para realizar el análisis

Una vez que se conoce que tipo de datos se necesitan, se debe tomar en cuenta cuantos datos son necesarios para realizar el análisis de minería de procesos, para lo cual existen dos formas para conocer si una muestra es representativa según Fluxicon (2020): el primero es registrar todos los datos de las actividades de un periodo determinado de tiempo ya que brinda una imagen total de lo que está sucediendo en el periodo de tiempo seleccionado (figura 13).

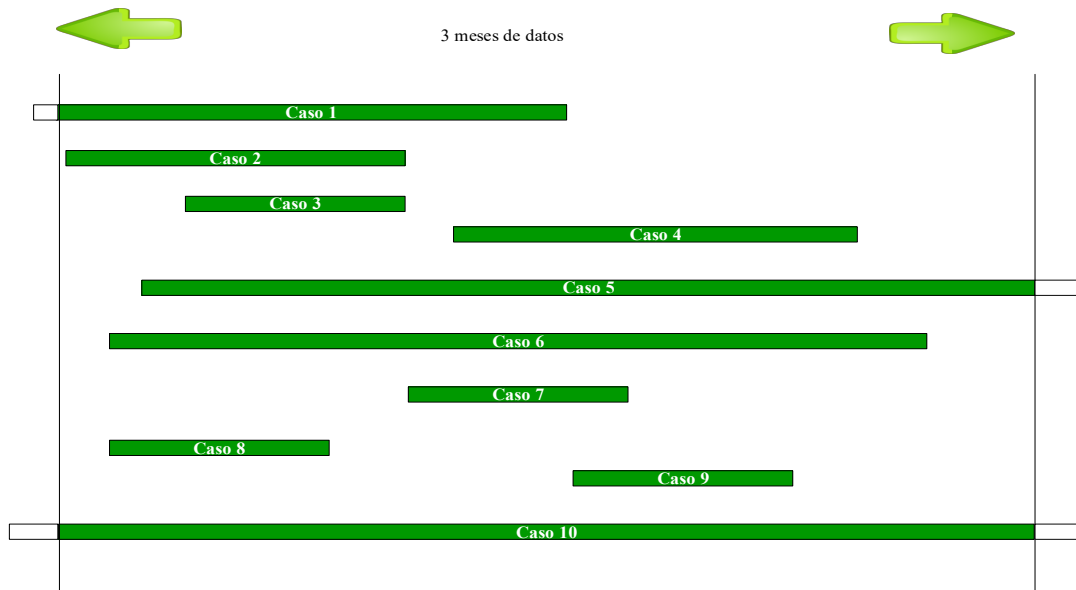


Figura 13. Registros de un tiempo específico.

Sin embargo, se debe contar con registros al menos de tres meses, dependiendo del tiempo de ejecución de una actividad de proceso, puede ser de un año o más se debe tener en cuenta cuanto tiempo tarda en realizar su proceso (ej. Si realiza una actividad en 6 meses no le servirá la información obtenida de 3 meses).

Por tanto, realmente depende del tiempo de ejecución de cada proceso, sin embargo, se puede utilizar la siguiente fórmula para conocer cuántos datos se necesitan extraer:

- Fórmula: periodo de tiempo = tiempo de finalización esperado \*4 \*5.
- La línea de base es el tiempo de finalización del proceso esperado para un caso típico.
- El \*4 asegura que hay tantos datos que se pudiera ver cuatro casos que se iniciaron y se completaron uno después del otro.
- El \*5 representa los casos ocasionales de larga duración (regla 20/80), además asegura que se ven los casos que tardan hasta cinco veces más en la ventana de tiempo extraída.

Ejemplo: si el tiempo de finalización de un caso es de 6 días se tendría el siguiente resultado:

Periodo de tiempo =  $6 * 4 * 5$ .  $6 * 4 = 24$   $24 * 5 = 120$  días.

El segundo es seleccionar todos los casos (figura 14) que comenzaron o terminaron en un periodo de tiempo determinado (ej. Periodo de un mes, tres meses, seis meses, un año, etc.) para extraer todas las actividades que ocurrieron en el caso.

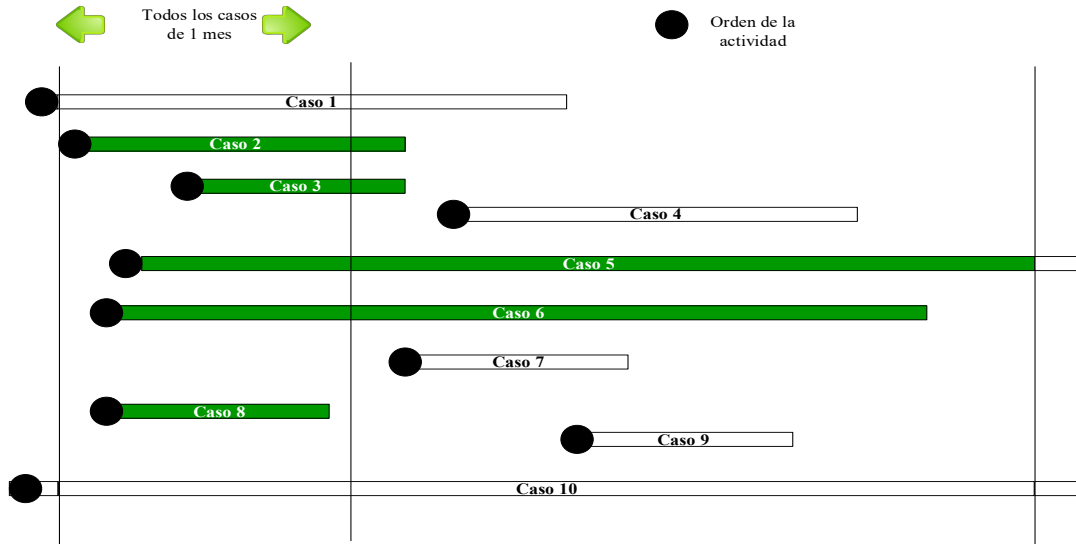


Figura 14. Registro de eventos de un periodo determinado.

### 1.5. Herramientas informáticas para la minería de procesos

Existen diversas herramientas para realizar análisis de minería de procesos, cada una de ellas cuenta con funcionalidades específicas que deben ser consideradas para cada caso de estudio por tal motivo se describen a continuación para una mejor elección: (Arias y Rojas, 2016; CELONIS, 2024; G2, 2024; Grupo de ciencia de procesos y datos Universidad RWTH de Aquisgrán, 2024; Pérez, 2015 y Verbeek, HMW., Buijs, JCAM., Boudewijn, F., Dongen, V. y Aalst, V., 2010).

- **QPR Process Analyzer:** es una aplicación de paga que permite analizar y visualizar los flujos de procesos de negocio de una organización, la arquitectura empresarial y la administración de desempeño, puede integrar MS-Excel con MS-SQL, ofrece la integración de datos a través de conectores de sistemas operativos como SAP, Oracle y ServiceNow, combina inteligencia artificial avanzada, aprendizaje automático y predicciones con el conjunto de gráficos y análisis más completos para la industria.

- **CELONIS:** es un software que cuenta con versión gratuita y de paga, es una tecnología avanzada que tiene la capacidad de ejecución de proceso de datos en tiempo real para tomar acciones inmediatas y decisivas. El sistema entrelaza los procesos interconectados identificando y eliminando ineficiencias corporativas que aumenta la toma de mejores decisiones comerciales y activa las acciones correctas automáticamente; además, permite realizar monitoreo, facilita el análisis en diferentes niveles con funcionalidades de inteligencia de negocios, estadísticas, manejo de indicadores de desempeño y cuenta con funciones de auditoría de procesos para el control de los mismos creando informes, diseño y modelado de procesos para la visualización de resultados y seguimiento de cambios a través de cuadros de mando (dashboard).
- **XESame:** es una aplicación de versión gratuita que está integrada en PROM, admite la extracción de un registro de eventos de diferentes fuentes como: tablas de bases de datos o archivos de texto; convierte los registros de eventos en un formato llamado extensible event stream (EXS), esta herramienta permite a los expertos del dominio especificar cómo se debe extraer el registro de eventos de los sistemas existentes y convertirlo a XES Prom 6, siendo esta una funcionalidad innovadora para la minería de procesos.
- **PROM:** ofrece una arquitectura que incorpora nuevas funcionalidades a los algoritmos ya existentes en forma de plug-in, es una solución informática de uso libre, creada con el fin de disponer de un espacio donde sea posible publicar los principales avances en minería de procesos a través de algoritmos de procesamiento y análisis de datos. Contrario a DISCO en esta herramienta se encuentran cientos de posibilidades de algoritmos que fueron creados para diferentes propósitos de uso para lo cual se debe tener conocimiento profundo de cada uno de ellos, para saber cómo se usan y que resultados se obtendrán en las diferentes notaciones (BPMN, Petri Net, UML, CNet, entre otras.)

- **DISCO:** Es una solución informática de paga para minería de procesos enfocada a la visualización de resultados enfocado en el comportamiento de los eventos de cada proceso on simulación de movimiento en tiempo real. Adicionalmente permite la consulta de un conjunto de estadísticas, que se presentan desde un resumen general hasta el detalle de actividades con la posibilidad de filtrar el conjunto de datos según los campos (criterios) disponibles para la carga y su posterior análisis; además utiliza el algoritmo fuzzy miner que es uno de los más destacados en el ámbito educativo, se especializa en el descubrimiento y análisis de performance de los procesos permitiendo la aplicación de filtros sobre los criterios de análisis y se destaca por la velocidad, procesamiento y generación de resultados; ofrece una interfaz amigable para el usuario, tiene funcionalidades para realizar filtros, obtener estadísticas e información relacionada con los casos y las variantes.

Para este caso de estudio se utilizará la herramienta informática DISCO puesto que a pesar de que es una herramienta de paga cuenta con un entorno amigable y por tanto fácil de usar, también se utilizara PROM que es una herramienta de software libre que cuenta con más complementos de análisis.

# **CAPÍTULO II**

**Algoritmos de**

**la minería de**

**procesos**

## **2. Algoritmos y herramientas de modelado de la minería de procesos**

Las técnicas de minería de procesos utilizan los datos que se encuentran almacenados en las base de datos del sistema de información de una empresa, del cual se eligen los elementos necesarios (el nombre del proceso, fecha de inicio y termino del proceso y otros elementos se requieran analizar) para generar un *event log* y elegir la técnicas de minería de procesos (algoritmo Alpha, minado heurístico, basado en regiones, algoritmo genético o fuzzy miner) más adecuada a aplicar para después generar los modelos de procesos (redes Petri, modelo BPMN, redes causales, etc.) y poder visualizar el análisis ya sea de la vista global del proceso, la frecuencia, los tiempos, las variantes alternativas del proceso o identificación de cuellos de botella (Pérez, Izquierdo, Pupo y Orellana, 2018).

En el sector empresarial y en ocasiones en el académico se aplican diferentes técnicas de minería de procesos (algoritmos), por lo que en los siguientes apartados se describirán y ejemplificaran algunas de ellas para entender cómo se descubren los modelos.

### **2.1. Algoritmos de modelado**

La minería de procesos cuenta con técnicas para descubrir modelos de proceso, lo que implica la implementación de algoritmos que conviertan los registros de eventos en modelos que sean capaces de representar el comportamiento visto en los registros. Para ello se dispone de varios algoritmos que se definen a continuación (Bogarín, 2018; Diez, 2016; González, 2016 y Loaiza, 2018):

#### **2.1.1. Algoritmo básico (Algoritmo “Alpha”)**

El algoritmo Alpha es uno de los primeros algoritmos de descubrimiento de procesos que abordan la concurrencia, este algoritmo trabaja en relación con la dependencia entre los eventos de un registro a fin de que se tenga una representación visual del proceso, sin embargo, presenta problemas para lidiar con el ruido y el comportamiento infrecuente o incompleto (Bogarín, 2018, González, 2016 y Van der Aalst, 2021).

El algoritmo Alpha consta de ocho pasos los cuales se muestran a continuación:

$$1. - T_L = \{ t \in T \mid \exists \sigma \in L \ t \in \sigma \},$$

$$2. - T_l = \{ t \in T \mid \exists \sigma \in L \ t = first(\sigma) \},$$

$$3. - T_o = \{ t \in T \mid \exists \sigma \in L \ t = last(\sigma) \},$$

$$4. - X_L = \{ (A, B) \mid A \subseteq T_L \wedge A \neq 0 \wedge B \subseteq T_L \wedge B \neq 0 \wedge$$

$$\forall a \in A \ \forall b \in B \ a \rightarrow_L b \wedge \forall a_1, a_2 \in A \ a_1 \#_L a_2 \wedge \forall b_1, b_2 \in B \ b_1 \#_L b_2 \},$$

$$5. - Y_L = \{ (A, B) \in X_L \mid \forall (A', B') \in X_L \ A \subseteq A' \wedge B \subseteq B' \Rightarrow (A, B) = (A', B') \},$$

$$6. - P_L = \{ p_{(A, B)} \mid (A, B) \in Y_L \} \cup \{ i_L, o_L \},$$

$$7. - F_L = \{ (a, p_{(A, B)}) \mid (A, B) \in Y_L \wedge a \in A \} \cup \{ (p_{(A, B)}, b) \mid (A, B) \in$$

$$Y_L \wedge b \in B \} \cup \{ (i_L, t) \mid t \in T_l \} \cup \{ (t, o_L) \mid t \in T_o \},$$

$$8. - \alpha(L) = (P_L, T_L, F_L).$$

En el siguiente apartado se describe brevemente la funcionalidad de cada uno de los pasos del algoritmo Alpha.

1. Definir todos los eventos en el registro de eventos.
2. Definir todos los primeros eventos destacados.
3. Definir todos los últimos eventos.
4. Calcular los posibles conjuntos de a y b (todos los eventos dentro de A y B deben ser independientes entre sí y todos los eventos deben ser casualmente relacionados con los eventos en B).
5. Caída de conjunto no máximo, es decir, se tienen que eliminar los conjuntos duplicados.  
se corrigió la redacción
6. Crear lugares para todos los estados de inicio y finalización.
7. Dibujar las conexiones pertinentes y por ultimo
8. Retornar la red Petri final

Para mostrar cómo se aplica el algoritmo Alpha a un conjunto de registro de eventos se utilizará el siguiente ejemplo (Coursera, 2020 y Van der Aalst, 2016).

Se tiene el siguiente registro:

$$L1 = [(a, b, c, d, e, f, b, d, c, e, g), \\ (a, b, d, c, e, g)^2, \\ (a, b, c, d, e, f, b, c, d, e, f, b, d, c, e, g)]$$

Donde L es igual al *event log*, cada una de las letras del ejemplo son las actividades o eventos, ahora bien, todas las letras contenidas dentro de un paréntesis indican una secuencia de actividades y el subíndice revela las veces que se ejecuta una secuencia de actividades, en el ejemplo anterior hay un subíndice dos (a, b, d, c, e, g)<sup>2</sup> el cual indica que esa secuencia se ejecutó dos veces.

Pasos:

1.- Definir todos los eventos.

$$TL = \{a, b, c, d, e, f, g\}$$

2.- Definir los eventos iniciales.

$$Tl = \{a\}$$

3.- Definir los eventos finales.

$$To = \{g\}$$

4.- Para calcular los posibles conjuntos de A y B es necesario utilizar la matriz de huella (Footprint) la cual ayuda a saber las relaciones que hay entre las actividades o eventos, para lo cual hay que tener en cuenta las siguientes relaciones:

- **Relación sucesión directa (>):**

En este caso se puede dar la relación donde la sucesión es directa de un evento con el siguiente como en el ejemplo: a es seguida de b, b es seguido de c, c por d y así sucesivamente.

$$[(a, b, c, d), (e, d, a, b), (a, e)]$$

A > B

A > E

B > C

C > D

D > A

E > D

- **Relación causalidad directa ( $\rightarrow$ ):**

Es representada con una flecha hacia la derecha ( $\rightarrow$ ) donde un evento (x) puede ser seguido por un evento (y), pero (y) nunca es seguido por (x), por ejemplo, a es seguido por b, pero b no puede ser seguido por a.

$\rightarrow$                        $\rightarrow$

[(a, b, c, d), (e, d, a, b), (a, e)]

- **Relación paralelismo ( $||$ ):**

Se da cuando la relación directa de la sucesión se mantiene en ambas direcciones es decir al leer la sucesión de izquierda a derecha y viceversa, por ejemplo, cuando a es seguida de b y b es seguida de a se dice que son paralelos.

$\rightarrow$                        $\leftarrow$

[(a, b, c, d), (e, d, b, a), (a, e)]

A || B

- **Relación independencia o relación de elección (#):**

Esta relación se da cuando un evento es independiente y/o no tiene relación con algún evento. Para el siguiente ejemplo, se puede decir que A tiene sucesión directa con B, por lo tanto, no hay relación y es independiente.

[(a, d, c, b), (a, c, d, b), (a, e)]

(a # b)

Estas relaciones se utilizan para aprender patrones en el proceso, en base a los cuatro casos anteriormente descritos se puede llenar la matriz de huella (tabla 3) y observar las relaciones que hay entre los eventos para conocer los posibles conjuntos.

Tabla 3. Matriz de huella (Footprint).

	a	b	C	d	e	f	g
a	#	→	#	#	#	#	#
b	←	#	→	→	#	←	#
c	#	←	#		→	#	#
d	#	←		#	→	#	#
e	#	#	←	←	#	→	→
f	#	→	#	#	←	#	#
g	#	#	#	#	←	#	#

- Primero se deben anotar todos los eventos que tengan una relación o sucesión directa con causalidad y son todos aquellos eventos que se muestran con el siguiente símbolo →.

XL= {(a, b) (b, c) (b, d) (c, e) (d, e) (e, f) (e, g) (f, b)}

- En segundo lugar, se deben escribir los conjuntos que tengan relación pero que a su vez sean independientes.

((a, f) (b) ((e) (f, g)

5.- Se deben eliminar los conjuntos duplicados para dejar los máximos, se tiene que checar cuales son los conjuntos que se repiten.

XL= (a, b) (b, c) (b, d) (c, e) (d, e) (e, f) (e, g) (f, b) ((a, f) (b) ((e) (f, g)

Dando como resultado los siguientes conjuntos:

$$YL = \{(b, c) (b, d) (c, e) (d, e) ((a, f) (b)), ((e) (f, g))\}$$

6.- En el siguiente paso se deben crear lugares (P) para cada uno de los conjuntos que quedaron.

$$PL = \{P((b), (c)),$$

$$PL =$$

$$P((b), (d)),$$

$$PL = P((c), (e)),$$

$$PL = P((d), (e)),$$

$$PL = P(((a), (f)), (b)),$$

$$PL = P((e), ((f), (g)))\}$$

Los arcos sirven para ver todas las conexiones del lugar inicial hasta el final; para obtener los arcos se tienen que sacar cada una de las actividades que se encuentran en los conjuntos resultantes de los lugares.

7.- En base a los resultados antes obtenidos se crearán los arcos (F) para cada uno de los conjuntos, se saca el primer evento o actividad seguido del conjunto como estaba originalmente con los lugares, después se coloca el conjunto, pero con la diferencia de que al final se saca el último evento y así sucesivamente con cada una de las actividades de los demás conjuntos quedando de la siguiente manera.

$$PL = \{P((b), (c)),$$

$$FL = (b, P((b), (c))), (P((b), (c)), c),$$

$$PL = P((b), (d)),$$

$$FL = (b, P((b), (d))), (P((b), (d)), d),$$

$$PL = P((c), (e)),$$

FL= (c, P ((c), (e)), (P (c), (e), e),

PL= P((d), (e)),

FL= (d, P ((d), (e)), P (d), (e),

PL= P(((a), (f)), (b)),

FL= (a, P ((a), (f)), (b))), (f, P((a), (f)), (b)), (P ((a), (f)), (b), b)),

PL= P((e), ((f), (g))).

FL= (e, P ((e), ((f), (g))), (P (e), ((f), (g)), f), (P (e), ((f), (g)), g).

Al juntar todos los arcos resultantes quedaran de la siguiente manera:

FL= {(b, P ((b), (c))), (P ((b), (c)), c), (b, P ((b), (d)), (P (b), (d), d), (c, P ((c), (e)), (P (c), (d, P ((d), (e)), P (d), (e), (a, P ((a), (f)), (b))), (f, P((a), (f)), (b)), (P ((a), (f)), (b), b)), (e, P ((e), ((f), (g))), (P (e), ((f), (g)), f), (P (e), ((f), (g)), g)}

Se agrega iL al inicio (iL, a) y oL al fin (g, oL).

Finalmente se obtiene la red Petri con lugares, transiciones y arcos.

$$\alpha(L) = (P_L, T_L, F_L)$$

Donde  $P_L$  son lugares (places),  $T_L$  trazas o transiciones y por último  $F_L$  son los arcos.

**8.-**Para terminar, se dibujan las conexiones y se obtiene la red Petri.

Cuando se habla de huellas se habla de ciertos patrones que se reflejan en el registro como en el modelo y estos son el punto de partida para el algoritmo Alpha; el registro de eventos y el modelo de procesos tienen exactamente la misma huella (figura 15).

L1= [(a, b, c, d, e, f, b, d, c, e, g),

(a, b, d, c, e, g)2,

(a, b, c, d, e, f, b, c, d, e, f, b, d, c, e, g)]

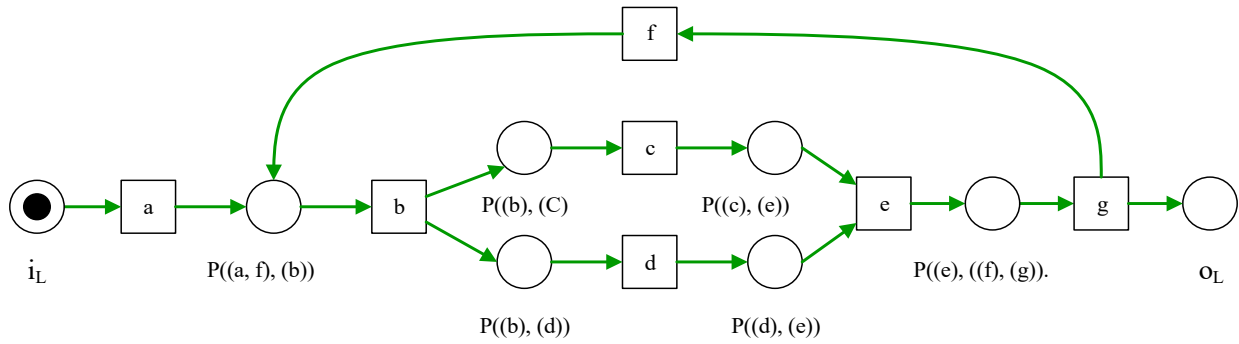


Figura 15. Resultado: red Petri.

Con la intención de tener un mejor entendimiento del algoritmo Alpha se desarrollan a continuación dos ejercicios.

**Ejercicio 1.**

$L2 = [(a, c, d)45, (b, c, d)42, (a, c, e)38, (b, c, e)22]$

1.-  $TL = \{a, b, c, d, e\}$

2.-  $Tl = \{a, b\}$

3.-  $To = \{d, e\}$

4.- Matriz de huella (tabla 4)

Tabla 4. Matriz de huella Ejercicio 1.

	a	b	c	d	e
a	#	#	→	#	#
b	#	#	→	#	#
c	←	←	#	→	→
d	#	#	←	#	#
e	#	#	←	#	#

$XL = \{(a, c), (b, c), (c, d), (c, e), (c, (d, e)), (d, (a, b))\}$

5.-  $YL = \{(a, c), (b, c), (c, (d, e)), (d, (a, b))\}$

6.-  $PL = \{P((a), (c))\}$

$PL = P((b), (c)),$

$PL = P(c, (d), (e)),$

$PL = P(c, ((a), (b)))\}$

7.-  $FL = \{(a (P((a), (c))), (P(a), (c), c)),$

$(b, P((b), (c))), (P((b), (c), c)),$

$(c, (P(c), (d), (e))), (P((c), (d), (e)), d), (P((c), (d), (e)), e),$

$(c, P(d, ((a), (b))), (P(c, ((a), (b), a)), (P(c, ((a), (b), b))))\},$

$(iL, (a, b)), ((d, e), oL).$

$\alpha(L) = (PL, TL, FL)$

8.- Conexión de la red Petri (figura 16)

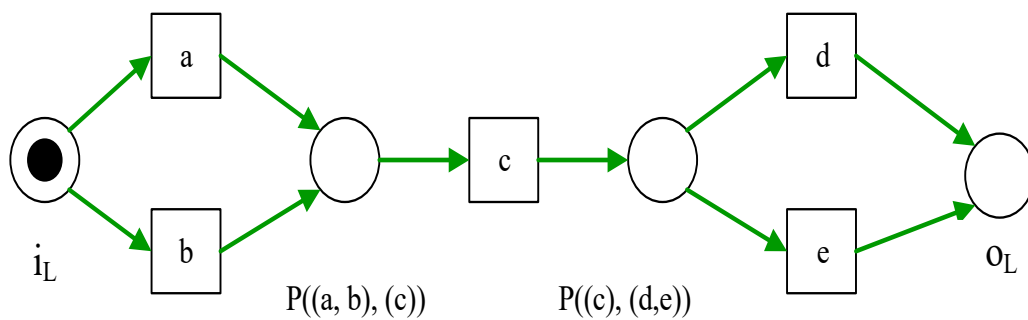


Figura 16. Red Petri ejercicio 1.

**Ejercicio 2.**

$L3 = [(a, b, e, f)^2, (a, b, e, c, d, b, f)^3, (a, b, c, e, d, b, f)^2, (a, b, c, d, e, b, f)^4, (a, e, b, c, d, b, f)^3]$

1.-  $TL = \{a, b, c, d, e, f\}$

2.-  $Tl = \{a\}$

3.-  $TO = \{f\}$

4.- Matriz de huella (tabla 5)

Tabla 5. Matriz de huella ejercicio 2.

	a	b	c	d	e	f
a	#	→	#	#	→	#
b	←	#	→	←		→
c	#	←	#	→		#
d	#	→	←	#		#
e	←				#	→
F	#	←	#	#	←	#

$XL = \{(a, b), (a, e), (b, c), (b, f), (c, d), (d, b), (e, f), ((a, d), (b)), ((b), (c, f))\}$

5.-  $YL = \{(a, e), (c, d), (e, f), ((a, d), b), ((b), (c, f))\}$

6.-  $PL = \{P(a, e), P(c, d), P(e, f), P((a, d), b), P((b), (c, f))\}$

7.-  $FL = \{(a, P(a, e)), P((a, e), e), (c, P(c, d)), P((c, d), d), (e, P(e, f)), P((e, f), f), (a, P(a, d), b), (d, P(a, d), b), P((a, d), b), (b, P(b), (c, f)), P((b), (c, f), c), P((b), (c, f), f)\}$

$(iL, a), (f, oL) \quad \alpha(L) = (PL, TL, FL)$

### 8.- Conexión de la red Petri (figura 17)

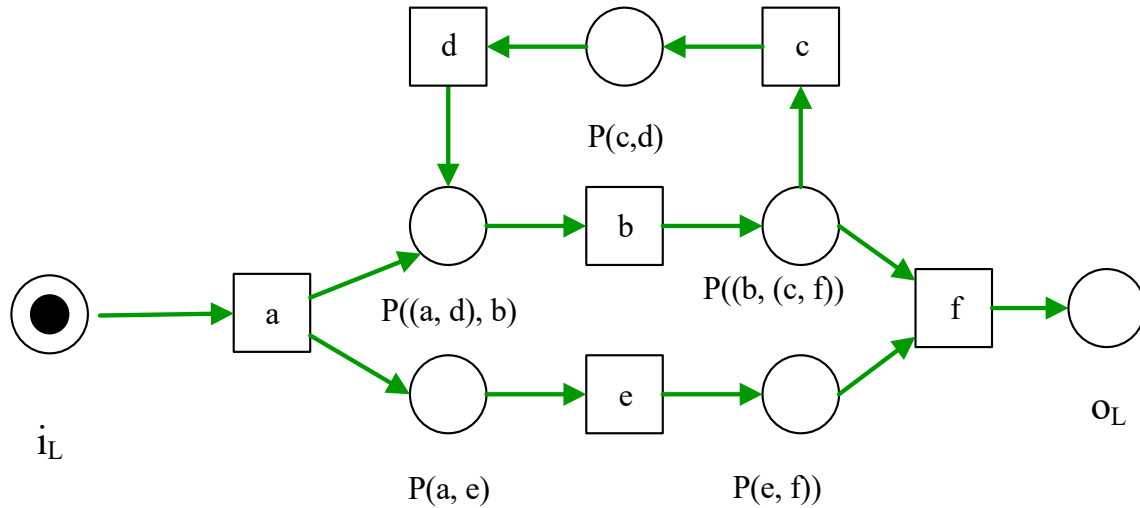


Figura 17. Red Petri ejercicio 2.

#### 2.1.2. Limitaciones del algoritmo Alpha

Para conocer cuando existe alguna limitación en el algoritmo primero se debe tener un registro de eventos y conocer las secuencias de las actividades para cada caso, aplicar el algoritmo Alpha para obtener una red Petri la cual es típicamente una red de flujo de trabajo con un inicio y un fin y analizar su estructura. ¿En qué situaciones el algoritmo Alpha no produce el resultado esperado?, a continuación, se presentan algunos casos donde se presentan limitantes del algoritmo Alpha.

- Lugares implícitos:

Ocurre cuando se aplica el algoritmo Alpha a un registro y se obtiene un modelo que contiene lugares duplicados ( $P1$  y  $P2$ ) o que no tienen algún propósito, puesto que si se eliminan no alteran el comportamiento del modelo debido a que las actividades no están influenciadas por esos lugares y lo único que provocan es la complejidad del modelo y aportan algo (figura 18).

$L4 = [(a, c, e, g)2, (a, e, c, g)3, (b, d, f, g)2, (b, f, d, g)4]$

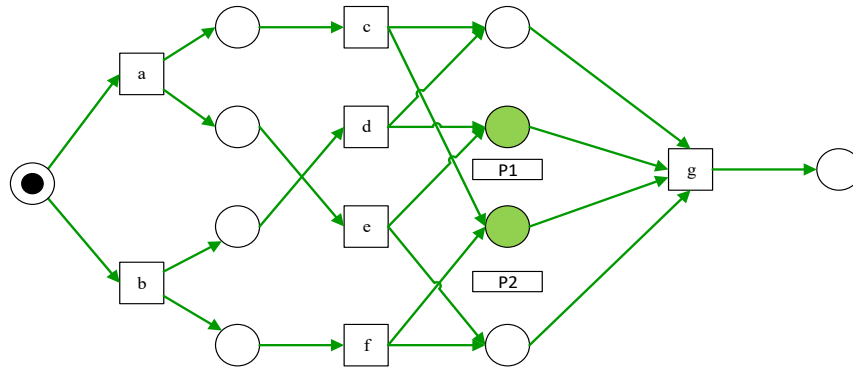


Figura 18. Lugares implícitos.

- **No se puede hacer frente a los bucles:**

Esto sucede cuando se tiene un registro que repite alguna actividad que lleva por nombre bucle de longitud uno (figura 19) o en su caso bucle de longitud dos (figura 20) el cual indica que no existe sucesión o relación directa con las demás actividades para poder formar la red Petri. En los dos ejemplos siguientes (figura 19 y 20) existe paralelismo, en el primer ejemplo se da con la actividad b y en el segundo con las actividades bc, en ambos existe un error, las actividades no tiene relación con las demás y por tal motivo el modelo no se forma de manera adecuada.

Ejemplo 1.

$L5 = [(a, c)2, (a, b, c)3, (a, b, b, c)2, (a, b, b, b, b, c)1]$

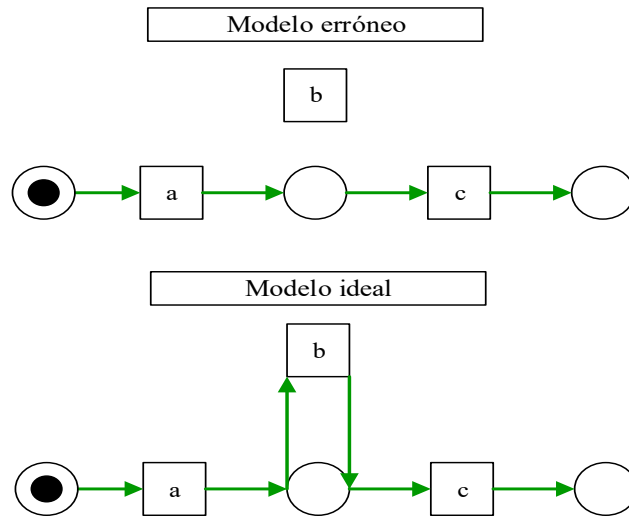


Figura 19. Bucle longitud uno.

$L6 = [(a, b, d)3, (a, b, c, d)2, (a, b, c, b, c, b, d)]$

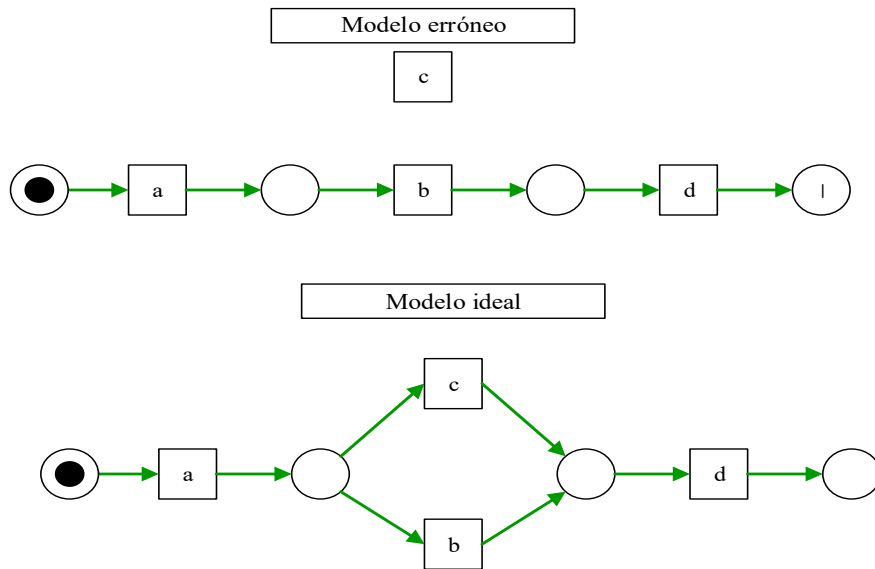


Figura 20. Bucle longitud dos.

- Dependencias no locales:

Se refiere a la dependencia que pueda existir entre dos actividades, es decir, si no se descubren lugares (figura 21) que relacionen las actividades nunca habrá sucesión directa de una actividad a otra, en el ejemplo siguiente no existen lugares que unan a la actividad a y b ni en las actividades b y e.

$L7 = [(a, c, d)45, (b, c, e)42]$

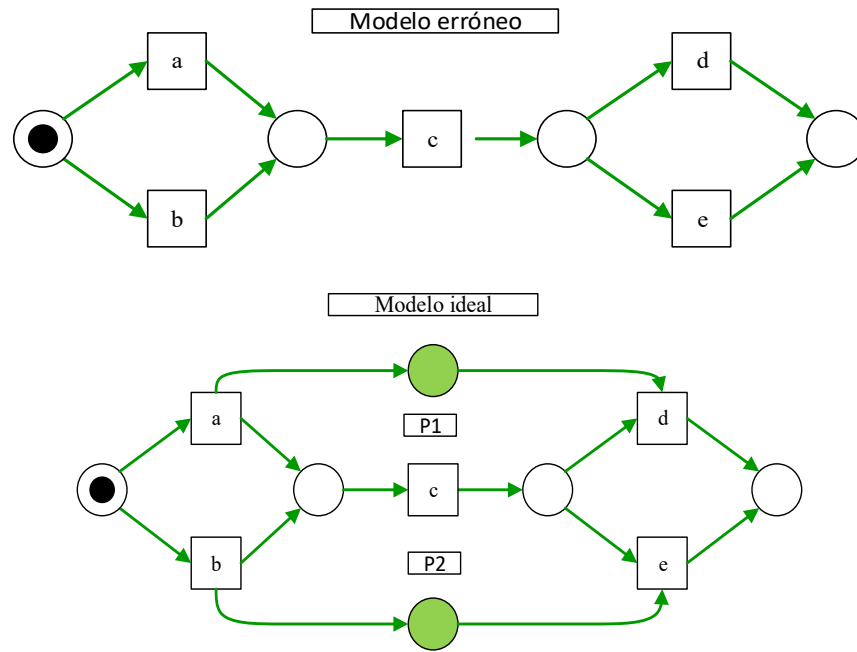


Figura 21. Dependencias no locales.

- Redes Petri que tienen una etiqueta de actividad única como en el siguiente caso (figura 22), solo existe la actividad a.

$L8 = [(a, a)55]$

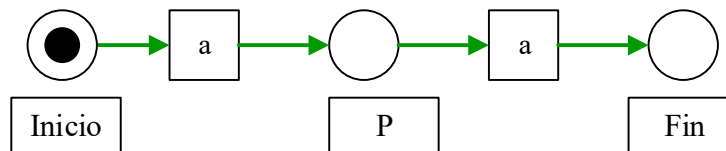


Figura 22. Red Petri con etiqueta de actividad única.

- Modelo en flor

Este tipo de modelos permiten cualquier comportamiento que involucre una actividad por lo que se considera insuficiente precisión, nula o con poca integridad, la cual muestra que no tiene un orden específico (figura 23).

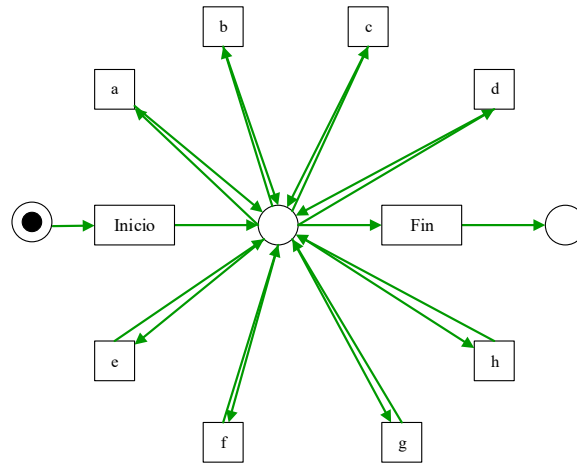


Figura 23. Modelo en flor.

El ruido es un problema que se presenta en los registros de eventos provocando comportamientos raros y registros poco frecuentes que no se requieren en el modelo lo cual complica la integridad del mismo, debido a estos inconvenientes han surgido los siguientes modelos refinados.

## 2.2. Algoritmos avanzados de descubrimiento

### 2.2.1. Algoritmo Alpha +

A partir del algoritmo Alpha surge el algoritmo Alpha+ el cual incorpora el tratamiento de ciclos cortos y constructores sin libre elección; este algoritmo supone una mejora sobre el algoritmo Alpha, puesto que es capaz de modelar bucles de cualquier longitud introduciendo dos nuevos casos: el primero “ $a\Delta b$ ”, se da cuando hay una secuencia “aba” y el segundo caso es “ $a\langle b$ ” se utiliza cuando no se encuentra una traza que contiene “aba” y “bab” de esta forma se asegura la modelación de bucles.

Esta mejora en el algoritmo se realiza bajo dos premisas: la primera es que el registro debe estar completo, es decir, debe contar con todas las trazas posibles del proceso y todas sus transacciones; la segunda premisa el modelo obtenido suele ser muy general y tiende a contemplar comportamiento adicional que no es mostrado en el registro.

### 2.2.2. Algoritmo minado heurístico

Este algoritmo usa una representación similar a las redes causales, toma las frecuencias de eventos y las secuencias para construir un modelo de procesos; el sesgo de representación

proporcionado por la red causal como el uso de frecuencias hacen que el enfoque sea mucho más robusto, la idea básica del algoritmo minado heurístico es que los patrones de frecuencia baja no deberían ser incorporados dentro del modelo (Coursera, 2021).

El gráfico de dependencia se usa para visualizar las frecuencias (sobre todo las frecuencias bajas) de las actividades; la medida de dependencia para descubrir la causalidad (relación) y las frecuencias ayudan a conocer la sucesión directa de un registro de eventos; es así que el minado heurístico emplea una fórmula (figura 24) que indica el número de veces que una actividad fue seguida por otra dentro de un registro, resume todos los rastros, además, cada actividad puede tener enlaces de entrada y enlaces de salida.

$$|a >_L b| = \sum L(\sigma) x |\{1 \leq i < |\sigma| \mid \sigma(i) = a \wedge \sigma(i+1) = b\}| \sigma \in L$$

Figura 24. Sucesión directa.

Para descubrir la causalidad se utiliza la fórmula de la medida de dependencia la cual se describe a continuación.

$|a \Rightarrow L^b|$  es el valor de relación de dependencia entre a y b:

$$|a \Rightarrow L^b| = \left( \frac{|a > L^b| - |b > L^a|}{|a > L^b| + |b > L^a| + 1} \right) \text{ si } a \neq b$$

Dónde:

L es un registro de eventos sobre "A" y "a" y "b" pertenecen a "A". (a,b ∈ A).

$|a > L^b|$  Es el número de veces que a es seguida directamente de b.

$|a \Rightarrow L^b|$  Produce un valor entre -1 y 1.

Al tener un modelo que tiene concurrencia accidentalmente las actividades se seguirán unas a otras, pero no se deben confundir con las causalidades, por lo cual al observar la fórmula anterior se tiene que, si "a" y "b" son diferentes se debe contar cuantas veces "a" es seguida por "b" menos el número de veces que "b" es seguido por "a" y se divide por la suma de esos dos números más uno, de esa manera siempre se obtiene un resultado entre cero y uno, si el número

está cerca de uno hay una causalidad muy fuerte, pero si el número es negativo o está cerca del cero hay una baja causalidad en esa dirección.

Así mismo, la fórmula trabaja con un caso especial en el que “a” es igual a “b”, al aplicar la formula donde a=b siempre se obtendría cero (figura 25), es decir, se cuenta cuantas veces “a” es seguido por sí mismo y dividido por el mismo número más uno, de nueva cuenta el resultado se encontrara en valores de cero y uno, cuanto mayor sea el número más probable es que haya dependencia o causalidad infinita.

$$|a \Rightarrow L^b| = \left( \frac{|a > L^a|}{|a > L^a| + 1} \right) \text{ si } a = b$$

Figura 25. Fórmula medida de dependencia.

Donde:

$|a \Rightarrow L^a|$  “a” es seguido por sí misma “a” indica que es un bucle con una fuerte dependencia reflexiva.

A continuación, se muestran dos ejemplos de minado heurístico: uno de sucesión directa y otro de medida de dependencia; en el primer ejemplo se tiene un gráfico de dependencia (figura 26) con dos números: el número 45 (rojo) indica cuantas veces “a” fue seguido por “b” en algún lugar del registro, el valor 0.98 (azul) muestra la medida de dependencia, el cual indica que hay una fuerte relación (entre más cerca este del valor 1 existe mayor relación).

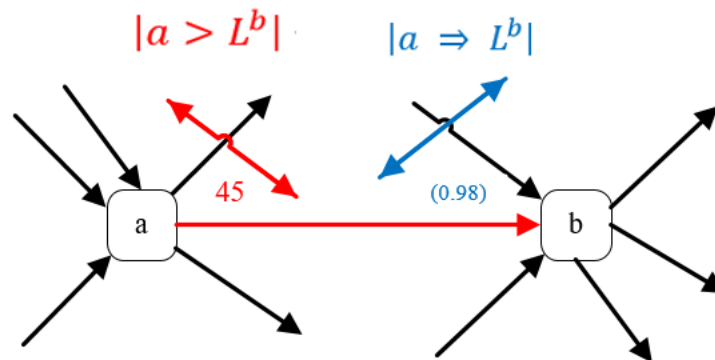


Figura 26. Gráfico de dependencia.

Las flechas representan los arcos y sirven para indicar si algún arco debe ser incluido o no en el modelo, esto dependerá de los valores que estén por encima de un cierto umbral el cual es establecido, por tanto, se eliminarán los arcos que no cumplan con este requisito.

En el segundo ejemplo existe una sucesión directa donde “a” es seguido por “b” diez veces, al aplicar la formula (figura 25) se tiene lo siguiente (figura 27):

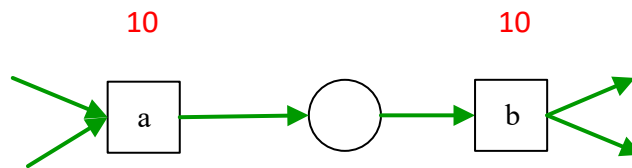


Figura 27. Patrón de secuencia.

$$|a > L^b| = 10 \qquad |b > L^a| = 0$$

$$|a \Rightarrow L^b| = 10/11 \qquad |b \Rightarrow L^a| = -10/11$$

El número rojo (10) indica que existe una frecuencia de diez, al aplicar la formula se le resta cero, se divide entre diez y se le suma uno al cero dando como resultado once, la medida de dependencia entre las actividades “a” y “b” es de 10/11.

Al trabajar con la dirección opuesta “b” nunca es seguido por “a”, por tanto, se tiene una frecuencia de cero, al calcular la medida de dependencia entre “b” y “a” que es la dirección inversa se obtiene un número negativo, es decir, el resultado es -10/11.

Para ambos ejemplos (figura 26 y 27) se alcanzan los dos umbrales (frecuencia y dependencia grafica), al incluir la causalidad se observa que el arco cumple con los requisitos de los umbrales la cual indica que “a” debe ser seguido directamente por “b” al menos una vez (a es seguido por b diez veces) y para la medida de dependencia debe ser al menos la mitad lo cual se cumple porque el resultado obtenido es de 0.9.

A continuación, se muestra un ejemplo aplicando el minado heurístico el cual consta de tres pasos: el primero es establecer los umbrales y contar las sucesiones directas; el segundo paso es

calcular las medidas de dependencia utilizando las fórmulas (figura 24 y 25) según sea el caso y por último dibujar el grafico de dependencia incluyendo los arcos que cumplen ambos umbrales.

Se tiene el siguiente registro:

$L1 = [(a, e)5 (a, b, c, e)10(a, c, b, e)10, (a, b, e)1, (a, c, e)1, (a, d, e)10, (a, d, d, e)2, (a, d, d, d, e)1]$

Primero se llena la tabla de frecuencia de actividades con respecto al registro (figura 28).

$ > L $	a	b	c	d	e
a	0	11	11	13	5
b	0	0	10	0	11
c	0	10	0	0	11
d	0	0	0	4	13
e	0	0	0	0	0

Figura 28. Frecuencia de actividades.

Segundo se aplica la fórmula antes mencionada (figura 25) a los resultados que se obtuvieron en las frecuencias de actividades (figura 28) para obtener las medidas de dependencia (figura 29).

$ \Rightarrow L  a$	a	b	c	d	e
a	$\frac{0}{0+1} = 0$	$\frac{11-0}{11+0+1} = .92$	$\frac{11-0}{11+0+1} = .92$	$\frac{13-0}{0+13+1} = .93$	$\frac{5-0}{5+0+1} = .83$
b	$\frac{0-11}{0+11+1} = -.92$	$\frac{0}{0+0+1} = 0$	$\frac{10-10}{10+10+1} = 0$	$\frac{0-0}{0+0+1} = 0$	$\frac{11-0}{11+0+1} = .92$

c	$\frac{0 - 11}{0 + 11 + 1} = -.92$	$\frac{10 - 10}{10 + 10 + 1} = 0$	$\frac{0}{0 + 0 + 1} = 0$	$\frac{0 - 0}{0 + 0 + 1} = 0$	$\frac{11 - 0}{11 + 0 + 1} = .92$
d	$\frac{0 - 13}{0 + 13 + 1} = -.93$	$\frac{0 - 0}{0 + 0 + 1} = 0$	$\frac{0 - 0}{0 + 0 + 1} = 0$	$\frac{4}{4 + 0 + 1} = .80$	$\frac{13 - 0}{0 + 13 + 1} = .93$
e	$\frac{0 - 5}{0 + 5 + 1} = -.83$	$\frac{0 - 11}{0 + 11 + 1} = -.92$	$\frac{0 - 11}{0 + 11 + 1} = -.92$	$\frac{0 - 13}{0 + 13 + 1} = -.93$	$\frac{0}{0 + 0 + 1} = 0$

Figura 29. Medidas de dependencia.

Tercero en base a la tabla de frecuencia de actividades (figura 28) y la tabla de medida de dependencia (figura 29) se deriva la llamada dependencia gráfica (figura 30) en esta se muestran los arcos que cumplen con los umbrales establecidos; un arco entre “x” y “y” solo será incluido si es mayor a los valores, para este ejemplo la frecuencia se usa un umbral de 2 y para la medida de dependencia 0.7.

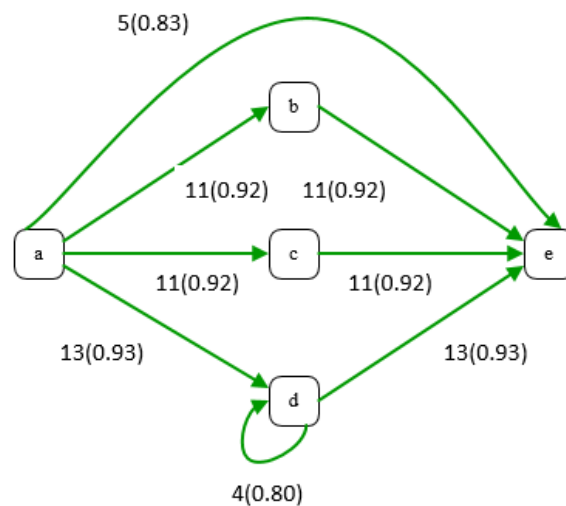


Figura 30. Grafica de dependencia usando umbrales de 2 y 0.7.

Al aumentar las medidas de los umbrales (5 para frecuencia y .9 medida de dependencia) se puede observar que dos arcos fueron eliminados (4(0.80) y 5(0.83)) del grafico anterior (figura 30) debido a que se encuentran por debajo del valor de los umbrales establecidos quedando el grafico de la siguiente manera (figura 31).

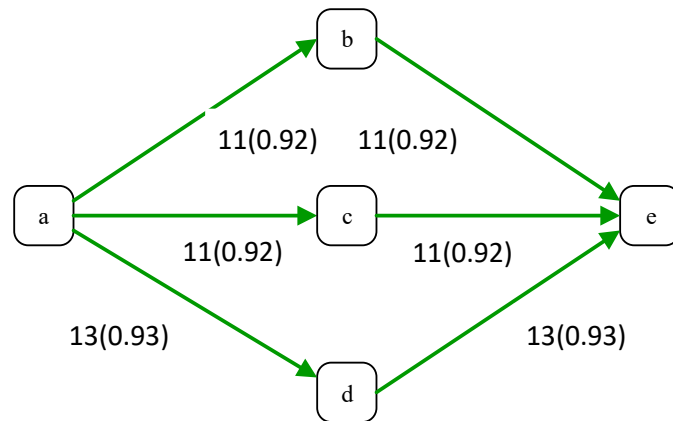


Figura 31. Gráfico de dependencia con umbrales de 5 y .9.

Los gráficos de dependencia son modelos que se obtienen de analizar un registro de eventos, estos modelos se van generando al ajustar o modificar los umbrales y dependerán del grado de concurrencia (dependencia entre las actividades) y la frecuencia con la que las actividades se involucran entre sí.

Para conocer si un rastro es ruido o simplemente es un patrón poco frecuente y saber qué actividades pueden ser aceptadas en el modelo e indicar la relación de dependencia que existe entre ellas existen se toman en cuenta tres parámetros para los umbrales.

- Umbral de dependencia
- Umbral de observaciones positivas
- Umbral relativo al mejor.

Con estos umbrales se puede indicar si se aceptan relaciones de dependencia: (i) señalan que una medida de dependencia debe estar por encima del valor del umbral, (ii) que la frecuencia debe ser mayor que el valor positivo del umbral de observaciones y (iii) la diferencia con la "mejor" medida de dependencia sea menor que el valor relativo al mejor umbral.

El minado heurístico maneja los siguientes parámetros predeterminados:



El registro cuenta con 1000 trazos generados aleatoriamente y el 5% de esos trazos contiene ruido; las actividades D1, D2 Y D3 no son presentadas en el registro (son tareas invisibles) (figura 32) y el ruido que existe genera 5 diferentes tipos de operación ya sea para:

- Eliminar la cabeza del trazo
- Eliminar la cola del trazo
- Eliminar una parte del cuerpo
- Eliminar un evento o
- Intercambiar dos eventos escogidos aleatoriamente

Durante la eliminación de operaciones al menos un evento y no más de un tercio de la traza se elimina, para incorporar el 5% de ruido en las trazas sin ruido se seleccionan aleatoriamente y luego se aplica una de las cinco operaciones descritas anteriormente (cada generación de operación tiene una probabilidad de 1/5).

#### **2.2.2.1. Minado heurístico con parámetros por default**

Al usar los parámetros predeterminados del minado heurístico: relativa al mejor umbral = 0.05, observaciones positivas = 3, umbral de dependencia = 0.9 y AND- umbral = 0,1 se observa que las dos relaciones de dependencia con frecuencia baja de D a K y de J a C no están en el modelo (figura 33).

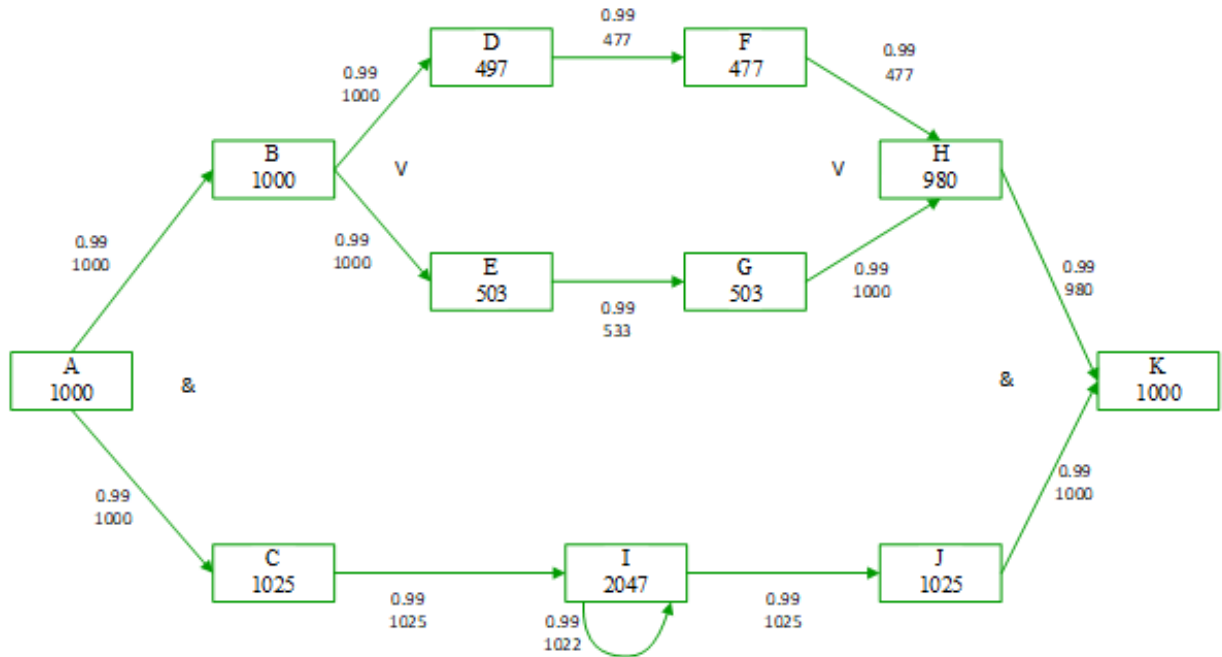


Figura 33. Modelo con 1000 registros y con parámetros de umbral predeterminados.

Al cambiar la configuración de los umbrales con los siguientes parámetros El resultado del minado heurístico: umbral relativo al mejor =0.20, umbral de dependencia= 0.85, se tiene el siguiente modelo (figura 34) sin ruido y tomando en cuenta que la relación de dependencia con frecuencias bajas son obtenidos con esta configuración, pero no estaba cuando se utilizaron los parámetros predeterminados.

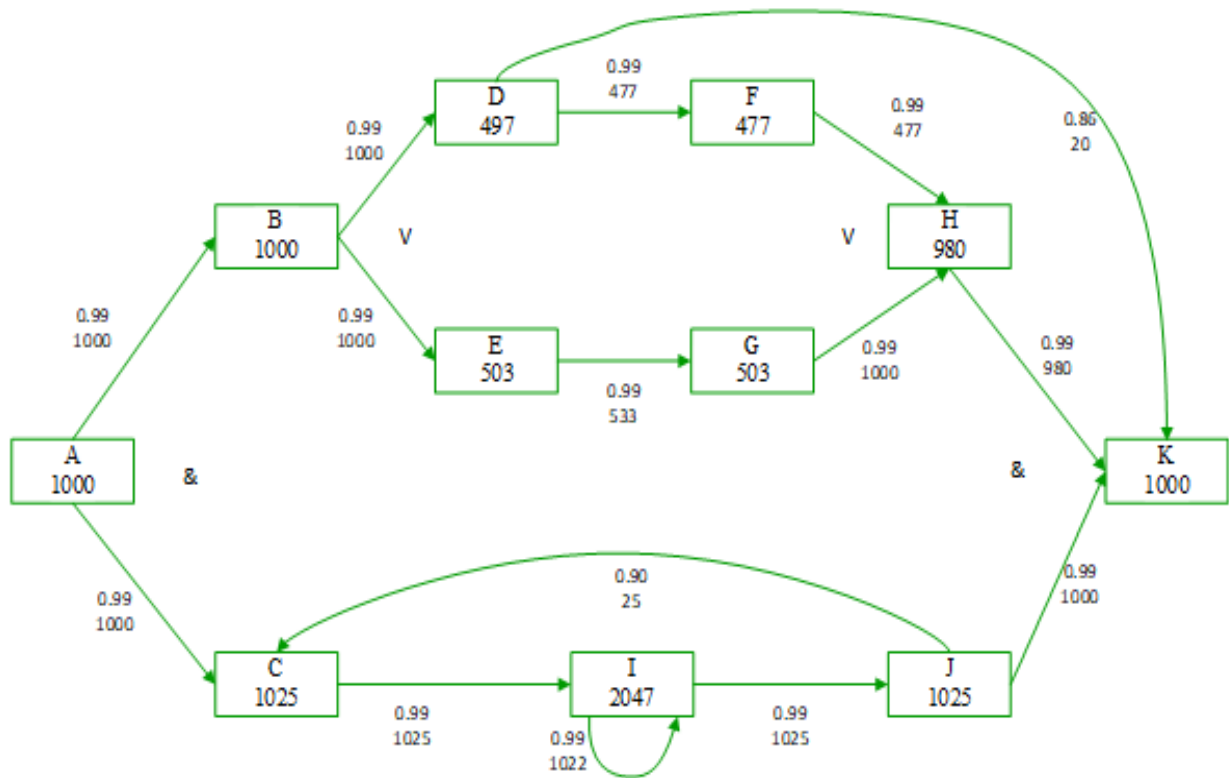


Figura 34. Modelo con umbrales umbral relativo al mejor =0.20, umbral de dependencia= 0.85.

Cuando se cuenta con un registro libre de ruido se pueden utilizar los parámetros extraordinariamente intolerantes: positivo umbral de observaciones =1000, umbral de dependencia=1.0 y relación con el mejor umbral= 0.00 el modelo resultante tendrá las conexiones correctas.

La función principal de los parámetros es la extensión del modelo sobre la base de un comportamiento poco frecuente, es decir, en base a los parámetros son los modelos que se crearan y las actividades que mostrara en él, una de las desventajas durante la minería de procesos es la combinación de patrones de baja frecuencia y el ruido (Werjters, Aalst y Medeiros, 2016).

Loaiza (2018) concluye que el minado heurístico es una alternativa del algoritmo Alpha, ya que es capaz de encontrar causalidad (relación entre actividades) obteniendo el descubrimiento de modelos desconectados o en forma de flor; es un algoritmo que se centra en el comportamiento

general y puede ignorar los sucesos excepcionales con los que se corre el riesgo de reducir la precisión y validez del descubrimiento.

### 2.2.2.2. Validez del descubrimiento

La eficacia de los modelos creados por los diferentes algoritmos de descubrimiento se mide en relación con cuatro criterios de calidad: el primero es el ajuste, el segundo la precisión, el tercero la generalización y el cuarto la simplicidad los cuales se describirán en el siguiente apartado (Diez, 2016 y Velázquez, 2017).

1. **Ajuste:** representa un gran número de posibles trazos, entre más grande sea el grado de ajuste mayor será el número de instancias explicadas por el modelo, conformado de cinco elementos:
  - Token Counter: Permite observar datos y transiciones perdidas que no se ejecutaron.
  - Failed Task: Permite ver las transiciones que no han sido completadas.
  - Remaining Tasks: Se identifican las transiciones que quedaron habilitadas pero que no fueron ejecutadas.
  - Path Coverage: Se identifican todas las transiciones que fueron ejecutadas.
  - Passed Edges: Indica la frecuencia durante la repetición de las instancias de cada proceso.
2. **Precisión:** cuanto más preciso sea el modelo, se comprime el comportamiento del proceso, no permite posibilidades lejanas al comportamiento del proceso visto en el registro de evento; conformado por cuatro características:
  - Always precedes: identifica las actividades que siempre precedieron entre sí en el registro.
  - Never precedes: determinar las actividades que nunca precedieron entre sí en el registro.
  - Always Follows: establecer las actividades que siempre siguieron a otra en el registro.
  - Never Follows: reconocer las actividades que nunca siguieron a otra en el registro.

3. **Generalización:** es la capacidad para representar un mayor número de posibles instancias o peticiones de proceso, cuanto más preciso sea, menor grado de generalización tendrá el proceso, está conformado por dos elementos:
- **Redundant invisible Tasks:** Permite identificar tareas duplicadas, es decir aquellas tareas que nunca se utilizan juntas en una de las posibles rutas de modelo.
  - **Alternative duplicate Tasks:** Permite identificar tareas invisibles, es decir tareas que no afectan el comportamiento del modelo y por lo tanto pueden ser eliminadas debido a que no tienen relevancia en el modelo.
4. **Simplicidad:** mide el grado de claridad del modelo, consiste en encontrar un modelo lo más sencillo posible.

### **2.2.3. Algoritmo minado inductivo básico e infrecuente.**

Este algoritmo funciona con una división recursiva (que se aplica de nuevo al resultado de haberlo aplicado previamente.), es decir que va creando las particiones de los estados hasta crear el modelo; en los siguientes apartados se describirán los dos tipos de minado inductivo existentes.

### **2.2.4. Algoritmo minado inductivo básico (Inductive miner IM)**

El algoritmo minado inductivo al igual que los otros algoritmos trabaja con un registro de eventos con la diferencia de que el registro se divide (Split log) para poder construir un árbol de procesos, cada parte del registro que ha sido dividido se trabaja por separado.

Usualmente este algoritmo se utiliza para analizar registros muy grandes; modela redes de flujo de trabajo o Workflow (Wf) por sus siglas en inglés, Redes Petri con tareas invisibles, registros con procesos lineales, registros con condiciones de ramificación o de división que pueden ser seguidos por una ruta alternativa de eventos invisibles (Nuritha, 2017).

Las redes de flujo de trabajo y los árboles de procesos están conformados de la siguiente manera:

Las redes de flujo de trabajo son una subclase de las redes Petri y se emplean para modelar los procesos de negocio los cuales están conformados por puntos de inicio (i) y fin (o) bien definidos en el proceso y se representa de la siguiente manera:

Red de flujo de trabajo  $(P, T, A, m_0, F, l)$  donde:

$P$  es un conjunto finito de lugares

$T$  conjunto finito de transacciones

$A \subset M((P \times T) \cup (T \times P))$  a son arcos y una relación múltiple de conjuntos de arcos

$i \in P$  tal que  $i = 0$  indica que existe una única entrada

$o \in P$  tal que  $o = 0$ . existe una sola salida

$i: m_0 = [i]$ . Indica el marcado inicial o entrada

$o: F = [0]$ . El marcado final o salida

Todos los lugares  $P$  y transacciones  $T$  son un camino de entrada (i) a la salida (o).

Los elementos de un diagrama de flujo de trabajo (figura 35) son los siguientes (Raeburn, 2024).

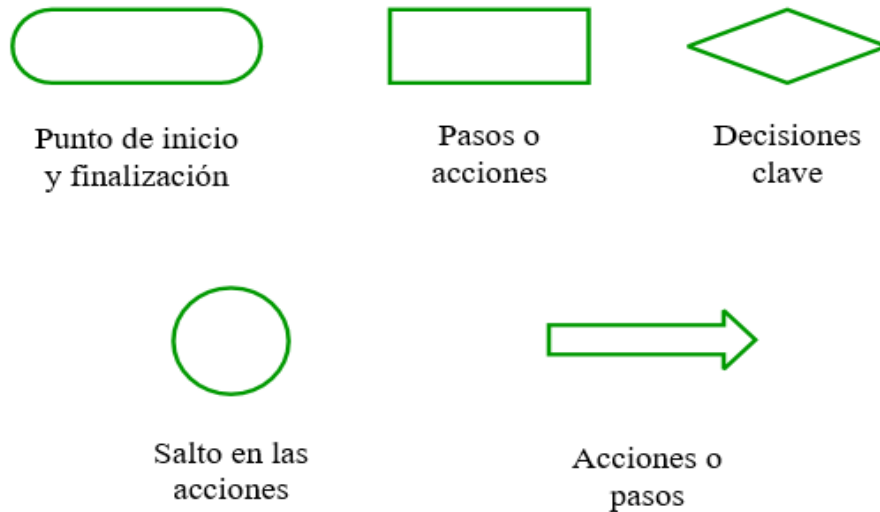


Figura 35. Elementos del diagrama de flujo de trabajo.

#### 2.2.4.2 Árboles de proceso

Otro tipo de modelado que emplea el algoritmo minado inductivo son los llamados arboles de proceso, donde los cuadros con alguna letra son las hojas del árbol y representan cada una de las actividades, los círculos son los nodos internos y se escriben con operadores que describirán el orden en el que se ejecutarán cada una de las actividades; existen cuatro tipos de operadores los cuales se describirán a continuación.

- Primer operador: composición secuencial ( $\rightarrow$ ) (figura 36), consiste en que todos los hijos se ejecutan de izquierda a derecha.

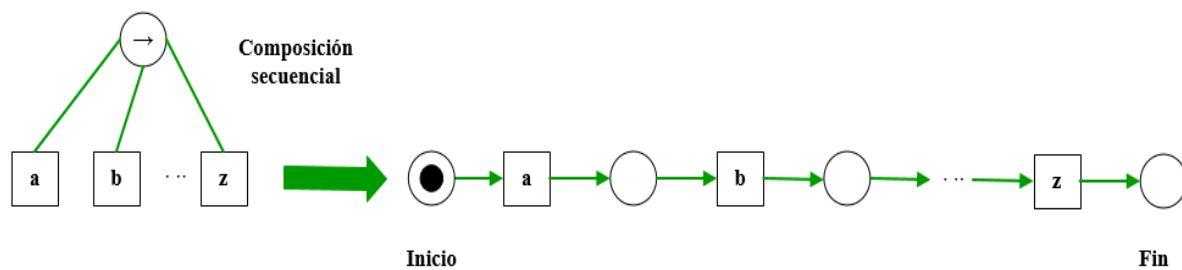


Figura 36. Operador secuencial.

- Segundo operador: elección exclusiva o xor ( $\times$ ) (figura 37), solo se ejecuta un hijo.

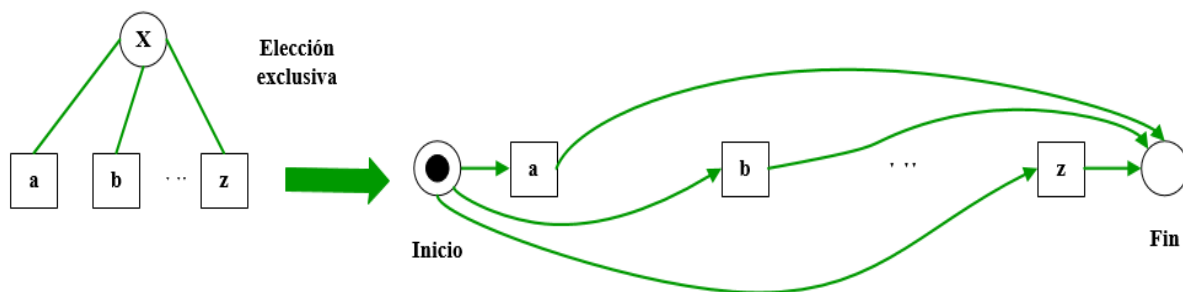


Figura 37. Operador XOR.

- Tercer operador: composición paralela o de concurrencia ( $\wedge$ ) (figura 38), todos los hijos se ejecutan y puede suceder en cualquier orden

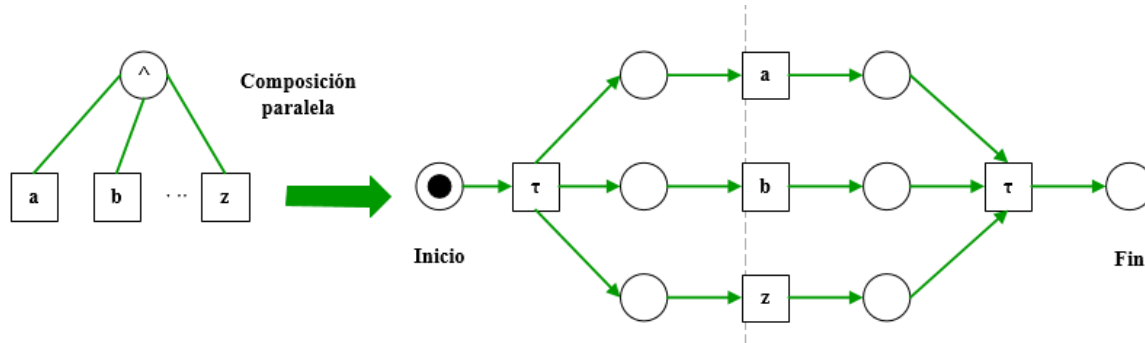


Figura 38. Operador paralelo.

- Cuarto operador: bucle de rehacer o simplemente bucle ( $\odot$ ) (figura 39), cuando existe el operador rehacer al menos dos hijos se pueden dividir en dos grupos, el primer hijo de la izquierda es llamado “do-parte” hacer parte y los demás hijos pertenecerán a la “parte de rehacer”.

Por el contrario, si el operador es de bucle se activa la “do-part” y se ejecuta al menos una vez y después uno de los hijos de la “parte de rehacer” se puede ejecutar y si fuese el caso debe seguirse como parte del “do-part”.

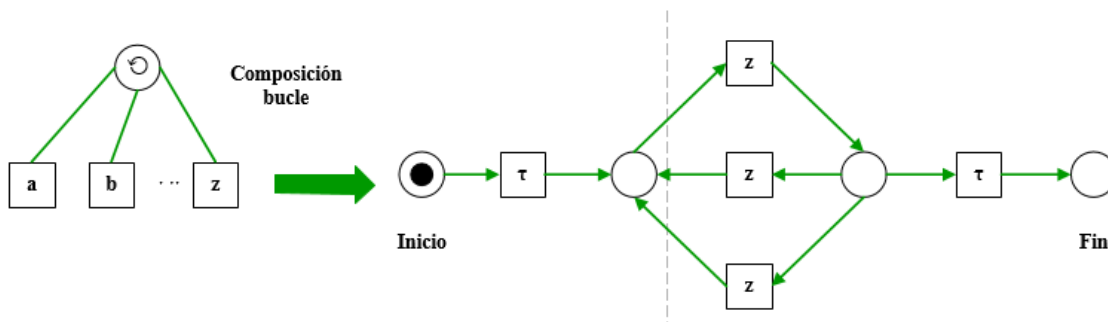


Figura 39. Operador bucle.

Según el tipo de operador a utilizar será la acción que se realizará dentro del modelo.

Un árbol de procesos se representa de la siguiente manera formal: checar si no falta algún punto

Si  $A$  es un conjunto finito de actividades  $T \notin A$ .  $\oplus = \{\rightarrow, x, \wedge, \odot\}$  es el conjunto del árbol de procesos de operadores.

Si  $a \in A \cup T$  entonces  $Q = a$  es un árbol de procesos.

Si  $n \geq 1$  y  $Q_1, Q_2 \dots Q_n$  son arboles de proceso y  $\{\rightarrow, x, \wedge\} \oplus \in$  entonces  $Q = \oplus (Q_1, Q_2 \dots Q_n)$  es un árbol de procesos.

Si  $n \geq 2$  y  $Q_1, Q_2 \dots Q_n$  son arboles de proceso entonces  $Q = \odot (Q_1, Q_2 \dots Q_n)$  es un árbol de procesos.

En la figura 40 se ejemplifica un árbol de procesos del registro de compra de boleto de una aerolínea, el proceso cuenta con ocho actividades (solicitud de registro, reiniciar solicitud, decisión, registrar boleto, examinar a fondo, examinar casualmente, pagar compensación y rechazar solicitud).

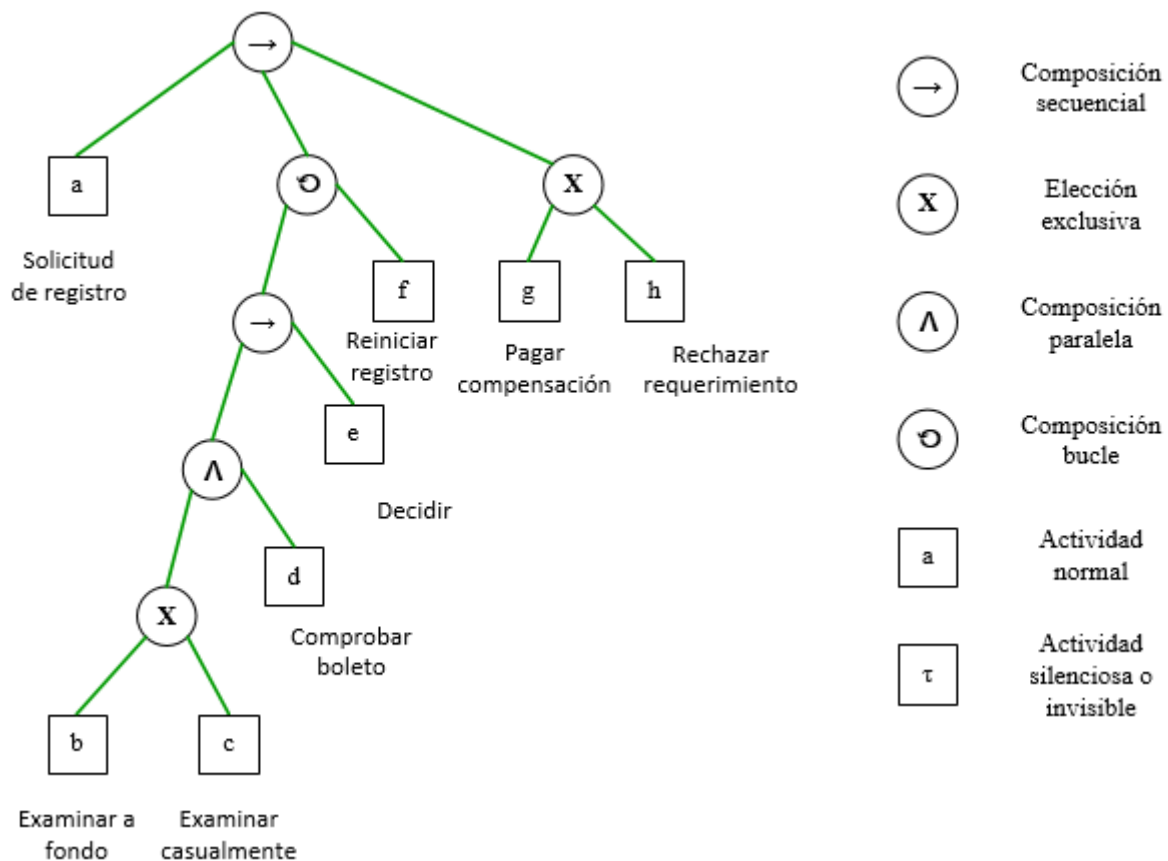


Figura 40. Árbol de procesos.

En el ejemplo anterior (figura 40) el proceso cuenta con ocho actividades (a, b, c, d, e, f, g, h); siempre comienza y termina con la actividad “a”, la notación del proceso es la siguiente:

$$\rightarrow (a, \odot (\rightarrow (\wedge(x(b, c), d), e), f), x(g, h))$$

Dado que el nodo raíz es un operador de secuencia ( $\rightarrow$ ) y el primer hijo es la actividad a, cada instancia de proceso de este árbol comenzara con la actividad “a”.

La actividad “a” es seguida por el subárbol que inicia con un bucle ( $\odot$ ), después de este se ejecuta el subárbol derecho que comienza con un xor ( $x$ ).

El ciclo comienza con el hijo que se encuentre más a la izquierda y puede retroceder sobre la actividad “f”. La parte do – (el hijo más a la izquierda del ciclo) comienza con una secuencia operador de bucle ( $\odot (\rightarrow (\wedge(x(b, c), d), e), f)$ ), por lo que primero se ejecuta el subárbol  $\wedge(x(b, c), d)$  y luego es seguido por la actividad e.

En el subárbol paralelo, d se ejecuta simultáneamente con xor elección entre b y c, después de ejecutar el subárbol de bucle, se ejecuta el subárbol más a la derecha el cual es  $x(g, h)$  para hacer una elección entre “g” y “h”.

En la figura 41 se observa un árbol de procesos con siete actividades cualesquiera (a, b, c, d, e, f y g), existe una secuencia en la que siempre inicia el proceso con la actividad “a” y terminan con la actividad “g”, al mismo tiempo se ejecutan actividades paralelas “b, c y d” y son dos actividades de elección “e y f” en la que solo se utilizara una de ellas según sea el caso; a continuación, se muestran las posibles rutas del proceso.

a, b, c, d, e, g

a, b, c, d, f, g

a, c, d, b, f, g

a, b, d, c, e, g

a, d, c, b, f, g

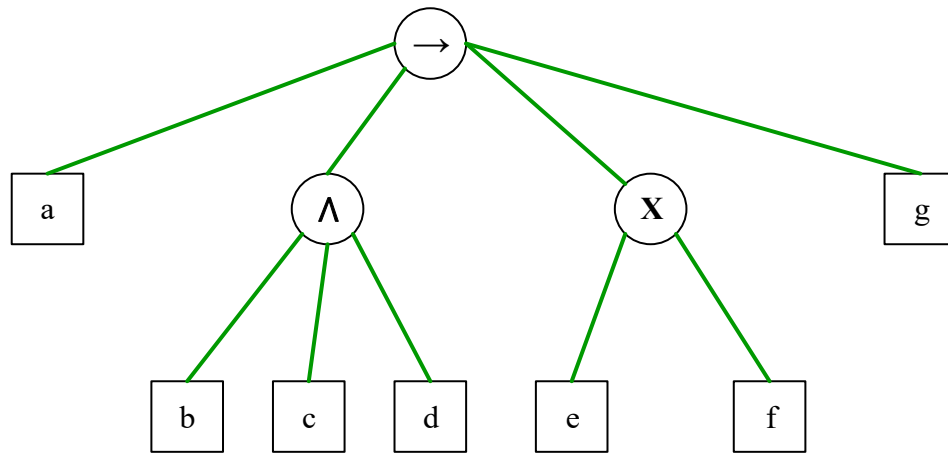


Figura 41.Árbol de proceso de siete actividades.

La conversión de los árboles de proceso a redes Petri es una de las funciones con las que cuenta el minado inductivo en el siguiente ejemplo se observa la conversión del árbol de procesos (figura 41) a un modelo red Petri (figura 42).

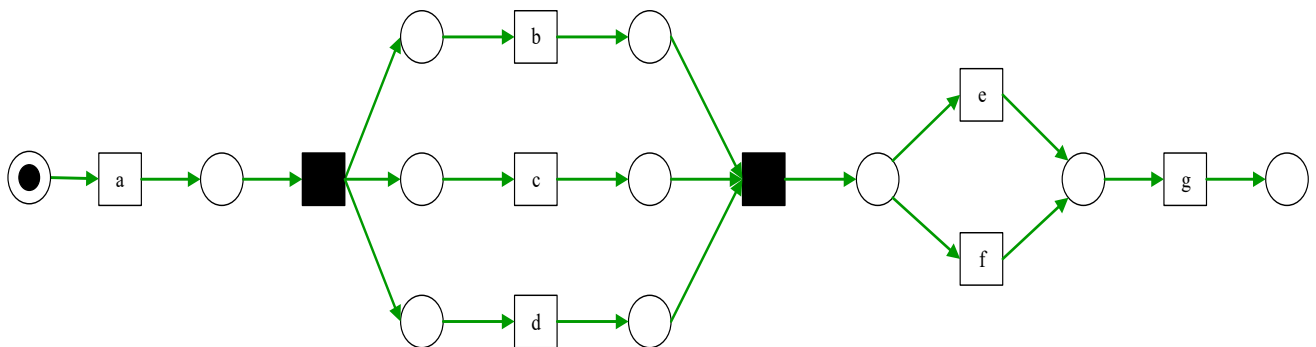


Figura 42.Conversion del árbol de procesos de 7 actividades a red Petri.

En el siguiente ejemplo se muestra la conversión de un árbol de procesos (figura 43) de cinco actividades (a, b, c, d y e) a una red Petri (figura 44), el proceso inicia con la actividad “a” y termina con la actividad “d”, existe una elección entre las actividades “e, b y c” y a la par se ejecutan las actividades paralelas “b y c”.

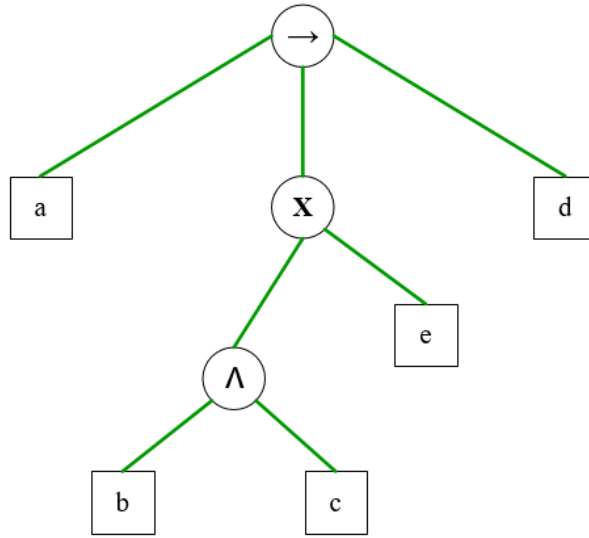


Figura 43. Árbol de procesos de cinco actividades.

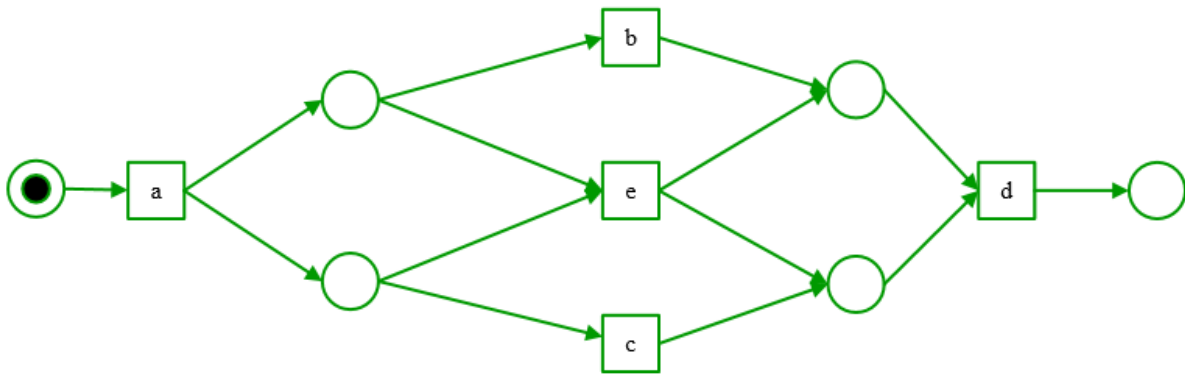


Figura 44. Conversión del árbol de procesos a red Petri.

Para aplicar el algoritmo minado inductivo básico se deben emplear tres pasos fundamentales los cuales (Figura 45) se pueden repetir o aplicar indefinidamente (Pohl, 2019).

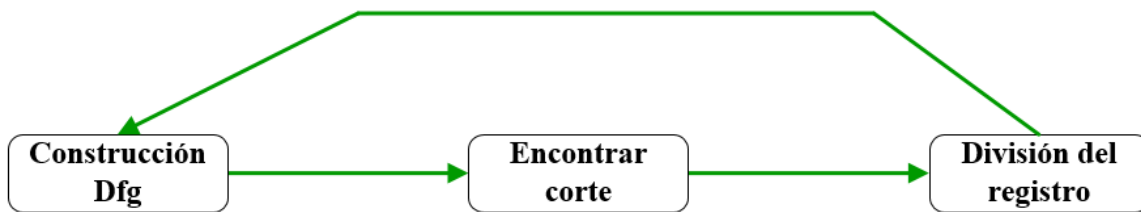


Figura 45. Pasos del minado inductivo.

1.- Primer paso: construcción del gráfico de seguimiento directo (Dfg).

El Primer paso consiste en calcular un gráfico de seguimiento directo o Dfg (Dfg por sus siglas en inglés Directly Follows Graph) basado en el registro de eventos, compuesto por nodos y arcos, cada nodo corresponde a una actividad dentro del registro, el número total de nodos será igual al número de las diferentes actividades que se encuentran en el registro, los arcos son los enlaces que unen las actividades.

2. Paso dos: detección de cortes.

Consiste en encontrar un corte dentro del gráfico, el algoritmo debe devolver un operador y una partición de corte, el cual dividirá el registro de eventos dentro de uno o más sub-bloques (este paso es llamado división del registro); el algoritmo es recursivo puesto que repite los mismos pasos para cada sub-bloque, es decir, para cada sub-bloque se calcula el gráfico, busca un corte y divide el registro de acuerdo con el corte encontrado.

Existen cuatro diferentes tipos de cortes que aplica el algoritmo en los gráficos de seguimiento directo y tienen el siguiente orden:

- Primer corte: elección exclusiva o xor:

Para detectar un corte xor, primero se calculan los componentes que están conectados dentro del Dfg, después se mide la longitud de los componentes conectados, si esa longitud es mayor que uno indica que los componentes conectados están en más de un conjunto, lo que lleva a obtener un corte xor y los componentes conectados son devueltos como partición cortada.

- Segundo corte: corte de secuencia.

Primero se unen todos los grupos más cercanos en parejas, es decir, se dividen las actividades en conjuntos de tal modo que no debe haber dos actividades en conjuntos diferentes que estén emparejados, después se fusionan los nodos dispersos en pares; los conjuntos resultantes se ordenan y si uno encuentra un corte de secuencia se regresa una partición cortada.

- Tercer corte: corte concurrente o paralelo.

Las actividades se dividen en subconjuntos, se fusionan todos los pares de los subconjuntos que no están completamente conectados en ambas direcciones; después se fusionan los pares de conjuntos donde están contenidas las actividades en el primer y en el segundo conjunto. Por último, se buscan conjuntos que no tengan actividad inicial ni actividad final y se fusiona en otro conjunto arbitrario. Si la partición con los subconjuntos resultantes tiene una longitud mayor que uno, se obtiene el corte concurrente o paralelo y regresa una partición cortada.

- Cuarto corte: corte de bucle.

Para este corte el algoritmo crea un conjunto que cuenta con todas las actividades de inicio y fin, después se crean conjuntos a partir de todas las actividades restantes, si se encuentra un corte se devuelve la partición resultante como cortada, por lo que no puede haber conexión entre un conjunto con un índice pequeño con un conjunto de índice más grande.

A continuación, se definen algunas notaciones que se emplean para usar los cuatro cortes:

→ Representa un gráfico de seguimiento directo.

$\sum$  (→) describe todas las actividades dentro del DFG.

$a \rightarrow b$  indica que la actividad “a” es seguida directamente por la actividad “b” en el registro de eventos por lo menos una vez.

$a \rightarrow + b$  revela que la actividad “a” es seguida en algún momento por la actividad “b” en el registro, pero pueden ser ejecutadas entre a y b.

$a \circlearrowleft b$  indica que existe una relación mínima entre las actividades “a” y “b” así que si “a” se ejecuta dos veces “b” pudiera ejecutarse entre esas dos ejecuciones de a.

3. Paso tres: división del registro.

Ya que se encontró un corte (paso 2) el algoritmo minado inductivo divide el registro en varios sub-bloques, para cada uno de los sub bloque la recursividad continúa, de acuerdo con el corte encontrado será la función de división que se aplicara, para lo cual existen cuatro diferentes funciones de división.

- Primera división: División de elección exclusiva.

Para hacer el corte de particiones el registro primero se divide en sub-bloques, se repite y hace la misma implementación de particiones con todos los trazos del registro, agrega cada trazo y actividades en el registro utilizando la siguiente función (figura 46).

```
function xorSplit (L, (x,  $\Sigma_1, \dots, \Sigma_n$ ))
  Vi:  $L_i \leftarrow [t \in L \mid \forall e \in t: e \in \Sigma_i]$ 
  return  $L_1, \dots, L_n$ 
end function
```

Figura 46. División de elección exclusiva.

- División de secuencia

Lo que hace la función es tomar el registro y un corte de partición como entrada, para cada corte se crea un sub-bloque y hace una iteración a través de todos los rastros, para cada trazo se crea un nuevo rastro vacío, tan pronto como una actividad este contenida en el trazo se comienza a añadir esta y las siguientes actividades en la partición de corte, si se encuentra una actividad no contenida se crea una nueva traza vacía y se vuelve a iterar sobre los eventos de la siguiente traza. Se utiliza la siguiente función (figura 47) cuando un corte de secuencia es encontrado.

```

function xorSPLIT (L(x,  $\Sigma_1, \dots, \Sigma_n$ ))
   $\forall i: L_i \leftarrow [t \in L \mid \forall e \in t: e \in \Sigma_i]$ 
  return L1, ..., Ln
end function

```

Figura 47. División de secuencia.

- División simultánea

La función (figura 48) toma el registro y un corte de partición como entrada, después los sub-bloques son construidos y se iteran a través de las huellas del registro, se checa si cada actividad está contenida en el trazo, si no es así se agrega en el índice del registro.

```

function CONCURRENTSPLIT (L, ( $\Lambda, \Sigma_1, \dots, \Sigma_n$ ))
   $\forall i: L_i \leftarrow [t \mid \Sigma_j \mid t \in L]$ 
  return L1, ..., Ln
end function

```

Figura 48. División simultánea.

- División de bucle

La función (figura 49) toma un corte de partición y el registro como entrada para la creación de los sub-bloques e itera sobre las particiones cortadas, por cada partición se crea un nuevo sub-bloque, después se itera sobre el sub-bloque, para cada partición cortada se itera sobre las huellas del registro para buscar secuencias de actividades, todas las secuencias encontradas se agregan como rastro al sub bloque actual.

$$\begin{aligned}
& \text{function } \text{LOOPSPLIT} \left( L, (\mathcal{O}, \Sigma_1, \dots, \Sigma_n) \right) \\
& \quad \forall i: L_i \leftarrow [t_2 \mid t_1 \cdot t_2 \cdot t_3 \in L \wedge \\
& \quad \quad \Sigma([t_2]) \subseteq \Sigma_i \wedge \\
& \quad \quad (t_1 = \in \vee (t_1 = \langle \dots, a_1 \rangle \wedge a_1 \notin \Sigma_i)) \wedge \\
& \quad \quad (t_3 = \in \vee (t_3 = \langle a_3, \dots \rangle \wedge a_3 \notin \Sigma_i)) \\
& \quad \text{return } L_1, \dots, L_n
\end{aligned}$$

Figura 49. División de bucle.

La recursividad llega a su fin cuando se detectan los llamados casos base, es decir, cuando no se encuentra ningún tipo de corte en el registro se aplica el fall through (caída a través), existen seis caídas diferentes que se deben aplicar en el siguiente orden específico.

1. Rastros vacíos: se aplica cuando el registro contiene trazas vacías.
2. Actividad una vez por traza: esto pasa cuando una actividad aparece con precisión una vez en cada trazo de un registro de evento, si hay varias actividades se aplica esta caída y se elige una traza arbitraria o parcial.
3. Actividad concurrente: esta caída intenta encontrar un corte omitiendo una actividad del registro, comienza con la iteración de todas las actividades para una nueva actividad, por lo que un nuevo registro es construido y filtrado, si un corte puede ser encontrado en el Dfg del registro esta caída es aplicada, se crea un nuevo segundo registro agregando una nueva traza y se vuelve a filtrar una actividad del registro para encontrar un corte de secuencia y así sucesivamente para cada registro que haya.
4. Bucle tau estricto: para aplicar esta caída se crea un segundo registro en base al registro actual, este se construye dividiendo las trazas, el resultado es de obtenido de cada ocurrencia de una actividad final seguida de una actividad inicial y se almacena en el segundo registro; si después

de la iteración de todos los trazos en el registro existen más trazos el comportamiento del bucle se presenta y la recursividad continua.

5. Bucle Tau: es similar al bucle estricto el cual crea un segundo registro, itera todos los trazos y estos se dividen en una actividad de ocurrencia, el resultado de las trazas es almacenado, si se tienen más trazas en el registro esta caída se aplica.

6. Modelo de flor (figura 50): esta caída se aplica a cualquier evento del registro sin trazas vacías, basado en todas las actividades se retorna un modelo en flor que tiene todas las actividades del registro de eventos; si se aplica un modelo de flor caída, la recursividad siempre será hasta el fin.

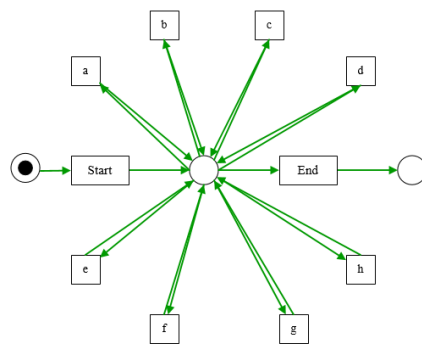


Figura 50. Ejemplo de modelo en flor.

Sin embargo, el modelo de flor (figura 50) no es un buen modelo de proceso, por lo que primero se deben aplicar otras fallas que permitan continuar con la recursión del minero inductivo básico.

Una debilidad del minado inductivo es que no toma en cuenta las frecuencias, lo que conduce a obtener modelos de procesos de menor calidad si existe un comportamiento poco frecuente o con desviación en el registro; por lo que el minado inductivo poco frecuente es un segundo algoritmo inductivo que es capaz de reducir ese comportamiento y ofrece funciones para filtrar comportamientos desviados e infrecuentes.

#### **2.2.4.3. Minado inductivo infrecuente.**

El minado inductivo poco frecuente (figura 51) se compone de la misma recursión utilizada en el minado inductivo básico con la diferencia de que se agregan dos pasos más al mismo (filtro del Dfg y corte infrecuente) los cuales se describen a continuación.

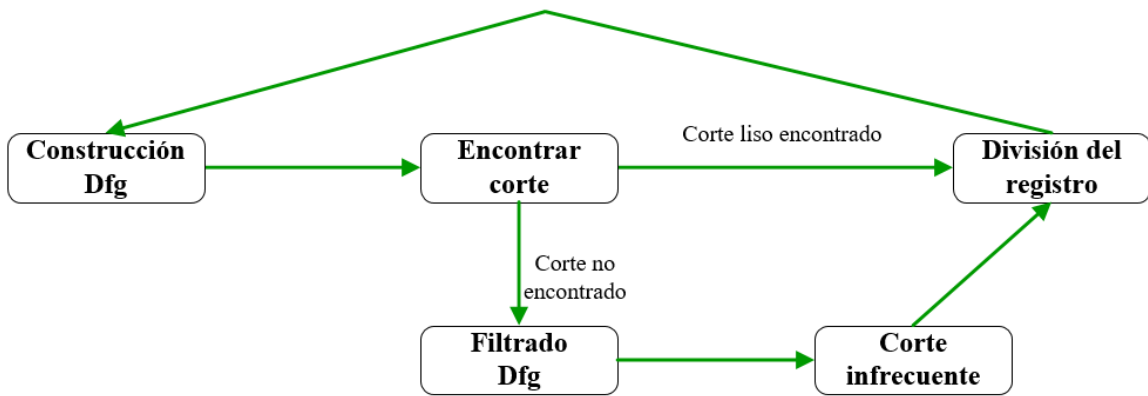


Figura 51. Minado poco infrecuente.

**Primero.** se calcula el Dfg (Grafico de seguimiento directo).

**Segundo:** se realiza una detección de corte del Dfg, para la detección de corte se utilizan las mismas funciones que en el minado inductivo básico (xor, secuencia, paralelo y bucle), si se encuentra un corte se divide y la recursividad continúa con los sub-bloques resultantes.

- Si no se encuentra ningún corte, el Dfg es filtrado basado en un parámetro de umbral f.
- Una vez que se filtra el Dfg, se aplica nuevamente la detección de corte utilizando las mismas funciones de detección utilizadas en el minero inductivo básico.
- Si se encuentra un corte se divide el registro y la recursividad continúa en los sub-bloques resultantes.
- Si no se encuentra ningún corte, se aplica una caída a través.

La tabla 6 ofrece una descripción general de las funciones que se pueden utilizar para el algoritmo minero inductivo y las funciones que se implementaron según Pohl, 2019.

Tabla 6. Funciones.

Funciones	Nuevas funciones
Corte exclusivo	Filtro Dfg
Secuencia de corte	División de corte exclusivo
Corte concurrente o paralelo	División de secuencia

Corte de bucle	División de bucle
División concurrente	Fall through – trazo vacío
Fall through - una actividad por traza	Caso base – una sola actividad
Fall through - actividad concurrente	
Fall through – bucle Tau estricto	
Fall through – bucle Tau	
Fall through – modelo en flor	
Caso base – registro vacío	

Nota: Elaboración propia tomado de Pohl, 2019.

### 2.2.5. Algoritmos basados en región.

El objetivo de este algoritmo es encontrar lugares determinados en los diagramas de transición aplicando abstracciones en el comienzo, final y en donde solo importe el siguiente evento o el anterior. Este tipo de algoritmo requiere un alto grado de prueba y error puesto que no existe una forma de abstracción, como resultado del algoritmo se obtiene una red Petri la cual debe ser comparada con la original a través de la huella de tal forma que se contemplen todas las trazas originales (González, 2016).

Para lo cual primero se debe realizar el diagrama de transición aplicando las abstracciones para después construir la red Petri equivalente al diagrama, a continuación, se muestra el siguiente ejemplo:

$$L1 = [(a, b, c, d)3, (a, c, b, d)2, (a, e, d)]$$

Tomando el registro anterior se tiene el siguiente diagrama de transición (figura 52).

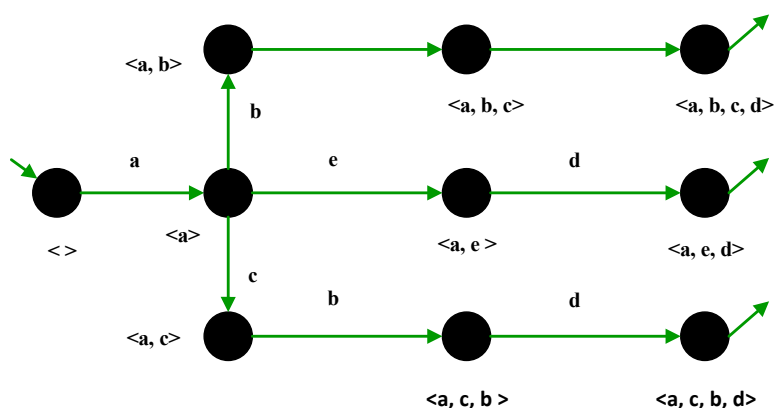


Figura 52. Diagrama por regiones.

Al realizar las abstracciones se tiene el siguiente diagrama (figura 53).

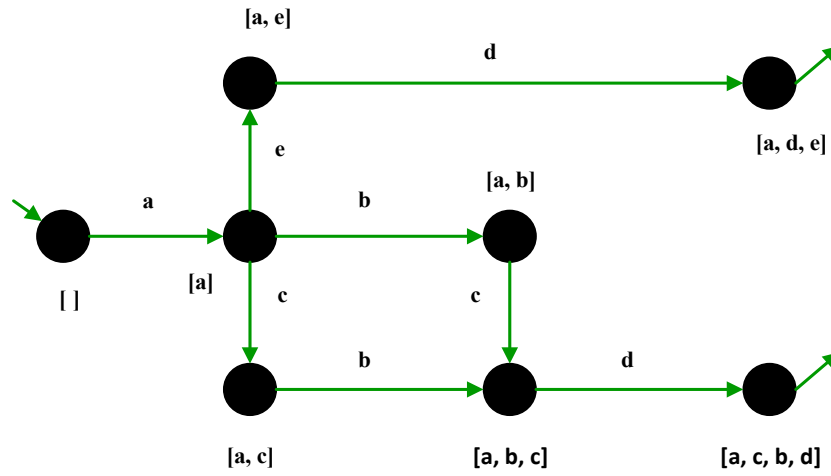


Figura 53. Diagrama resultante de las abstracciones.

A partir del diagrama final (figura 54) se crea la red Petri equivalente (figura 55).

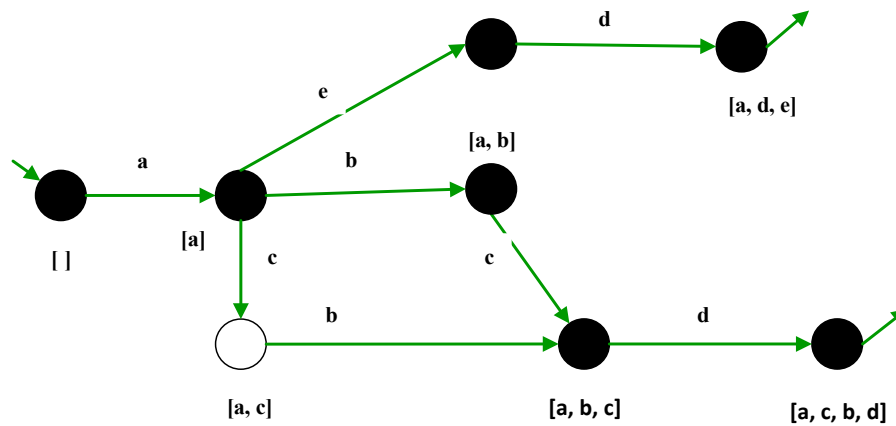


Figura 54. Diagrama final.

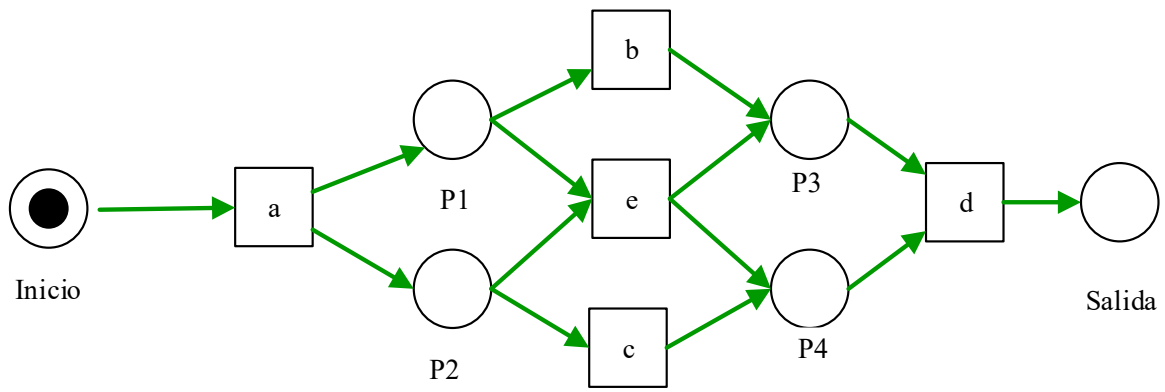


Figura 55. Red Petri equivalente al diagrama final.

# **CAPÍTULO III**

## **Modelos de procesos para la minería de procesos**

### 3. Modelos de procesos

Los modelos son el resultado que se obtienen de hacer un análisis de un proceso con los algoritmos, los cuales son una representación visual de un proceso de negocio; como lo son los diagramas de flujo y la notación universal de procesos de modelado o BPMN, redes Petri, entre otros, estos modelos sirven para identificar los flujos de trabajo y hacer mejoras en los procesos de una organización. Se pueden obtener dos tipos de modelos, el primer modelo muestra el proceso tal como se obtiene sin implementar ningún cambio y el segundo modelo muestra el proceso una vez realizado los cambios, modificaciones o mejoras.

Los modelos de proceso se pueden crear de forma manual, sin embargo, es un método lento y menos efectivo en comparación de hacer uso de un software de modelado de procesos con el que se pueden crear y ejecutar fácilmente modelos de proceso y diagramas de flujo de manera sistematizada y para identificar áreas de mejora (Process Maker, 2020).

El modelado de procesos de negocios puede incluir desde diagramas simples y dibujados a mano hasta algunos más complejos con elementos expandibles que brinden suficientes detalles de la implementación. La idea detrás de la gestión de procesos de negocio es crear un ciclo de vida de mejora continua (Lucid software Inc., 2024).

Van der Aalst (2011) alude que los modelos de procesos ayudan a las organizaciones a administrar la complejidad brindando información y la documentación de los procedimientos, además, los modelos son utilizados para deducir la planificación y control de proceso, estos modelos generalmente se utilizan en la administración de operaciones las cuales aplican técnicas de análisis específicos y se usan para responder a preguntas específicas. Por lo tanto, la creación de modelos a mano es una tarea difícil y es propensa a errores típicos los cuales incluyen:

- Un modelo con una versión idealizada de la realidad.
- Incapacidad para capturar apropiadamente el comportamiento humano, es decir, un modelo matemático puede modelar maquinas o personas que trabajan en una línea de montaje, sin embargo, cuando se modelan personas involucradas en múltiples procesos y expuestas a múltiples prioridades ya no lo son.

- Dependiendo de los datos de entrada y las preguntas a responder se tiene que elegir el nivel de abstracción adecuado, puesto que el modelo puede ser demasiado abstracto y no responder a preguntas relevantes.

La minería de procesos puede resolver esos inconvenientes y aminorar las cargas de trabajo debido a que facilita la construcción de mejores modelos en menor tiempo, la minería descubre y genera procesos de modelo de manera automática.

Para representar el comportamiento que se obtiene de aplicar los algoritmos a los registros de eventos de forma gráfica se utilizan varios modelos como lo son: las redes Petri, los BPMN (Business Process Model Notation), las redes causales, modelos EPC, BPMN, Petri Net, UML, CNet, entre otras.

A continuación, se describirá como trabaja cada uno de los modelos más utilizados por la minería de procesos.

### 3.1. Modelo red Petri

Bogarin (2018) indica que las redes Petri son gráficos con dos tipos de nodos enlazados por arcos, el primer tipo de nodo se llama lugar y está representado por un círculo y estos pueden almacenar múltiples valores llamados tokens, el segundo tipo de nodo son las transiciones, se representan con cuadros e identifican elementos activos de la red (figura 56).

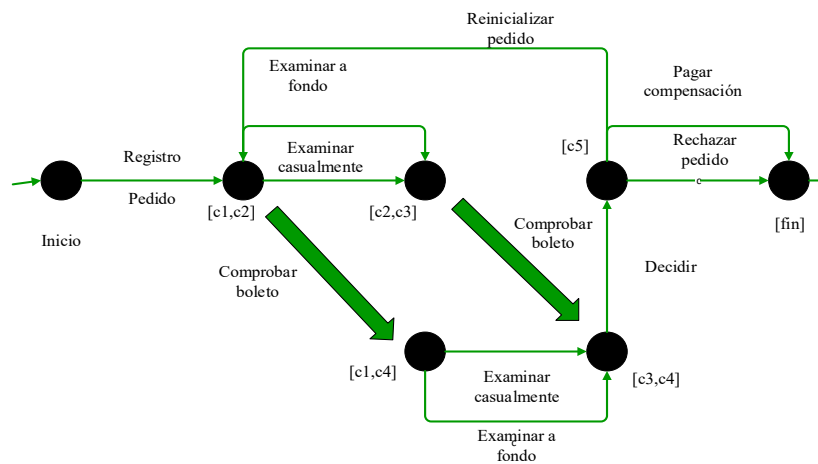


Figura 56. Red Petri.

Coursera (2021) y González (2016) indican que las redes Petri son un sistema de modelado de procesos formado por cuatro elementos: el primero las transiciones, segundo los lugares, tercero los arcos y por último y no menos importante los tokens (figura 57).

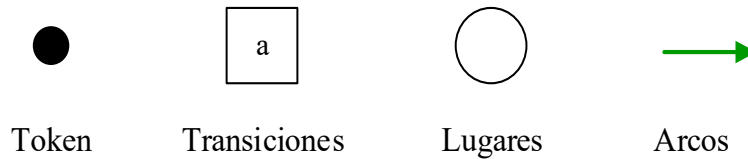


Figura 57. Elementos de los modelos BPMN.

### 3.1.1. Patrones que forman una red Petri

Los cuatro patrones en los que se puede expresar una red Petri al momento de modelar son:

- **XOR Separación**

Se da cuando una actividad pasa por un lugar y en la salida da como resultado dos actividades diferentes (figura 58).

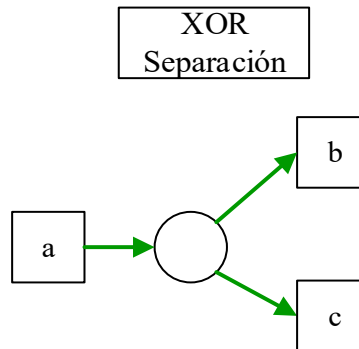


Figura 58. Patrón XOR separación.

- **XOR Unión**

Sucede cuando se tienen dos actividades que pasan por el mismo lugar y como resultado se tiene una actividad en la salida (figura 59).

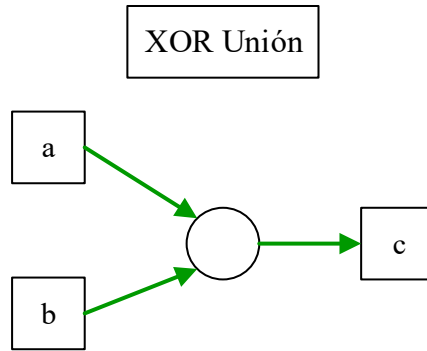


Figura 59. Patrón XOR unión.

- **AND Separación**

Al tener una actividad de entrada y pasar por dos lugares se separa con sus respectivas actividades en la salida (figura 60).

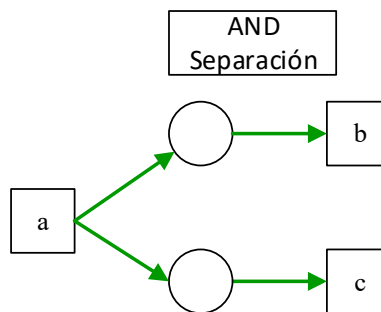


Figura 60. Patrón AND separación.

- **AND Unión**

Cuando hay dos actividades de entrada que pasarán por sus respectivos lugares cada una en la salida se unen y darán como resultado una única actividad (figura 61).

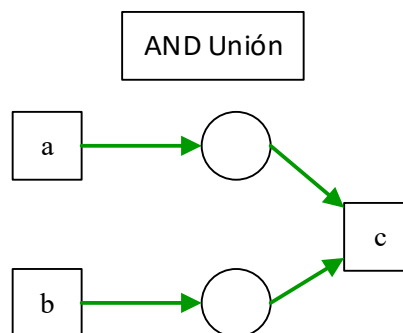


Figura 61. Patrón AND unión.

### 3.1.2. Desplazamiento del token a través del modelo

Las fichas (token) se sitúan dentro de los lugares y se desplazan a lo largo de los arcos hacia las transiciones y de regreso a lugares, durante el desplazamiento las fichas podrán multiplicarse en el caso de que se pase de un lugar a varios, o reducirse cuando se necesiten varias fichas para activar una transacción que solo produzca un lugar.

Las fichas se ejecutan según los eventos en el registro, a continuación, se muestra un ejemplo por pasos de la secuencia de un token a través de una red Petri (figura 62,63,64, 65 y 66).

Se tiene el siguiente registro y como se observa en la figura 62 el token se encuentra en el inicio del modelo.

[a, b, c, d] y [a, c, b, d]

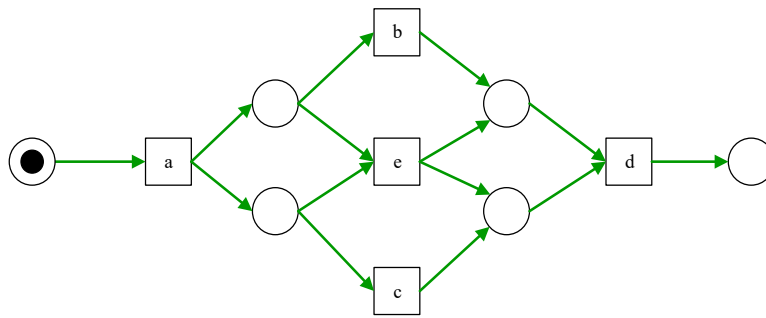


Figura 62. Paso 1 desplazamiento del token.

Al desplazarse el token por la actividad “a” y después a los lugares se observa que la ficha se multiplica para colocarse en dos lugares (figura63).

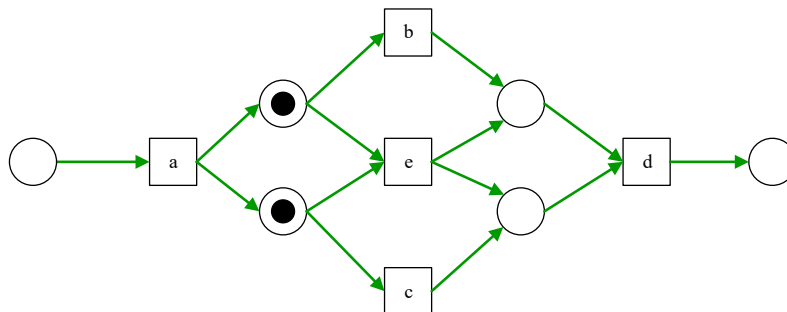


Figura 63. Paso 2 desplazamiento del token.

Los tokens se desplazan uno a la vez pasando por las actividades “b, ó e, y “e ó c” para ocupar la posición en el siguiente lugar (figura 64).

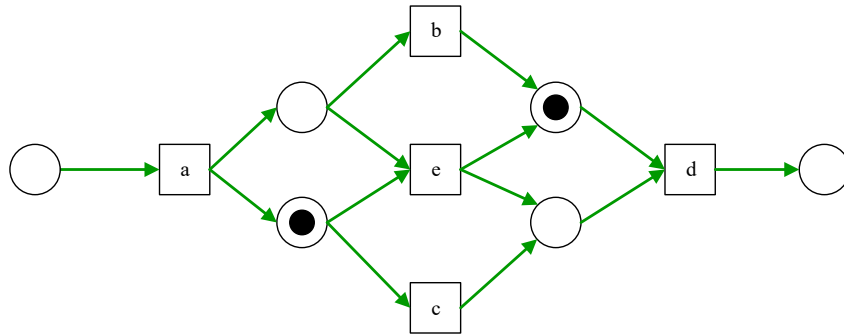


Figura 64. Paso 3 desplazamiento del token.

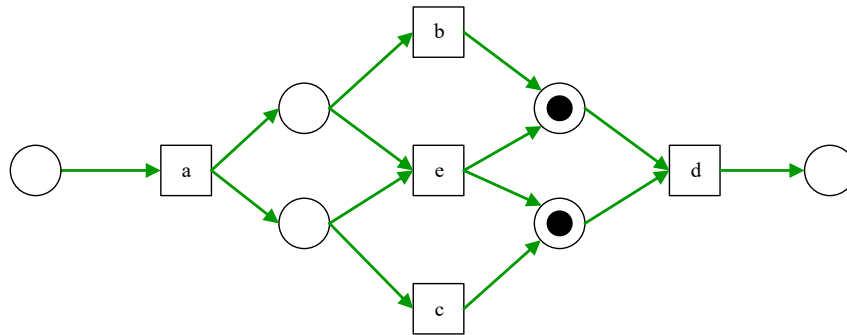


Figura 65. Paso 4 desplazamiento del token.

Por último, el token vuelve a ser uno y ocupa su lugar en la casilla de finalización debido a que ya no existen más actividades ni lugares a ocupar (figura 66).

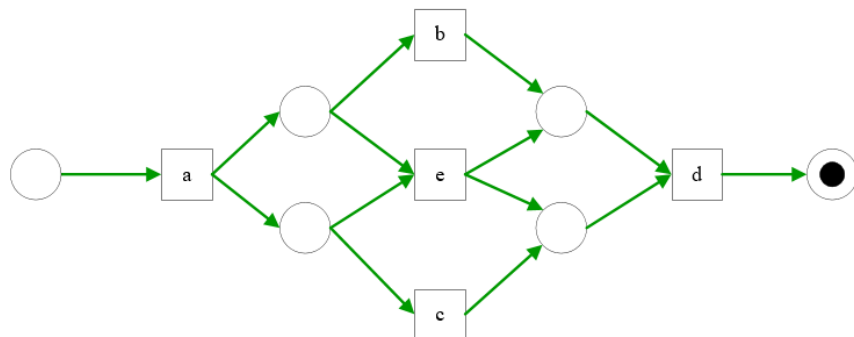


Figura 66. Finalización del desplazamiento del token.

### 3.2. Modelo BPMN

Los modelos BPMN (Business Process Model Notation) son una notación gráfica que describe la lógica de los pasos de un proceso, este modelo fue diseñada para coordinar la secuencia de los procesos y los mensajes que fluyen entre los usuarios de las diferentes actividades, además, utiliza un Diagrama de Procesos de Negocio (DPN) que sirve para representar la secuencia de las actividades que ocurren durante un proceso utilizando la técnica de “Flow chart”, además de incluir toda la información necesaria para el análisis (Nextech, 2021).

González (2016) indica que el modelado BPMN es similar a las redes Petri, los elementos de notación (figura 67) cambian y las tareas se ejecutan cumpliendo las compuertas lógicas.

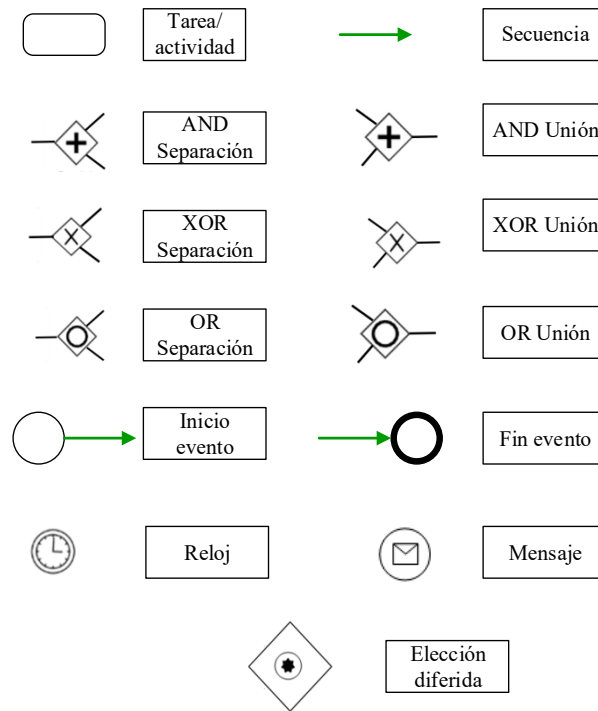


Figura 67. Notación modelos BPMN.

En los modelos BPMN existen algunos eventos que pueden ser desencadenados ya sea por el reloj o por un mensaje, también puede haber construcciones más avanzadas como la elección diferida, es decir, en la figura 68 existe una carrera entre el tiempo y el mensaje, si el mensaje es recibido antes de un punto en particular en el tiempo se realiza la actividad “z” de lo contrario se realiza la actividad “y”.



Figura 68. Elección diferida.

Van der Aalst (2021) indica que los BPMN son utilizados para modelar procesos y también como en las redes Petri tienen una semántica de tokens, es decir, poseen un inicio, van pasando a través de las actividades hasta llegar a su destino (figura 69).

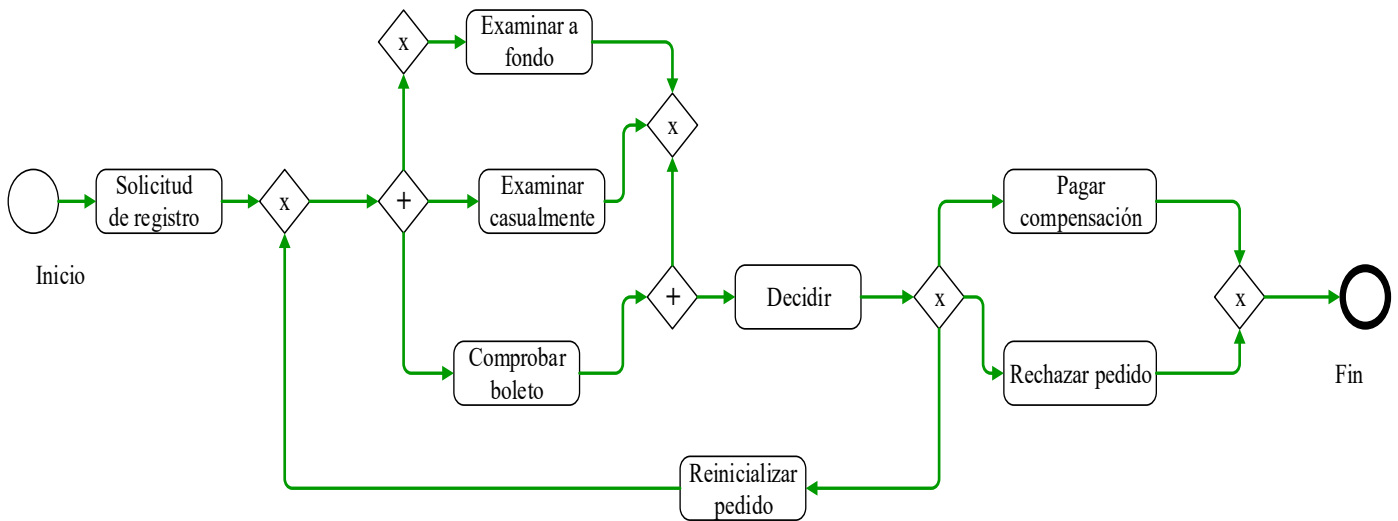


Figura 69. Modelo BPMN.

Cabe señalar que para que los modelos sean sólidos deben tener un inicio y un fin bien definidos.

### 3.3. Modelo EPC

El modelo EPC o Cadena de procesos impulsada por eventos, utiliza funciones que corresponden a las actividades, conectores que refieren a las puertas de enlace de un modelo BPMN y además emplea colores en las notaciones del modelo (figura 70 y 71) (Van der Aalst, 2021).

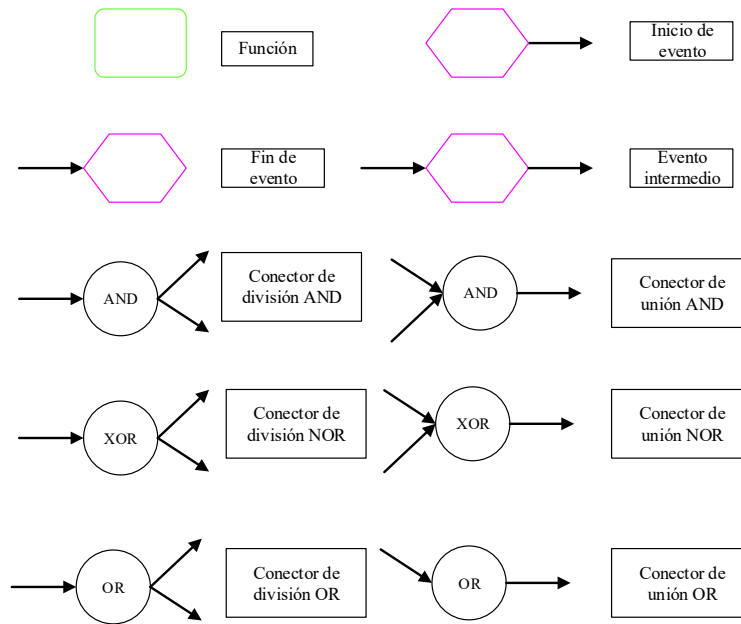


Figura 70. Notación EPC.

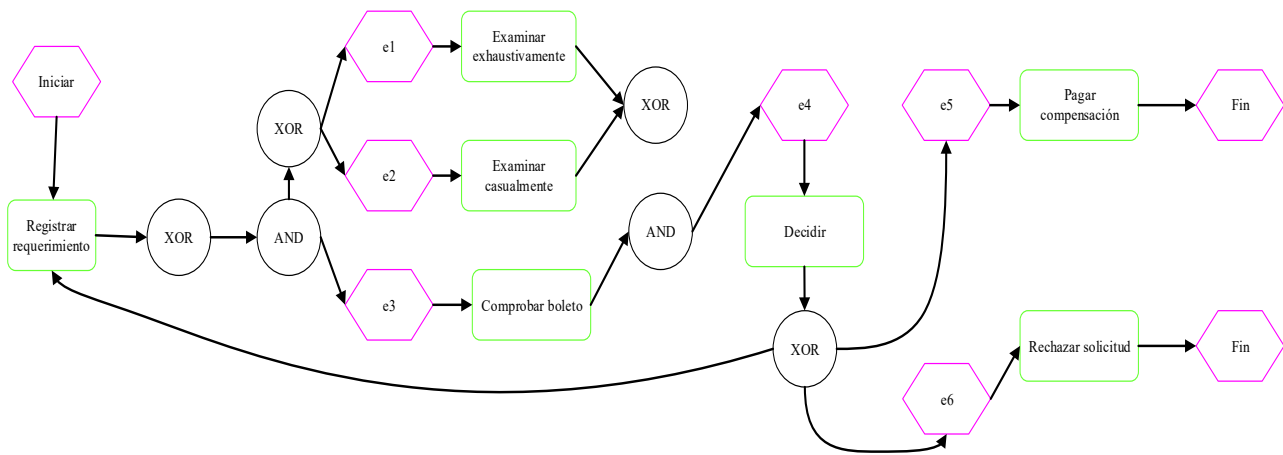


Figura 71. Modelo EPC.

### 3.4. Modelo red causal (Causal nets o C-Nets)

Las redes causales o C-nets utilizan transiciones, tareas y funciones, a través de modelos que utilizan los lugares (Petri nets), condiciones (Yawl), conectores y eventos (EPC), compuertas y enlaces (BPMN), todos esos elementos conectan actividades, pero no dejan alguna marca en el registro de eventos por lo que necesitan ser deducidos analizando el comportamiento (figura 72 y 73).

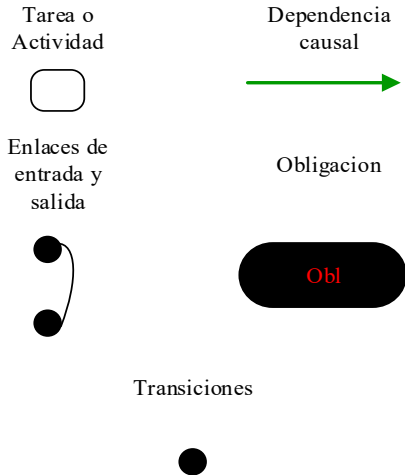


Figura 72. Notación Redes Causales.

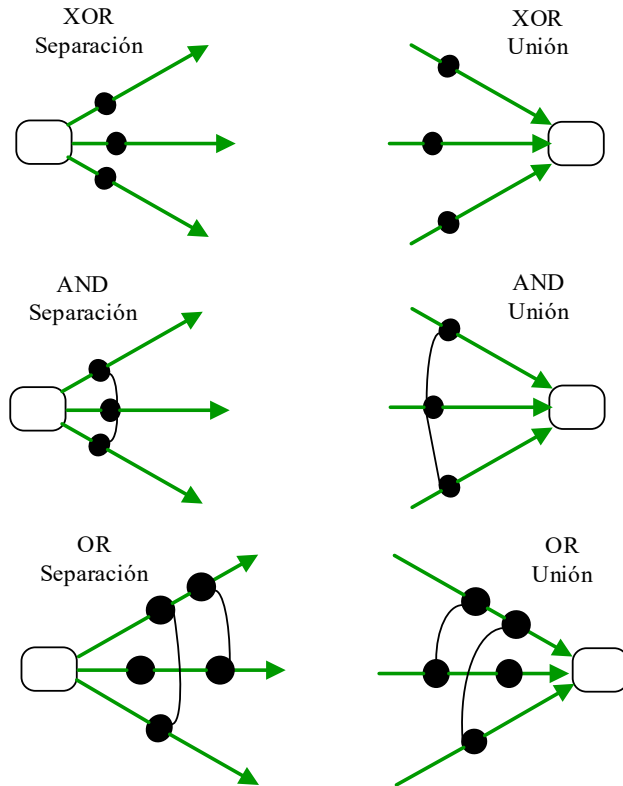


Figura 73. Compuertas.

Desde que el registro no provee información concreta acerca de los lugares, condiciones, conectores, compuertas y eventos, algunos algoritmos de la minería usan una representación que consiste de solo actividades sin conexión de elementos como lo son los C-Nets (Coursera, 2021 y Van der Aalst, 2016).

Las redes causales (C-nets) son una representación adaptada a la minería de procesos, es un gráfico donde los nodos representan las actividades y los arcos representan las dependencias causales, cada actividad tiene un conjunto de posibles enlaces de entrada y posibles enlaces de salida que revelan el posible comportamiento, en él se agregan más trayectorias que no son ejecutables pero que muestran los posibles comportamientos que pudieran existir en el (figura 74).

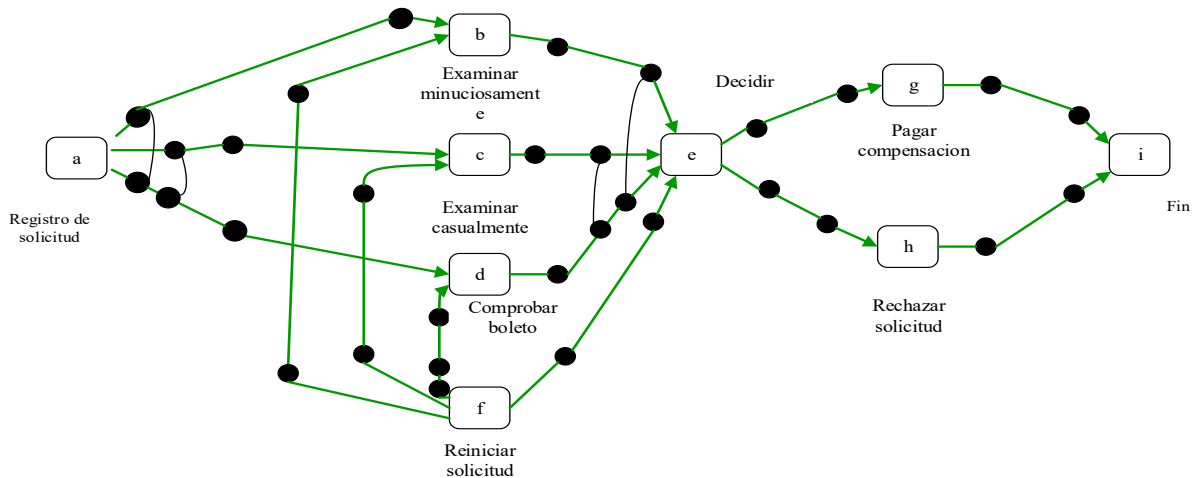


Figura 74. Ejemplo de red causal.

Los C-nets son utilizados por varias técnicas de minería de procesos como lo es el minado heurístico, además trabaja con los BPMN, las redes Petri, los EPC y es capaz de modelar con las compuertas XOR, AND, OR, pero no con transiciones silenciosas o duplicadas, evitando los modelos no firmados (modelos en el que se evita tener puntos muertos).

A continuación, se muestra de manera formal la notación C-nets el cual da a conocer si el comportamiento de seguimiento encaja con el del modelo, para lo cual debe existir una actividad inicial, una final, dependencias y conexiones entre los diferentes elementos y enlaces; al tener una actividad de entrada será la primera que ocurrirá y por tanto debe pasar en la primera posición del rastro, no en ninguna otra posición posterior y al ejecutarse tiene que elegir un enlace de salida, conocer las posibles secuencias.

### Notación formal de la C-net

$A \subseteq A$  es un conjunto finito de actividades.

$a_i \in A$  es el inicio de la actividad.

$a_o \in A$  es el fin de la actividad.

$D \subseteq A \times A$  es la relacion de dependencia.

$$AS = \{X \subseteq (A) \mid X = \emptyset \vee 0 \notin X\}^2$$

$I \in A \rightarrow AS$  define el conjunto de posibles enlaces de entrada por actividad

$O \in A \rightarrow AS$  define el conjunto de posibles enlaces de salida por actividad

Tal que

$$D = \{(a_1, a_2) \in A \times A \mid a_1 \in \cup_{as \in I(A_2)} as\}$$

$$D = \{(a_1, a_2) \in A \times A \mid a_2 \in \cup_{as \in I(A_1)} as\}$$

$$\{a_i\} = \{a \in A \mid I(a) = \{\emptyset\}\}$$

$$\{a_o\} = \{a \in A \mid O(a) = \{\emptyset\}\} \text{ y}$$

todas las actividades en el grafico  $(A, D)$  son un sendero de  $a_i$  a  $a_o$ .

En el siguiente ejemplo (figura 75) se describirá brevemente como trabajan las redes causales.

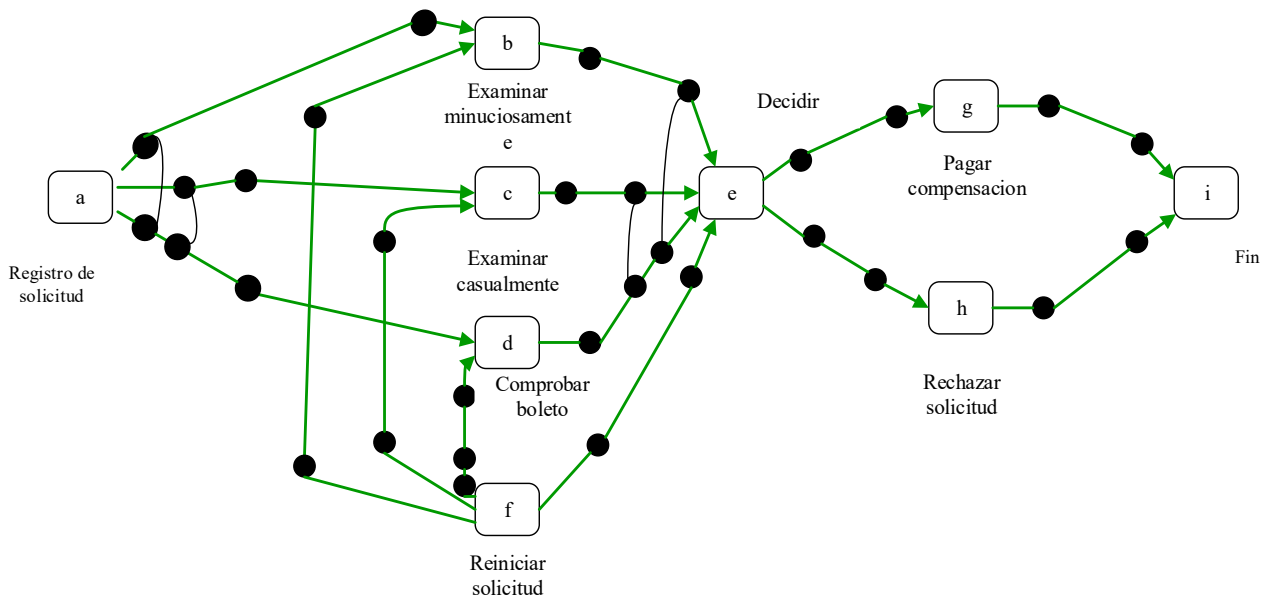


Figura 75. Ejemplo de red causal.

Al tener una red causal se observa que se tiene una actividad de entrada “a” (figura 76).

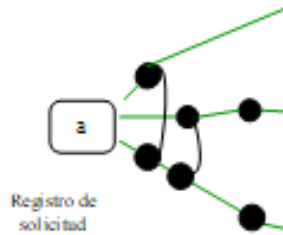


Figura 76. Entrada a.

Existen dos posibilidades de enlace de salida: “b, d” ó “c, d” (círculos color rojo) de las cuales solo se debe elegir uno. Al elegir el enlace (b, d) se crea una obligación para esos enlaces (óvalo con las letras Obl) (figura 77), los círculos rojos indicaran la secuencia que se eligió; al ir haciendo elecciones las obligaciones se van sucediendo, es decir, las obligaciones se van eliminando de la posición en la que estaban y se crean en la nueva elección de salida (los enlaces de entrada eliminan las obligaciones y los enlaces de salida los crean).

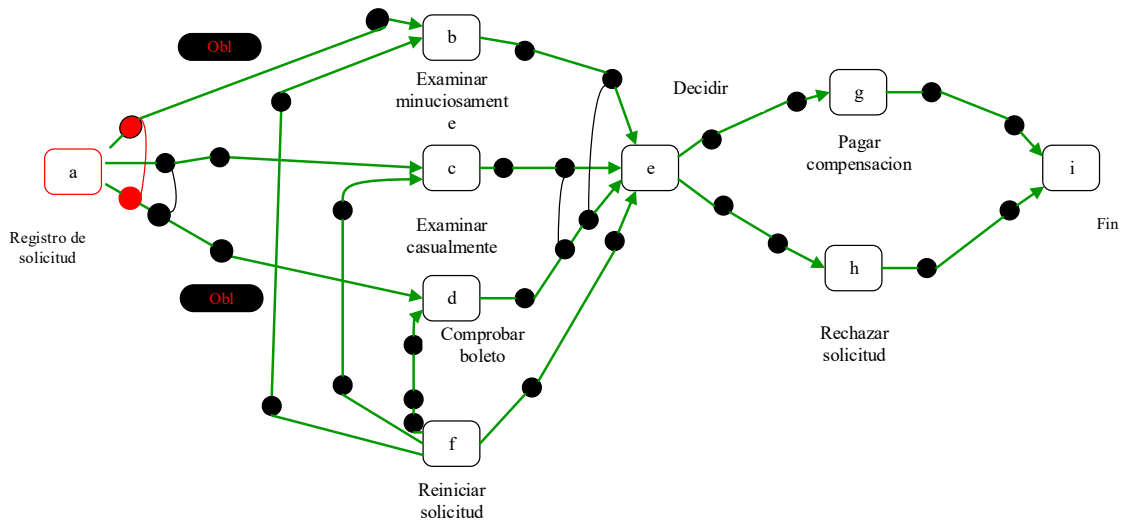


Figura 77. Elección de enlace y creación de obligaciones.

Al elegir “b” se elimina la obligación entre “a y b” y se crea una nueva obligación para “b y e” y en el futuro se tendrá que ejecutar y coincidir con uno de los enlaces de entrada de “e” (figura 78).

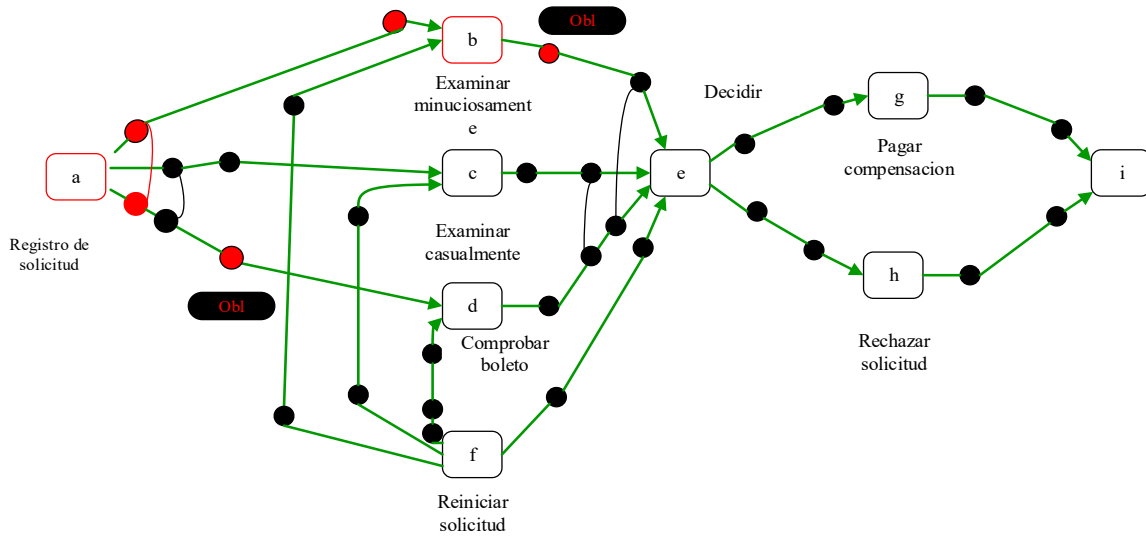


Figura 78. Ejecución de b y creación de una nueva obligación.

Después se ejecutará “d” y se eliminará la obligación de entrada entre “a y d” y se creará una nueva obligación de salida para “d y e” (figura 79).

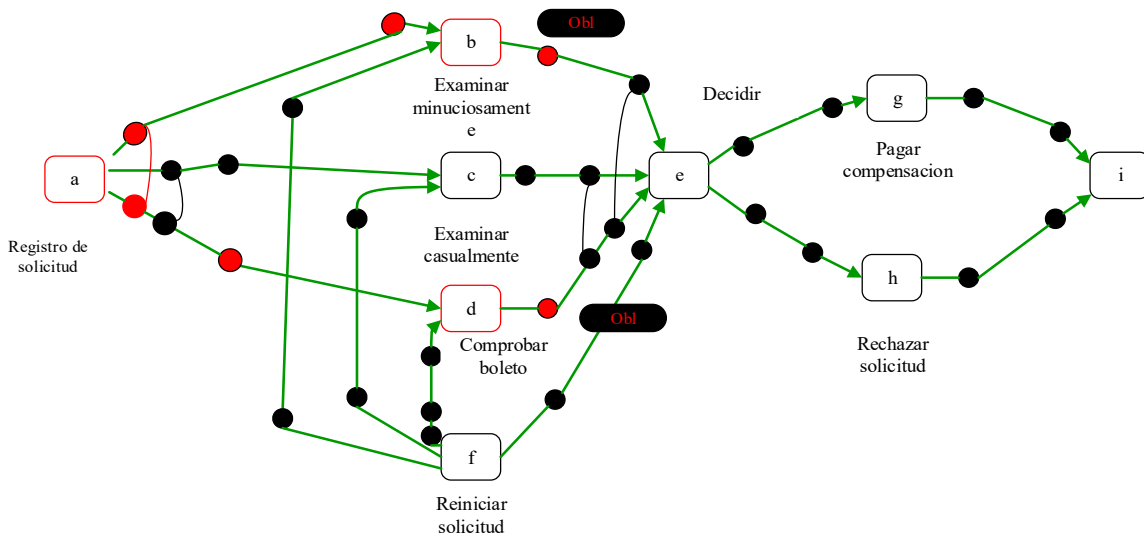


Figura 79. Ejecución de d y creación de obligaciones.

Después se ejecutará “e”, se eliminarán dos obligaciones simultáneamente (b y d), sin embargo, solo se eliminarán dos arcos entrantes de los tres que conectan y se creará una nueva obligación de salida para “e y g” por lo que “g” tendrá que suceder en el futuro (figura 80).

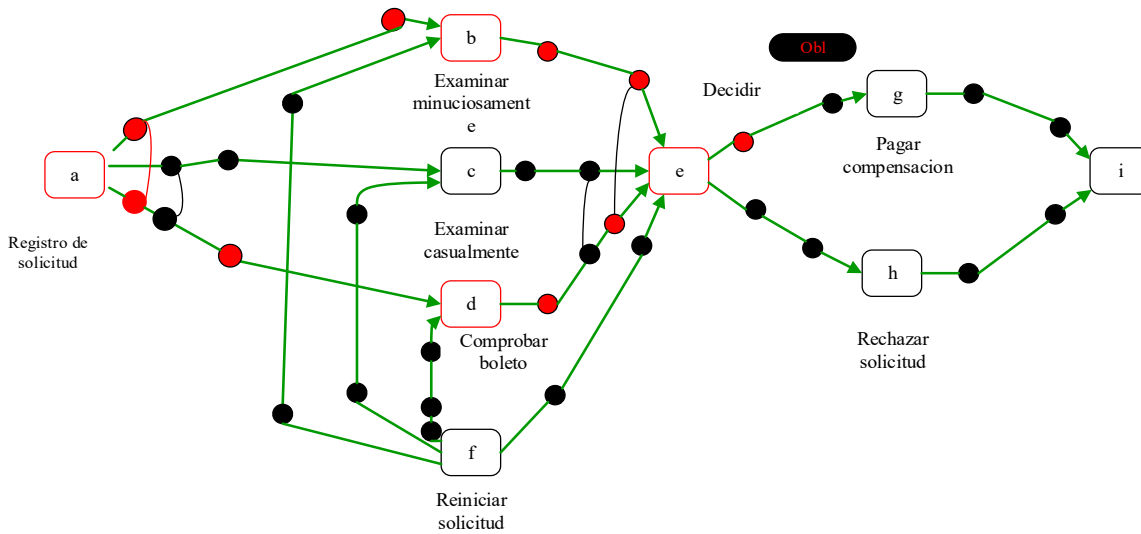


Figura 80. Ejecución de e y creación de obligación.

Se ejecuta “g”, se elimina la obligación de entrada entre “e y g”, se crea una nueva obligación de salida entre “g y i” (figura 81).

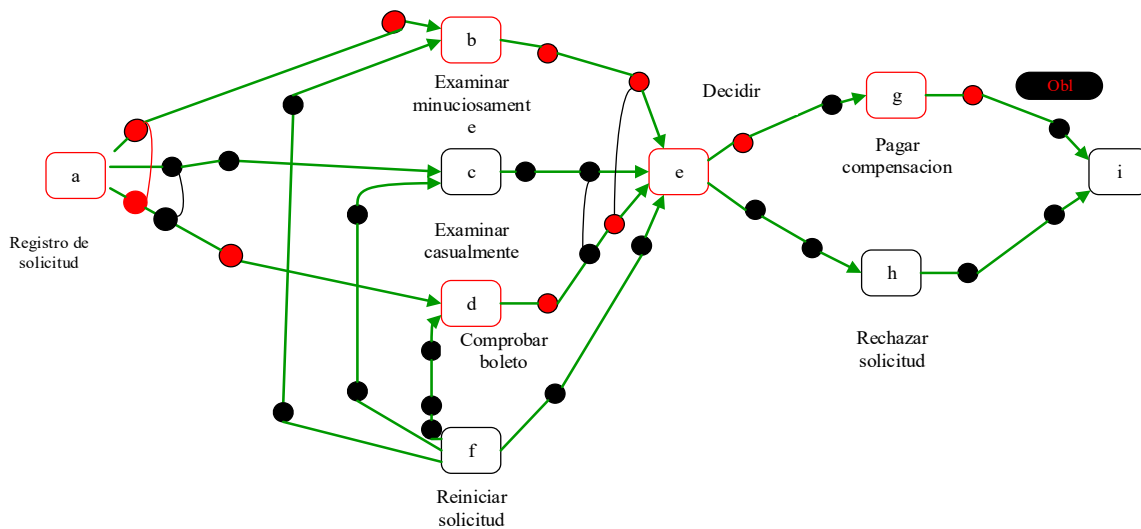


Figura 81. Ejecución de g y eliminación de obligación.

Se elimina la obligación entre “g é i” y se ejecutará “i” que es la actividad final, se debe llegar a la actividad final sin dejar alguna actividad sin ejecutar. (figura 82).

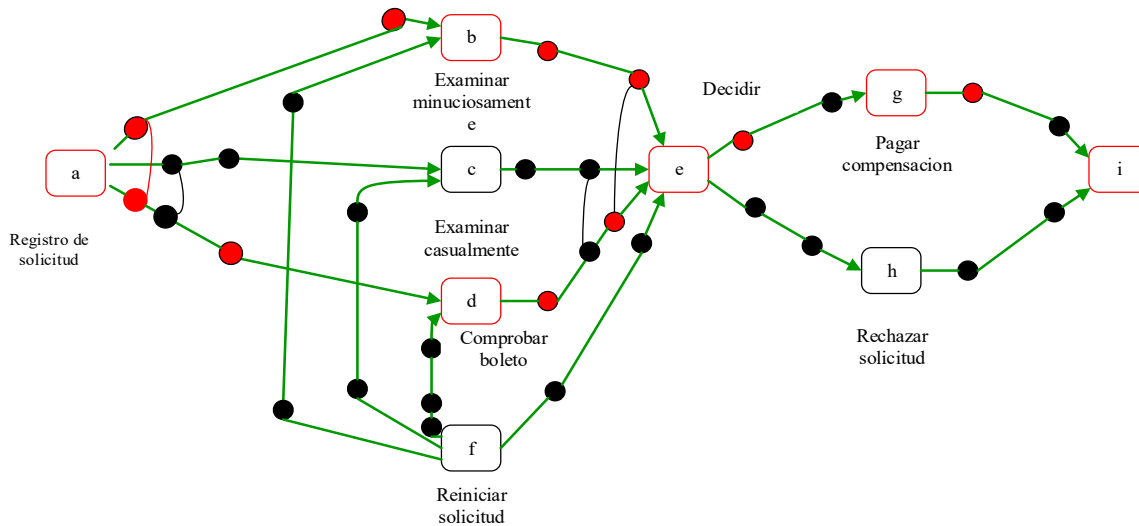


Figura 82. Actividad final.

La ejecución de las actividades dependerá de la ruta o camino a tomar, se deben crear y eliminar obligaciones según el modelo hasta llegar a la actividad final. Para ejecutar las actividades se deberá seguir las siguientes reglas del modelo:

- Empezar siempre con la actividad de inicio y solo puede suceder una a la vez.
- Al final siempre hay una actividad final designada.
- Las actividades intermedias no pueden ser actividades iniciales o finales.
- Las obligaciones son parecidas a los tokens y necesitan estar allí para ser consumidas por los enlaces de entrada.
- Al final no debe haber obligaciones restantes.
- Solo se describen posibles caminos sin decir cuando se está tomando una decisión

### 3.5. Diferencia entre de las C-nets y las redes Petri

Cada secuencia de enlace valida en la C-nets corresponde a una secuencia de disparo de la red de flujo de trabajo, va desde el marcado inicial hasta el marcado final llamado secuencia de disparo valida, cada ejecución de la red de flujo de trabajo empieza en el marcado inicial y termina en el marcador final, esto corresponde a una secuencia de enlace valida.

Cuando se habla de las redes causales solo son consideradas las secuencias de enlaces válidas y no requiere de un flujo de trabajo solido porque se usa la semántica declarativa (cuantas

secuencias validas hay en el modelo). Sin embargo, en la red Petri el flujo de trabajo pueden existir muchas ejecuciones que no corresponden a una secuencia de enlace valida.

### **3.6. Tipos de procesos**

Una vez modelado el proceso, se pueden tener dos tipos de procesos: el primero el proceso Lasaña en este proceso las trazas siguen un patrón determinado y un flujo de ejecución, el cual permitirá garantizar el entendimiento del proceso con poco esfuerzo, son procesos fáciles de descubrir y entender (figura 83).

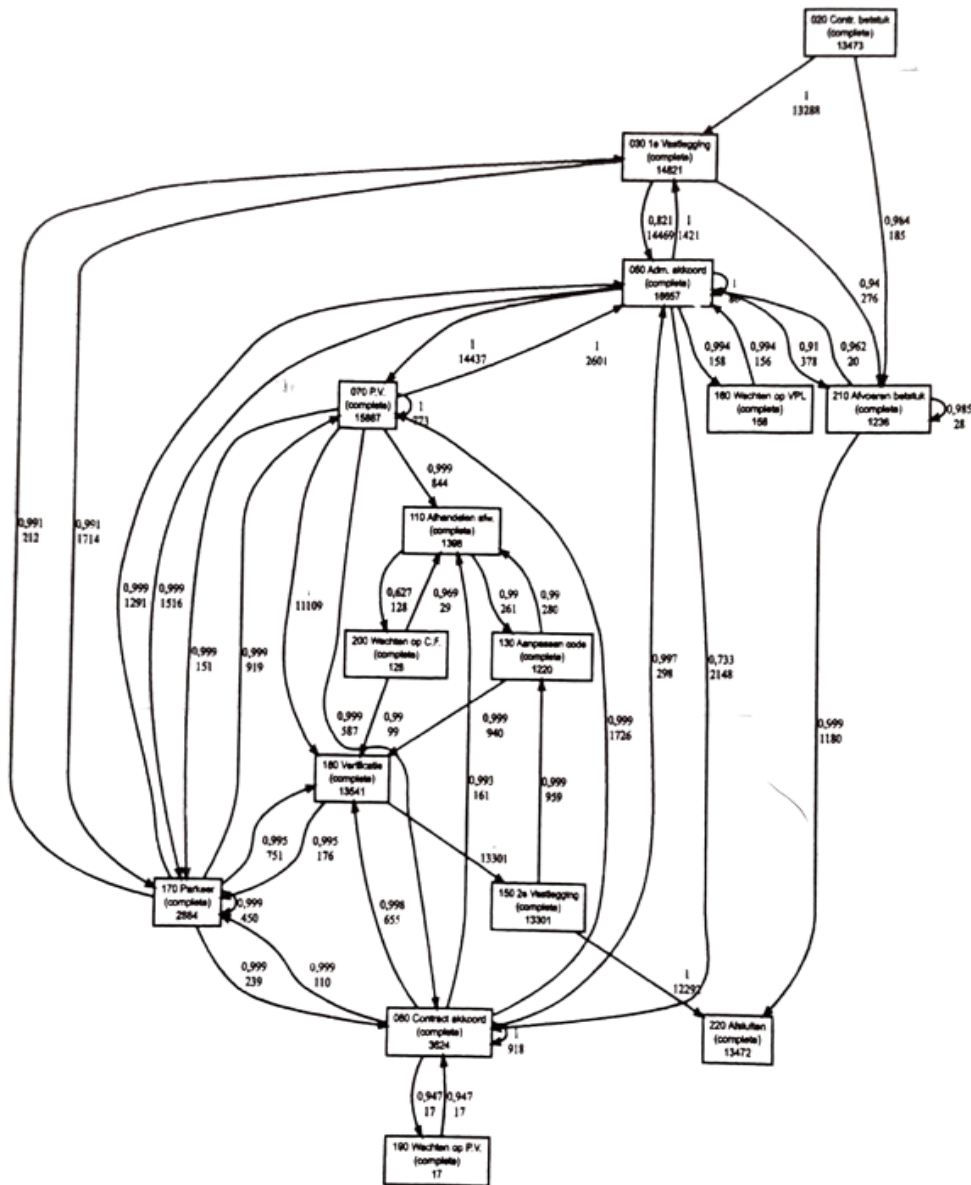
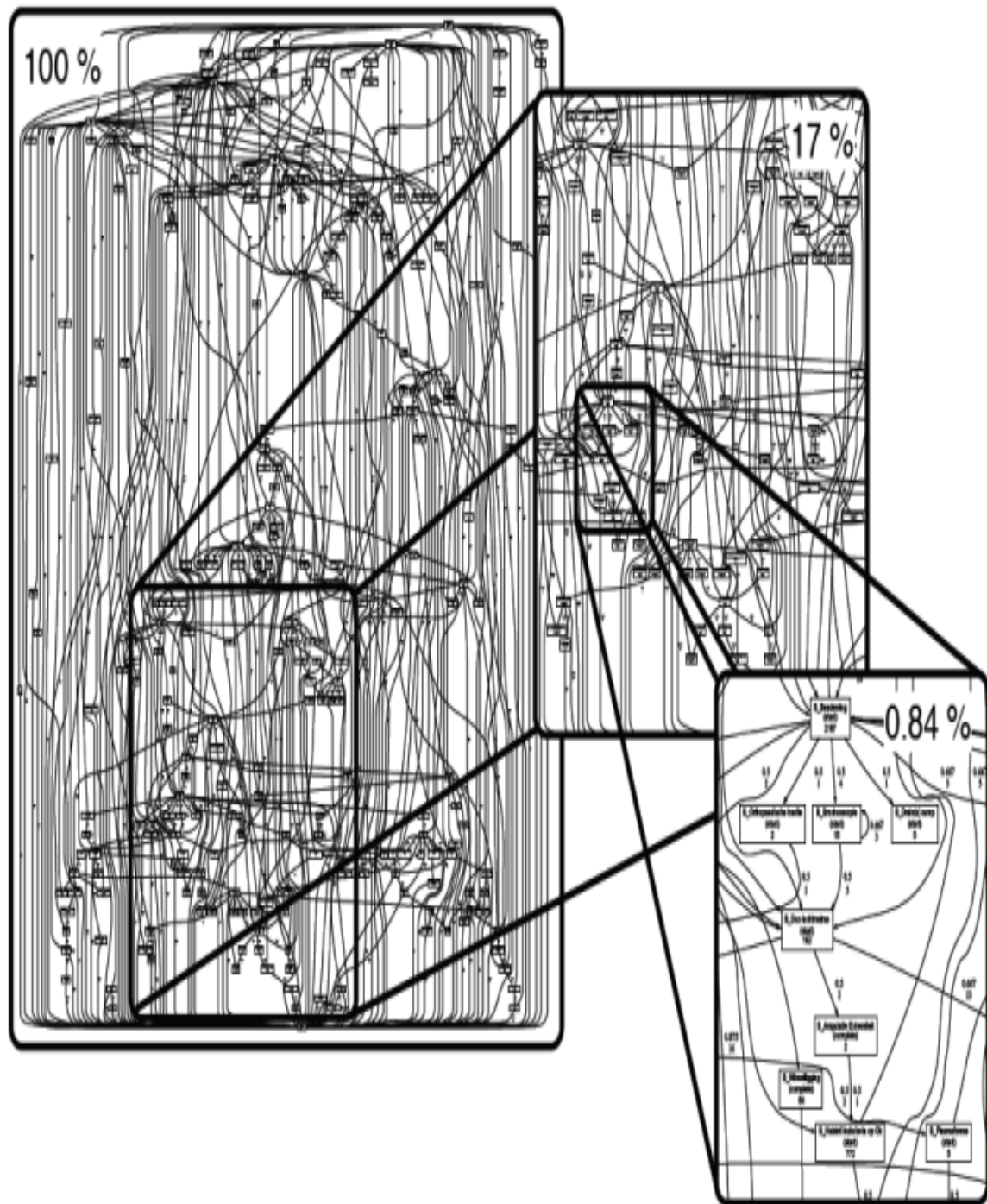


Figura 83. Ejemplo de proceso tipo lasaña.

El proceso Espagueti: no cuentan con un patrón o estructura por lo que solo algunas técnicas y algoritmos de minería de procesos los puede aplicar sobre ellos debido a su complejidad. Los datos que muestran este tipo de procesos incluyen la realización de clústeres o representación de aquellos caminos con una frecuencia de paso superior a la determinada (figura 84).



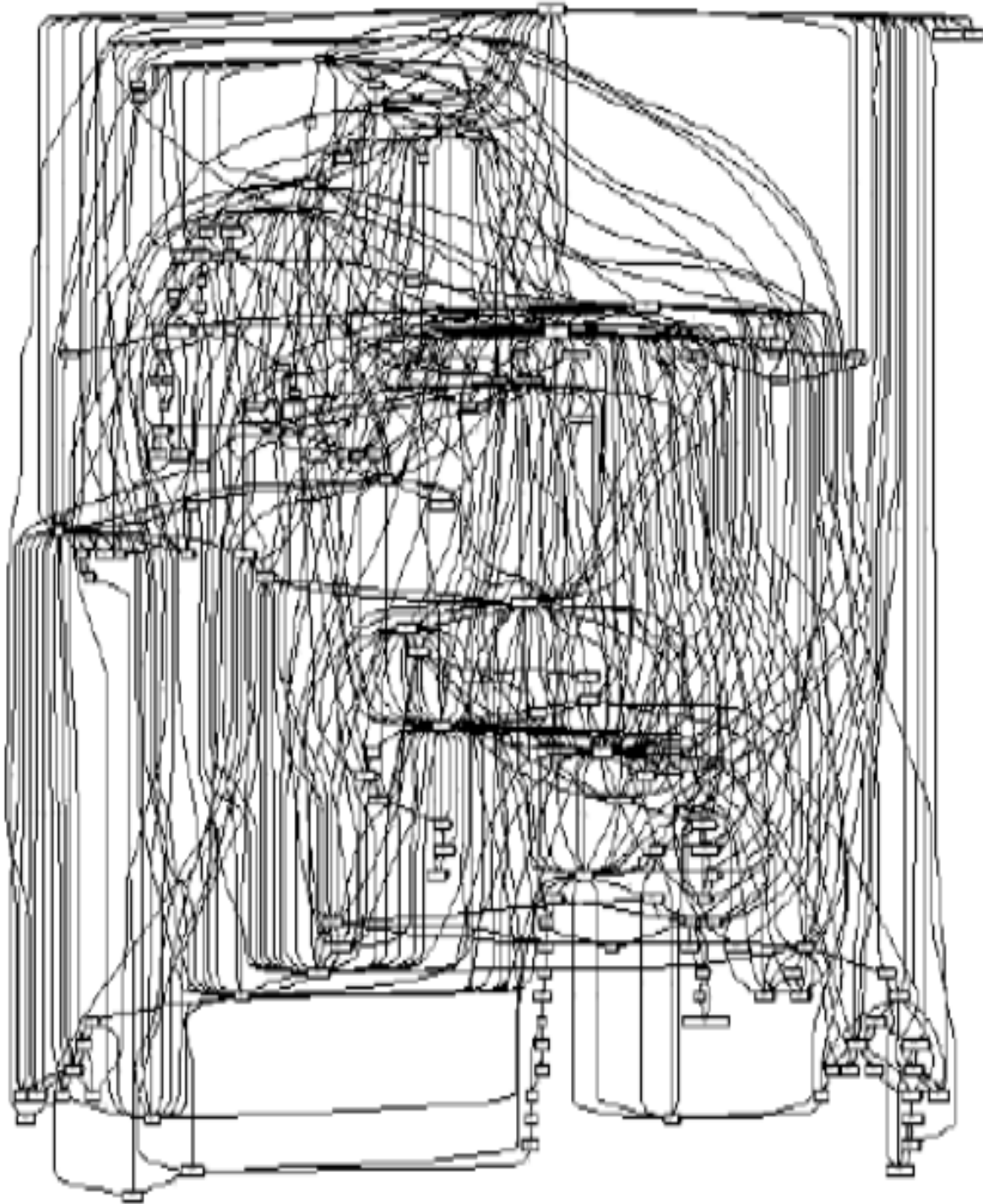


Figura 84. Ejemplo de proceso tipo espagueti.

**CAPÍTULO IV**

**Métodos para el**

**desarrollo de**

**proyectos de**

**minería de procesos**

#### **4. Modelos o métodos utilizados para llevar a cabo la minería de procesos.**

Autores referenciados dentro del campo (Hugo Santiago Aguirre Mayorga, Nicolás Rincón García, Wil van der Aalst, Enzo Edir Velásquez Lobatón, Claudia L. Hernández G\*, María Ximena Dueñas R., Brian Keith Norambuena, Vianca Vega Zepeda) de la minería de procesos proponen algunos modelos o métodos a aplicar con el objetivo de analizar el funcionamiento básico del proceso, información que además puede ser aprovechada en la minería de datos (hallar anomalías, patrones y correlaciones en grandes conjuntos de datos para predecir resultados), en el siguiente apartado se describirán (Aguirre y Rincón, 2015 y SAS, 2025).

Respecto a lo antes mencionado Aguirre y Rincón (2015) indican que se han desarrollado diferentes metodologías que definen una serie de etapas, pasos y lineamientos para aplicar la minería de procesos dentro de una organización, siendo el objetivo de esas metodologías facilitar la aplicación de las técnicas de minería de procesos para un mayor desarrollo de la disciplina en las organizaciones. Sin embargo, los resultados y desafíos para cada caso de aplicación son diferentes y se deben tener en cuenta las necesidades a cubrir para generar los resultados esperados aplicando la metodología que más le convenga a la organización, por lo que en el siguiente apartado se describen algunas metodologías aplicadas a la minería de procesos para su mejor entendimiento.

##### **4.1. Modelo de ciclo de vida L (Life Cycle Model)**

Este modelo es propuesto por Van der Aalst en el año 2011, va enfocado a las actividades y artefactos concretos de la minería de procesos (figura 85) y consta de cinco etapas.

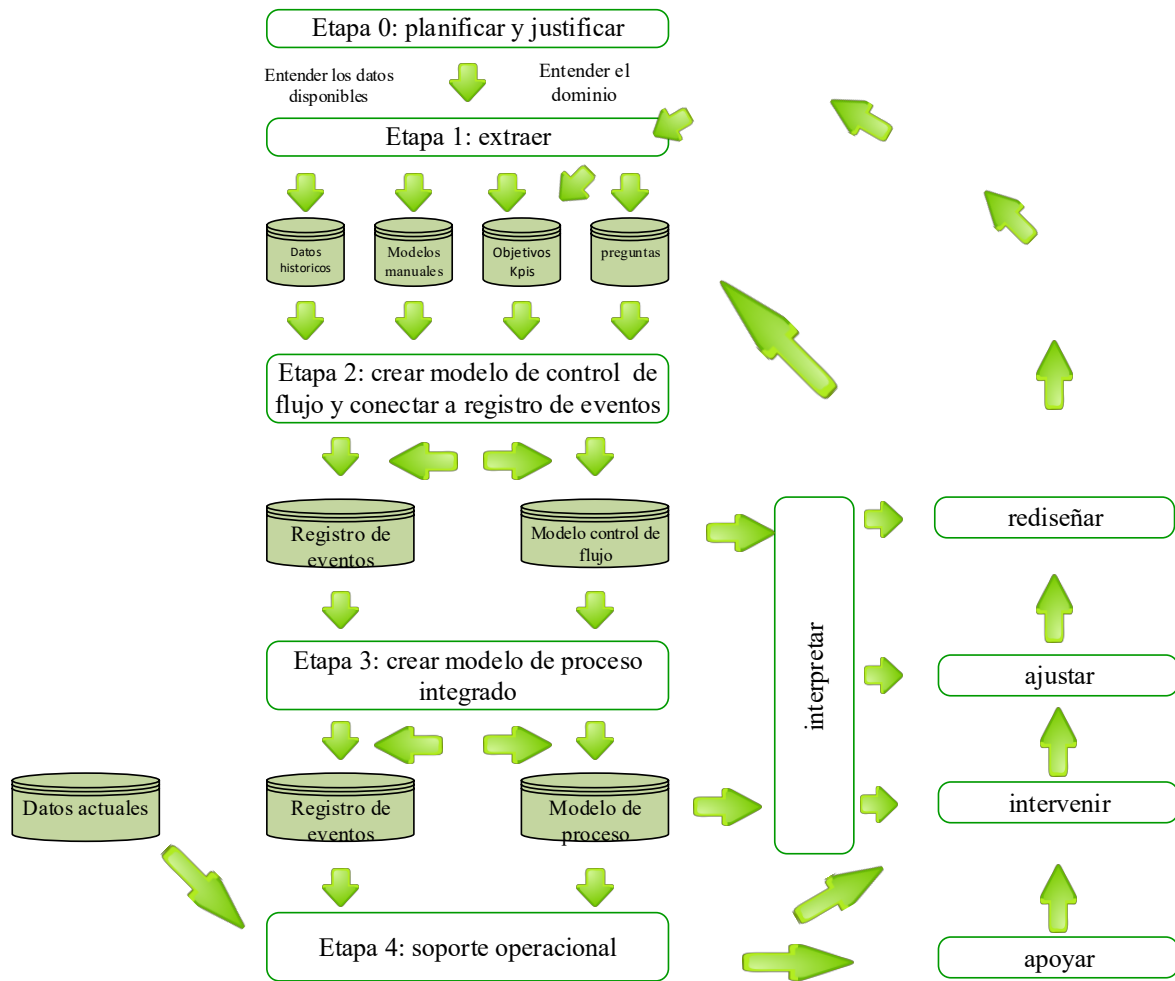


Figura 85. Modelo de ciclo de vida L (Life Cycle Model) propuesta por Van der Aalst, 2011.

**Etapa 0:** planificar y justificar un proyecto.

**Etapa 1:** extraer datos, eventos y objetivos necesarios a partir de los sistemas de información de los expertos del dominio y la gestión.

**Etapa 2:** Se crea un modelo de control de flujo y se relaciona con el registro de eventos; en esta etapa se utilizan técnicas de descubrimiento de procesos que dan como resultado un nuevo modelo.

**Etapa 3:** la relación entre el registro de eventos y el modelo se utiliza para extender el modelo existente o por lo contrario crear un modelo de proceso integrado, estos modelos son utilizados para apoyar las operaciones.

**Etapa 4:** el conocimiento extraído de los datos (de los eventos históricos) es utilizado para intervenir, predecir y hacer recomendaciones en los procesos operativos.

#### 4.2. Método de minería de procesos de Velázquez.

El método de Minería de Procesos propuesto por Velázquez (2017) inicia con la introducción a los procesos de negocio, después con la obtención de los datos, continúa con el procesamiento de los mismos y termina con la presentación de la información (figura 86); este método consta de seis etapas:



Figura 86. Diagrama flujo del Método de minería de procesos propuesta por Velázquez, 2017.

**Etapa 1 Introducción a los procesos de negocios:** esta etapa se utiliza para concientizar a los colaboradores sobre la importancia de los procesos, orientarlos en base al objetivo de la organización; puntos que son necesarios para tener resultados satisfactorios.

**Etapa 2 Obtención de datos:** consiste en conseguir los datos de diferentes fuentes (mensajes, documentos, bases de datos, etc.) para descubrir, monitorear y mejorar el proceso.

Las fuentes de datos se pueden encontrar dispersos por razones técnicas u organizativas y es por eso es necesario agruparlos para tener un punto de partida.

**Etapa 3 Registro de eventos:** en esta etapa se tiene que revisar que todos los eventos estén ordenados y tengan relación referente a un proceso; se tiene que hacer una evaluación de los datos para aprobar o rechazar la información.

**Etapa 4 Proceso de descubrimiento:** va enfocado a la perspectiva de control de flujo y de tiempo del proceso.

**Etapa 5 Análisis de conformidad:** el objetivo de esta etapa es encontrar similitudes y diferencias entre el modelo real que tiene la empresa y el modelo descubierto (creado a partir de los datos).

**Etapa 6 Alineación del negocio y auditoria:** el objetivo de esta etapa es asegurar que los sistemas de información y los procesos de negocio reales estén acorde a los lineamientos establecidos.

#### **4.3. Metodología de Aguirre.**

La metodología de Aguirre (2015) consta de cuatro etapas.

**Etapa 1 Definición del proyecto:** en esta etapa se debe entender el proceso y sus problemas primordiales con la finalidad de establecer los objetivos de mejora o las preguntas a responder en base a la Minería de Procesos.

**Etapa 2 Preparación de los datos:** se deben localizar los datos que son requeridos para el análisis y extraerlos del sistema de información para asegurar la calidad de los mismos y someterlos a análisis con ayuda de las técnicas de Minería de Procesos.

**Etapa 3 Análisis del proceso:** consiste en aplicar las técnicas de Minería de Procesos a los datos preparados para descubrir el modelo real del proceso, analizar su desempeño, la relación que hay entre las personas involucradas e identificar si se cumplen los procedimientos y reglas de la empresa.

**Etapa 4 Rediseño del proceso:** en esta etapa se establecen las posibles alternativas de mejora en base a los hallazgos de la etapa anterior, se debe evaluar la factibilidad para posteriormente implantar las mejoras.

Cada una de las etapas antes mencionadas tienen pasos específicos los cuales se describen en la figura 87.

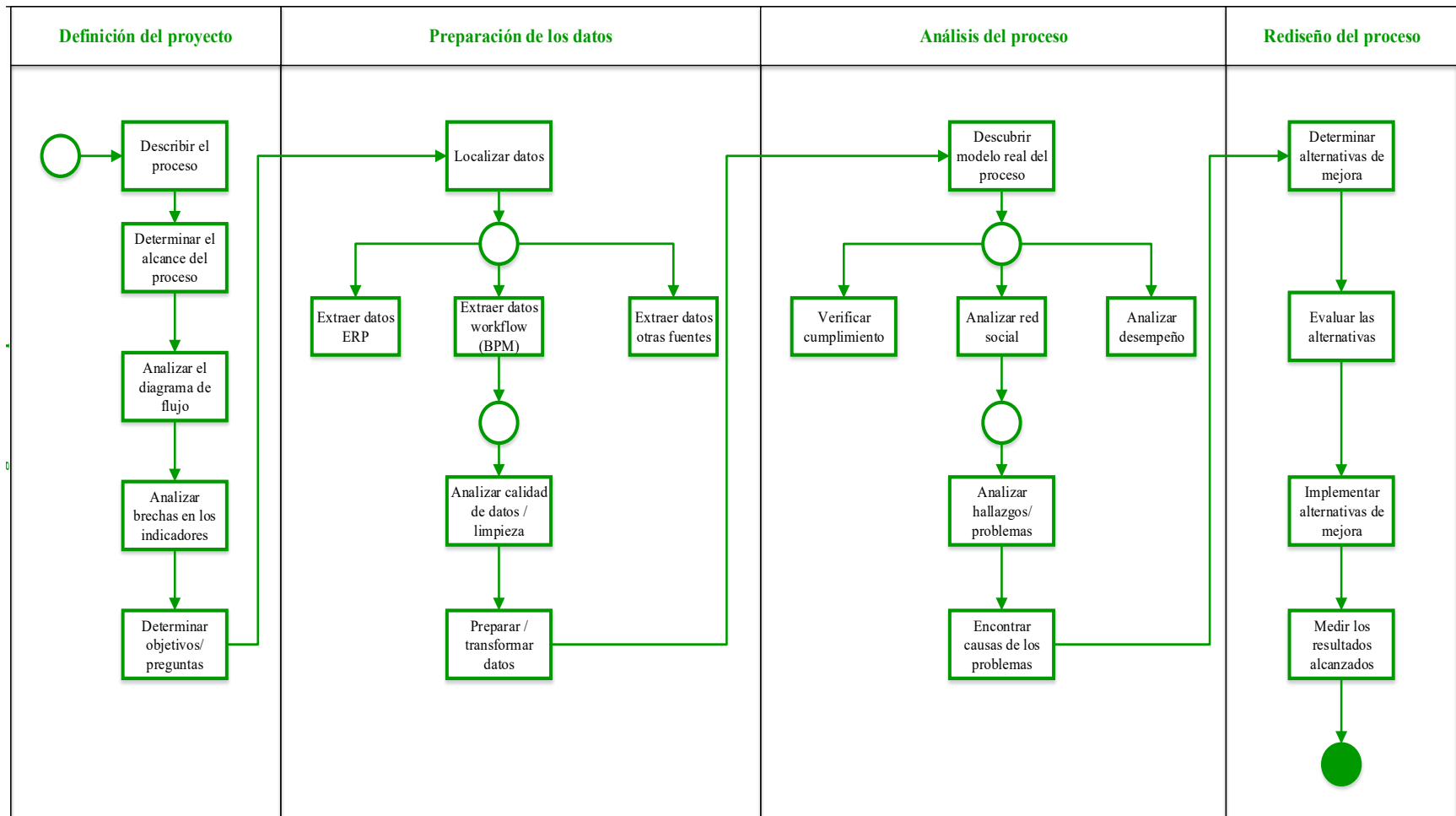


Figura 87. Metodología de Aguirre, 2015.

#### **4.4. Metodología para la aplicación de minería de procesos de Bozcaya.**

Para este caso de estudio se utilizará la metodología de diagnóstico de procesos de Bozcaya (2009) y retomada por Morales et al. (2022) dicha metodología resulta pertinente debido a que tiene un carácter exploratorio y se basa únicamente en los datos disponibles, permitiendo identificar aspectos relevantes con respecto a la ejecución de un proceso real. Consta de seis fases las cuales se describen a continuación.

1. **Preparación del registro:** La etapa inicial de este procedimiento consiste en la preparación de los registros para su análisis, en esta fase, se localizan, recopilan y depuran los datos obtenidos de los sistemas de información, para después transformarlos a un formato compatible con la minería de procesos. Como parte del preprocesamiento, se reconocen componentes fundamentales como las actividades realizadas, los eventos registrados y las marcas de tiempo. Adicionalmente, según sean los requerimientos del análisis, es posible integrar otros campos que son complementarios y brindan información valiosa para la investigación del proceso en cuestión.
2. **Inspección de registros:** En la fase de inspección de los registros, se recaban datos estadísticos del historial de eventos, este procedimiento comprende la detección de los procesos principales, así como la información de las tareas centrales y las tareas de apoyo. Asimismo, se cuantifica la cantidad de casos y roles involucrados en el proceso, el total de eventos registrados y los valores mínimo, máximo y promedio de eventos por caso, también, se determinan cuáles son los eventos de inicio y finalización, en conjunto con su frecuencia. De acuerdo con los resultados estadísticos obtenidos, en ciertas ocasiones es necesario filtrar el registro para excluir aquellos casos que se consideran incompletos.
3. **Control de análisis de flujo:** en esta fase se debe analizar el flujo del proceso para dar respuesta a la pregunta ¿cómo es el proceso real?, con ello, se lleva a cabo una verificación de conformidad, es decir, comprobar si el proceso obtenido del análisis

es igual al documentado por la empresa, donde cada caso del historial de eventos sea reproducible en el procedimiento definido por la organización. Cuando se detectan múltiples casos en el registro que no se alinean con el modelo, es necesario aplicar filtros para omitir los eventos de baja frecuencia durante la construcción del modelo. Esta práctica tiene el propósito de evitar la generación de modelos excesivamente complejos y caóticos, comúnmente conocidos como "modelos espagueti".

4. **Análisis de rendimiento:** los modelos generados en la tercera fase pueden emplearse para evaluar el desempeño del proceso, con el fin de detectar la posible existencia de cuellos de botella. Asimismo, permiten determinar la duración del ciclo de cada actividad de forma individual, así como del proceso en su conjunto. El conocimiento obtenido de este análisis proporciona información valiosa para identificar áreas con potencial de optimización.
  
5. **Análisis de roles:** La fase de análisis de roles consiste en examinar la información disponible sobre los distintos perfiles (sean personas u otros recursos), se identifica el responsable de la ejecución de cada tarea y se evalúa la productividad de los recursos involucrados en el proceso. La evidencia derivada de este análisis proporciona a los gestores una base sólida para tomar decisiones fundamentadas sobre la distribución de la carga laboral al interior del equipo y de la empresa.
  
6. **Transferencia de resultados:** En la fase final de transferencia de resultados, se consolida y redacta un informe con el propósito de comunicar de manera clara y estructurada los hallazgos y conclusiones derivados de la investigación a los tomadores de decisiones, ya sea la alta gerencia o los responsables de la administración de los sistemas y/o en su caso de los responsables de las áreas de la empresa.

## 5. Aplicación de la metodología

### 5.1. Caso de estudio

Para presentar la contextualización del caso con el que se trabajó, se buscó información disponible y pública del proceso de atención de servicios básicos al cliente en Internet, en donde se encontraron documentos relacionados a: la estructura organizacional de CFE, a sus objetivos y funciones; al proceso de la empresa, a los diferentes tipos de servicios; el proceso y procedimiento de atención, seguimiento y terminación de solicitudes; descripción de las solicitudes para los trabajos programados; y los roles de los usuarios para acceso al sistema de solicitudes de servicio la cual se describe a continuación.

#### 5.1.1. Definición del caso

La Comisión Federal de Electricidad (CFE) es una empresa pública, de carácter social que el gobierno federal creó el 14 de agosto de 1937, tiene por objetivo organizar y dirigir un sistema nacional de generación, transmisión y distribución de servicios de suministro de energía eléctrica al público; CFE maneja cuatro procesos: el primero es la generación, el segundo la transmisión, el tercero la distribución y el cuarto proceso y último es el suministro de servicios básicos de energía eléctrica (figura 89).

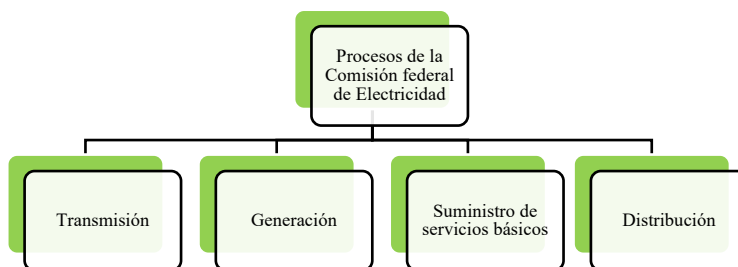


Figura 88. Procesos de CFE.

La Comisión Federal de Electricidad trabaja con ocho sistemas de información (figura 90), cada uno de los sistemas lleva a cabo diferentes procesos y por ende diferentes tipos de solicitud que son atendidos de acuerdo con las necesidades de los usuarios.

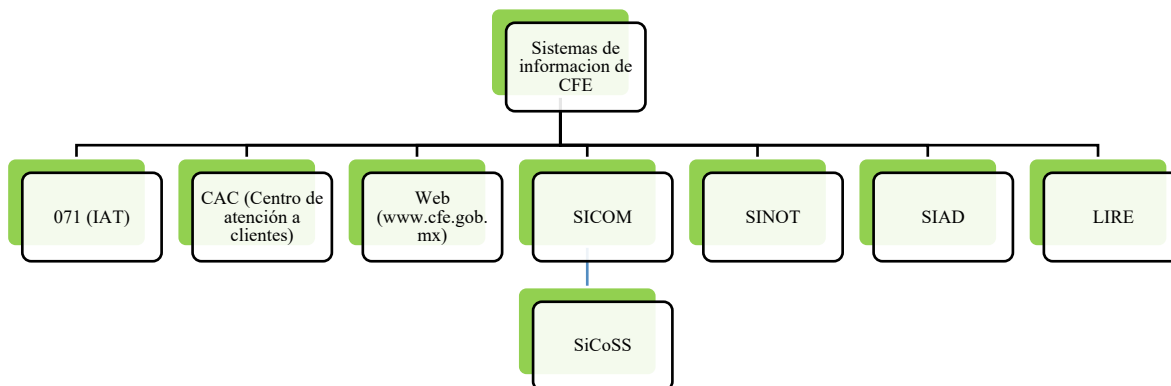


Figura 89. Sistemas para el registro de solicitudes.

Para realizar el presente proyecto se obtuvieron los datos del proceso de registro y atención de solicitudes de suministro de servicios básicos al cliente, este proceso tiene como objetivo alcanzar la satisfacción del usuario final en los servicios administrativos y de campo relacionados con el suministro eléctrico, por lo que se requiere un trabajo coordinado de las actividades que intervienen en la atención, desde el primer contacto con el usuario final hasta la atención de la solicitud del servicio requerido.

El sistema SICOM es la herramienta por medio de la cual se crean y administran usuarios (operador, agente, supervisor, oficinista, zonal, zonal divisional) con ciertos privilegios de acceso para poder ingresar al sistema SiCoSS (Sistema de Control de Solicitudes de Servicio) para atender y registrar las solicitudes del suministro de servicios básicos como lo es: la carga de lecturas, facturación de ciclos, emisión de cortes, reconexiones y desconexiones, altas, bajas y cambios de medidor, además de realizar modificaciones a la base de datos a nivel usuario, realizar la cobranza de facturación, de pagos de reconexiones, de focos, la supervisión de obras y contrataciones; cabe mencionar que este sistema es el vínculo con el cliente y el único mecanismo para el registro y atención de todas las solicitudes y quejas de los usuarios finales.

A través del SiCoSS se asegura la calidad de las actividades de registro y atención de solicitudes para aminorar las causas que obstaculizan una atención pertinente y oportuna a los requerimientos de los usuarios.

Las solicitudes que se registran en el sistema SiCoSS son: Contratación (C), Emergencias (E), Quejas (Q), Servicios (S) y Trabajos programados (T) (figura 91).

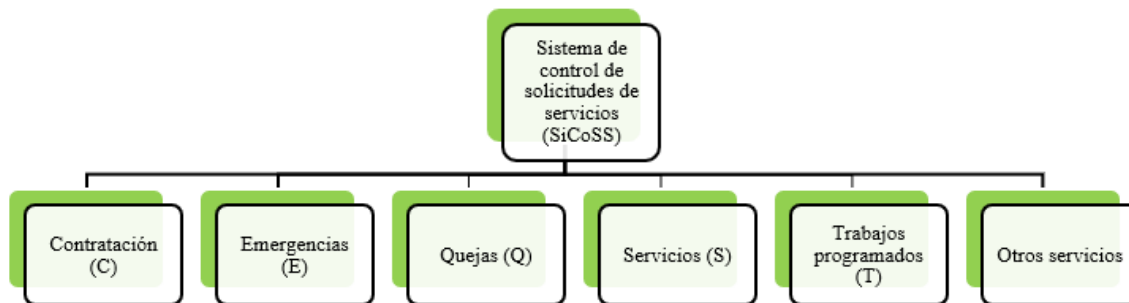


Figura 90. Tipos de solicitudes.

Para el presente proyecto de investigación se abordará solo con los trabajos programados, ya que son los datos que fueron proporcionados por un integrante del personal involucrado en el manejo del sistema de control de solicitudes.

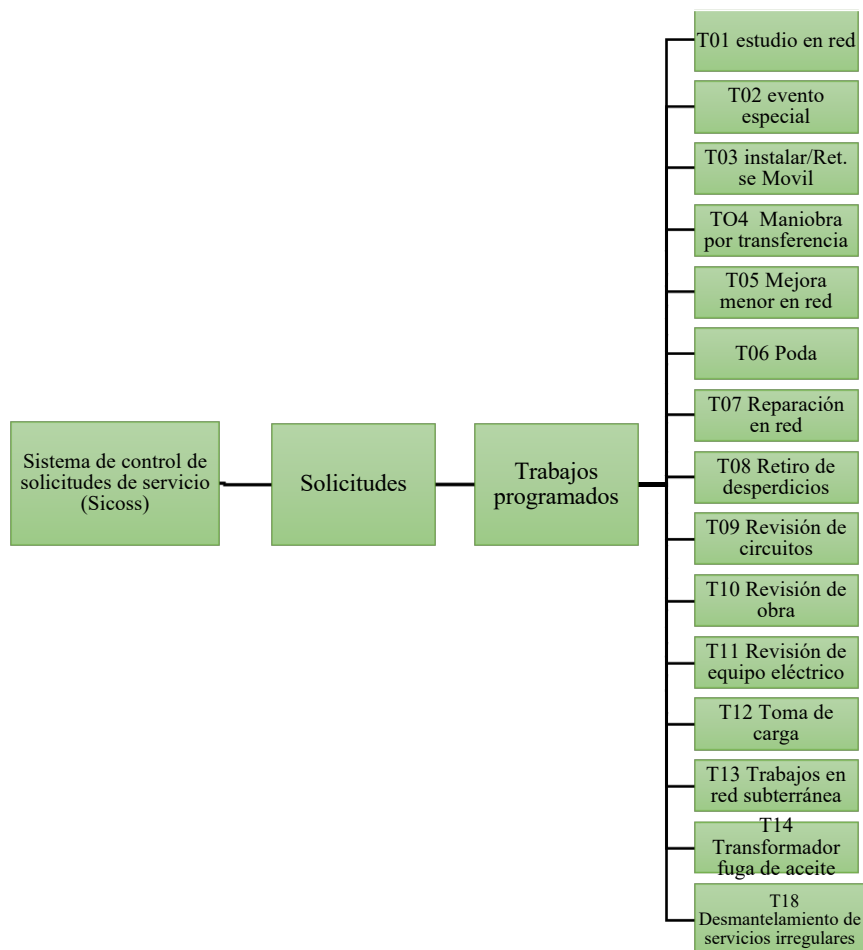


Figura 91. División de los trabajos programados.

Cada una de las solicitudes es autónoma, es decir, cada una de ellas cuenta con características y criterios particulares (Tabla 8) y de ello depende el tipo de centro que realiza la atención, por lo cual el flujo de la información es independiente entre cada una de ellas, además, se debe considerar que para los trabajos programados no hay cancelados, rechazados e impropcedentes, es decir, siempre son atendidas todas las solicitudes y no existe tiempo definido para su atención.

Tabla 7. Descripción de los trabajos programados.

Solicitud	Descripción	Criterios de Registro	Captura	Atención	Productividad (créditos)
<b>T01 Estudio en red.</b>	Estudio para mejorar la calidad del servicio.	Esta solicitud es solo para la ejecución de levantamientos de trabajos programados del área de distribución, CCC como mejoras a la red o acometidas.	CCC	CCC Planeación Distribución	20
<b>T02 Evento especial</b>	Establecimiento de una guardia operativa de una cuadrilla o pareja de lineros en algún evento importante como gira presidencial, elecciones, concentraciones sociales, informes de gobierno, atención de áreas con siniestro no eléctrico.	Son solo para la ejecución de trabajos programados en el área de distribución y CCC indicando el trabajo a realizar.	CCC	CCC Planeación Distribución	15
<b>T03 Instalar/ Ret. Se Móvil</b>	Instalación o retiro de una subestación móvil.	Sustitución de transformador de la red por daño o avería, para darle mantenimiento o cuando se requiera de incremento de capacidad temporal.	CCC	CCC Planeación Distribución	30
<b>TO4 Maniobra por transferencia</b>	Apertura y cierre de equipo de seccionamiento.	Maniobra en equipos de distribución para	CCC	CCC Distribución	18

		el restablecimiento del servicio o mantenimiento de subestaciones o líneas.			
<b>T05 Mejora menor a red</b>	Mejoras en la red de distribución, reubicación, cambio de calibres, instalación de caja antifraude, protección de líneas.	Reubicación de alguna retenida o aislamiento de red por construcción sin afectación al suministro o de usos ilícitos.	CCC	CCC Distribución	30
<b>T06 Poda</b>	Poda de árboles en redes de distribución por programa o a solicitud del usuario final.	Cortar ramas que afectan los bancos de la red de distribución en baja y media tensión.	CCC CAR CAC	CCC	12
<b>T07 Reparación en red</b>	Reparación programada a las redes de media tensión como reemplazo de aisladores, cortacircuitos, fusibles, crucetas, retenidas, postes, separadores, etc.	Reparación de área de Distribución y CCC. En el área de instrucciones deberá indicar el trabajo a realizar. Se utilizará cuando no halla afectación a usuarios o se hallan restablecido provisionalmente.	CCC	Distribución	30
<b>T08 Retiro de desperdicios</b>	Retiro de materiales y/o desechos de trabajo por distribución tales como ramas, escombros, postes, etc.	Por reporte de usuario y/o área de distribución.	CCC CAC CAR Distribución	CCC Distribución	4
<b>T09 Revisión de circuitos</b>	Solicitud para patrullaje programado de circuito.	Cuando es necesaria la revisión de un circuito o ramal en caso de algún evento transitorio en el mismo.	CCC	CCC Distribución	18
<b>T10 Revisión de</b>	Realizar revisión		CCC	Medición	28

<b>obra</b>	de obras de infraestructura eléctrica para su recepción, obras en instalaciones y/o infraestructuras en media tensión, transformadores, acometidas, redes de baja tensión y/o preparaciones domiciliarias.			Planeación	
<b>T11 Revisión de equipo eléctrico</b>	Revisión de equipo eléctrico: transformadores, capacitores, reguladores, cuchillas de operación y unipolar.	Revisión de algún equipo del circuito de distribución, derivado de una falla provocada por estos equipos o después de alguna transferencia de carga o apertura de alguna cuchilla en operación.	CCC	Se realiza la revisión minuciosa del equipo, para determinar la falla y en su caso hacer el reporte detallado para su envío a distribución.	10
<b>T12 Toma de carga</b>	Medir la tensión e intensidad de la corriente de un transformador o punto de la red. Cuando se requiere de conocer la carga que lleva alguna instalación.	Se relaciona con las solicitudes de contratación de más de 5kw.	CCC	CCC Distribución Medición	10
<b>T13 Trabajos en red subterránea</b>	Trabajos de mantenimiento programados en la red subterránea sea en media o baja tensión, como limpieza de transformadores, pedestales, limpieza de registros tipo pozo, conectadores, tapas de registro		CCC	CCC Distribución	30

	y/o cualquier otro trabajo donde no exista afectación al suministro eléctrico.					
<b>T14</b> <b>Transformador fuga de aceite</b>	Atención en campo de reportes del usuario final, ciudadanos y/o trabajadores referentes a la fuga de aceite de transformadores.	Al recibir un reporte de un transformador que está tirando aceite, o cuando el personal por este motivo requiere cambiar transformadores de distribución por fuga de aceite por corrosión en el tanque, no habiendo afectación en el suministro.	CAR CAC CCC	Distribución		10
<b>T18</b> <b>Desmantelamiento de servicios irregulares</b>	Desmantelar instalaciones abandonadas o en asentamientos irregulares sin la emisión de notificaciones de ajuste.	Programa de ciclo limpio.	CCC	CCC Medición		100

**Acometida:** Se llama acometida en las instalaciones eléctricas a la derivación desde la red de distribución de la empresa suministradora hacia la protección principal o medidor de la energía de la edificación o propiedad donde se hará uso de la energía eléctrica.

**CAC:** Centro de Atención a Clientes

**CAR:** Centro de Atención Regional

**CCC:** Centro de Continuidad y Conexiones

**KW:** Kilowatt

Para el correcto registro y termino de cada solicitud de trabajos programados dentro del sistema de información es fundamental que se especifique correctamente la fecha y hora de terminación, los trabajos realizados, los materiales utilizados, los medidores, los sellos instalados y la información necesaria para complementar la atención de la solicitud.

### 5.1.2. Proceso de atención de los trabajos programados

El proceso general documentado por CFE de cualquier solicitud está conformado por siete actividades (figura 81):

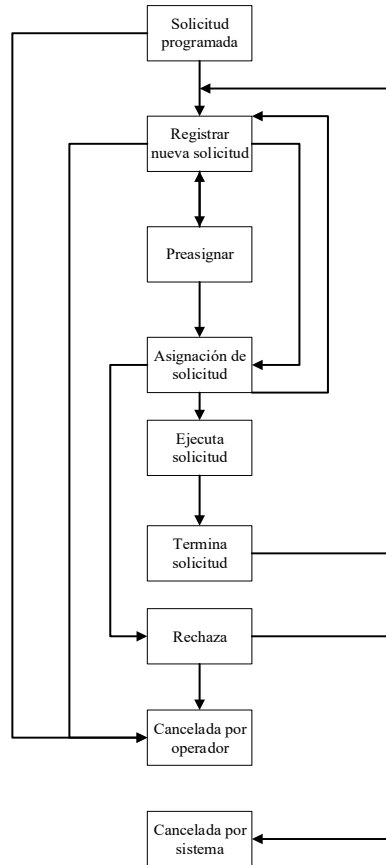


Figura 92. Proceso de una solicitud.

- 1.- Registrar una nueva solicitud de servicio, contrato, emergencia, trabajo programado, etc.
- 2.- Preasignar la solicitud para su atención.
- 3.- Asignar la solicitud a una cuadrilla para su atención.
- 4.- Ejecutar el trabajo asignado.
- 5.- Terminar la solicitud cuando fue atendida en campo.
- 6.- Rechazar la solicitud cuando se visitó el domicilio de la orden pero que por alguna causa no se pudo realizar o concluir el trabajo.
- 7.- Cancelar la solicitud cuando los datos son incorrectos.

## 5.2. Etapa 1 “Preparación del registro”.

Para crear el registro de eventos se tomaron los datos de las solicitudes de servicios básicos de atención al cliente que fueron proporcionados por parte de CFE (5119 registros de eventos, 1091 casos), los datos se encuentran en formato de texto (.doc) y constan alrededor de 37 campos divididos en siete secciones: información general, información del servicio, observaciones de registro, bitácora de movimientos, trabajos realizados, observaciones y bitácora de servicios (figura 94).

Solicitud de Suministro de servicios básicos al cliente de CFE

### INFORMACIÓN GENERAL

Solicitud	M2733017720	Situación	TERMINADA
Registrada como	T07 - REPARACION EN RED	Atendida como	T07 - REPARACION EN RED
Origen	CCC - CENTRO DE CONTINUIDAD Y CON	Centro	M27DJ - DISTRIB. CHAPINGO DM27

### INFORMACIÓN DEL SERVICIO

Nombre	DORANTES SANCHEZ ELEUTERIO	Teléfonos	0000000000-0000000000	Tarifa	01
Dirección	ATLAUTENCO #1 CP.56070	R.P.U.	513120200248	Hilos	1
Entre Calles	COLORINES			Visitas	1
Referencia		R.M.U.	XAXX010101	SSEBRA	S/N
Colonia	M27IM7 - BARRIO LA ASUNCION VMC	Código Postal	56070	Cuenta	31DM27A057310740
Población					
Municipio	TEPETLAOXTOC				
Evidencias	<a href="#">Click aquí para ver Fotos y Georeferencias</a>				

### OBSERVACIONES DE REGISTRO

SE ENCUENTRA CON APARTARRAYO VOLADO CENTRO EN TR DE 45 KVA

### BITACORA DE MOVIMIENTOS

Movimiento	Fecha	Hora	R.P.E.	Nombre	Login	Observaciones
CAPTURO	30-MAR-2019	10:11	9L7XN	MARQUEZ SANCHEZ CESAR	scd9l7xn	
PREASIGNO	31-ENE-2020	12:43	G274F	ALONSO HERNANDEZ LUIS	scdg274f	Cuad:M2745
ASIGNO	31-ENE-2020	12:47	G274F	ALONSO HERNANDEZ LUIS	scdg274f	
EJECUTO	31-ENE-2020	12:48	9L321	CEH CHI PAUL DAMIAN		CUADRILLA: M2745 VISITA: 1 NORMALIZO
TERMINO	31-ENE-2020	12:48	G274F	ALONSO HERNANDEZ LUIS	scdg274f	

### TRABAJOS REALIZADOS

Acción	Equipo	Ubicación	Tipo
--------	--------	-----------	------

### OBSERVACIONES

APARTARRAYOS REMPLAZADOS CON ANTERIORIDAD SERVICIO OK

## BITACORA DE SERVICIOS

#	Fecha	Servicio	Atendido por	Solicitud
1	17-ABR-2015	SCJ - CORTE BT MT X TERCE	J712J - MARQUEZ MARTINEZ ANDRE	M2730702548
2	21-ABR-2015	SC7 - RECONEXION MEDIDOR	J712J - MARQUEZ MARTINEZ ANDRE	M2730823899
3	27-MAR-2016	Q01 - NO LUZ	CANCELADA POR OPERADOR	M2732242268
4	30-ENE-2018	Q01 - NO LUZ	CANCELADA POR SISTEMA	M2732727134
5	31-ENE-2018	S08 - MEJORA DE ACOMETIDA	9K47V - URBINA GOMEZ CRISTOPHE	M2732728450
6	30-MAR-2019	E03 - SECTOR FUERA	9N0E7 - HUERTA AGUILAR JOSE LU	M2733017658

Figura 93. Solicitud de suministro de servicios básicos al cliente de CFE.

De las solicitudes proporcionadas solo se extrajeron catorce campos de cuatro secciones diferentes (tabla 9), campos que son importantes para identificar cada una de las solicitudes (id), conocer el tipo de solicitud que se trabaja (T01, T02, T03, T04 ... T18), las actividades que se realizan en cada solicitud y quien las ejecuta, cuanto tiempo tardan en concretar o finalizar una solicitud, que centro es el que atiende la solicitud (CAC: Centro de Atención a Clientes, CAR: Centro de Atención Regional, CCC: Centro de Continuidad y Conexiones), lo cual al realizar el análisis ayuda a conocer el proceso real, conocer cuáles son las actividades que tardan más en realizarse, identificar cuellos de botella o en su caso actividades repetitivas, averiguar qué personal es el que tiene más carga de trabajo o en su caso está más involucrado..

Tabla 8. Datos seleccionados para conformar el registro de eventos.

Registro de eventos	
Campos	Secciones
<b>Información general</b>	Solicitud
	Registrada como
	Atendida como
	Origen
<b>Observaciones de registro</b>	Centro
	Observaciones de registro
	Observaciones
<b>Bitácora de movimientos</b>	Movimientos
	Fecha
	Hora
	Observaciones
<b>Observaciones</b>	Nombre del personal que realizo la actividad
	Observaciones

Posteriormente, los datos de la tabla 9 se transcribieron de manera manual (tal cual están en la solicitud, sin modificación alguna, sin corrección ortográfica ni de ningún tipo) en una hoja de cálculo de Microsoft Excel para después guardarlos en formato utf8 Unicode (formato de codificación que permite representar cualquier carácter, incluyendo caracteres especiales, acentos, símbolos) (tabla 10).

Tabla 9. Extracto de datos obtenidos de las solicitudes.

ORDEN	TIEMPO DE INICIO	ACTIVIDAD	RESPONSABLE	CUADRILLAS	REGISTRADA COMO	ATENDIDA COMO	ORIGEN	CENTRO	OBSERVACIONES DE REGISTRO	OBSERVACIONES	CREDITOS
M2733017720	30/03/2019 10:11	CAPTUREO	MARQUEZ SANCHEZ CESAR		T07 - REPARACION EN RED	T07 - REPARACION EN RED	CCC - CENTRO DE CONTINUIDAD Y CON	M27DJ - DISTRIB. CHAPINGO DM27	AREA PRESENTA RED DE BAJA TENSION TIENE CABLEDO DE 1+1 NO. 6 SE REQUIERE REEMPLAZO FAVOR DE VERIFICAR SERVICIO SON 40 MTS	SE REQUIERE REEMPLAZO DE DOS CABLES DE BT POR CABLE MULTIPLE. SE PROGRAMA ATENCION	30
M2733017720	31/01/2020 12:43	PREASIGNO	ALONSO HERNANDEZ LUIS	Cuad:M2747							30
M2733017720	31/01/2020 12:47	ASIGNO	ALONSO HERNANDEZ LUIS								30
M2733017720	31/01/2020 12:48	EJECUTO	CEH CHI PAUL DAMIAN	CUADRILLA: M2747 VISITA: 1 NORMALIZO							30
M2733017720	31/01/2020 12:48	TERMINO	ALONSO HERNANDEZ LUIS								30
M2733019379	03/04/2019 09:53	CAPTUREO	PEREZ OROZCO MONICA		T07 - REPARACION EN RED	T07 - REPARACION EN RED	CCC - CENTRO DE CONTINUIDAD Y CON	M27DJ - DISTRIB. CHAPINGO DM27	SE ENCUENTRA CON APARTARRAYO VOLADO CENTRO EN TR DE 45 KVA	APARTARRAYOS REPLAZADOS CON ANTERIORIDAD SERVICIO OK	30
M2733019379	03/04/2019 09:53	PREASIGNO	PEREZ OROZCO MONICA	Cuad:M2745							30
M2733019379	11/09/2019 06:27	CAMBIO DE SITUACION	PEREZ OROZCO MONICA								30
M2733019379	09/01/2020 23:25	PREASIGNO	PEREZ OROZCO MONICA								30
M2733019379	09/01/2020 23:30	ASIGNO	PEREZ OROZCO MONICA								30
M2733019379	09/01/2020 23:30	EJECUTO	INTERIAN COCOM WILBERT	CUADRILLA: M2745 VISITA: 1 NORMALIZO							30
M2733019379	09/01/2020 23:30	TERMINO	PEREZ OROZCO MONICA								30

Cabe resaltar que el archivo cuenta con los tres elementos necesarios (id del caso, nombre de la actividad y el tiempo de ejecución) además de otros datos adicionales que se consideraron relevantes en el proceso para realizar el análisis de minería de procesos.

### 5.3. Etapa 2 “Inspección de registros”

En esta etapa se recolectaron las estadísticas de los registros, los cuales abarcan un periodo de dos años de la fecha 30/03/2019 10:11 al 09/11/2020 23:31 con un total de 990 casos, 5119 eventos y 112 personas involucradas dentro del proceso.

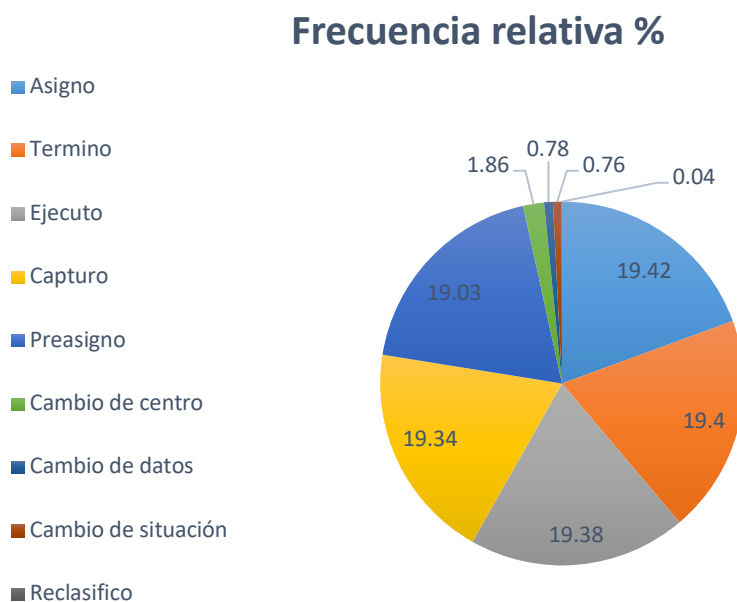
Además, se descubrieron 9 actividades diferentes de las cuales las primeras cinco actividades son las que mayor presencia tienen en el proceso: en primer lugar se tiene la actividad: “Asigno” que es la que mayor frecuencia tiene dentro del proceso con 994, en

segundo lugar: “Termino” cuenta con una frecuencia de 993, en tercer lugar “Ejecuto” con 992, en cuarto lugar la actividad “Capturo” con 990 y por ultimo “Preasigno” con 974 veces involucradas en el proceso, por el contrario el resto de las actividades (cambio de centro, cambio de datos, cambio de situación y reclasifico) son las que menor presencia tienen dentro del proceso (tabla 11).

Tabla 10. Actividades descubiertas en el proceso.

Actividad	Eventos por actividad	Frecuencia relativa %
<b>Asigno</b>	994	19.42
<b>Termino</b>	993	19.4
<b>Ejecuto</b>	992	19.38
<b>Capturo</b>	990	19.34
<b>Preasigno</b>	974	19.03
<b>Cambio de centro</b>	95	1.86
<b>Cambio de datos</b>	40	0.78
<b>Cambio de situación</b>	39	0.76
<b>Reclasifico</b>	2	0.04

En la gráfica 1 se observa el porcentaje de la frecuencia con la que se presentan cada una de las actividades que se descubrieron en el proceso (tabla 11).



### 5.4. Etapa 3 “Control de análisis de flujo”

En esta etapa se aplicaron los diferentes algoritmos mediante el uso de las herramientas informáticas para la minería de procesos para la obtención de los siguientes modelos:

El modelo Red Petri (figura 95) se obtiene a partir del algoritmo “Alpha”, este modelo está conformado por un inicio, transiciones, lugares y arcos, estos últimos se encuentran entre cada una de las actividades que están presentes en el flujo del proceso y la finalización del mismo.

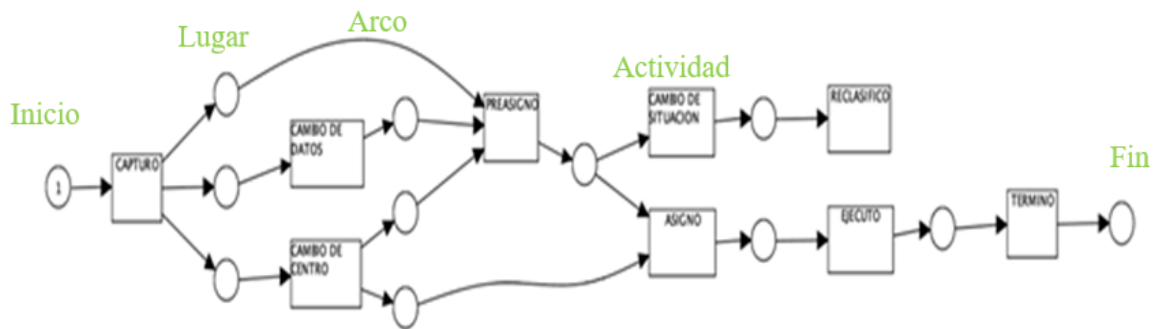


Figura 94. Modelo red Petri del proceso.

El algoritmo fuzzy miner da como resultado un modelo EPC (figura 96), este modelo da a conocer las actividades que están involucradas, muestra el flujo del proceso y sus conexiones (arcos), además, emplea colores para diferenciar las actividades con menor frecuencia y agruparlas; trabaja con la dependencia entre cada una de las actividades.

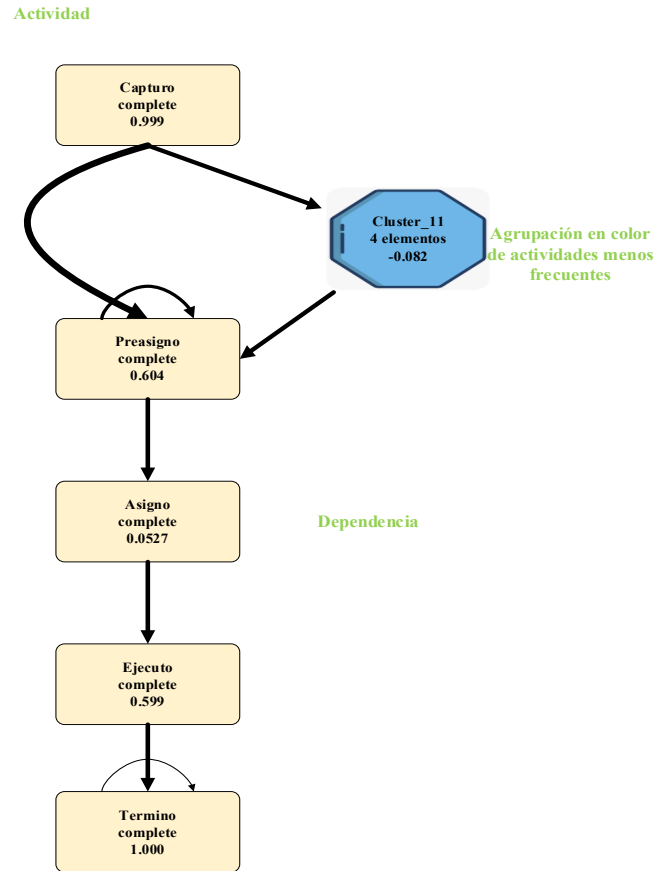


Figura 95. Modelo EPC

El modelo BPMN se obtiene a partir del algoritmo minado inductivo (figura 97), este modelo tiene un inicio, ejecuta las actividades utilizando las compuertas lógicas y muestra como es el flujo del proceso y donde finaliza; a diferencia del modelo fuzzy miner este modelo enmarca las actividades con menor frecuencia.

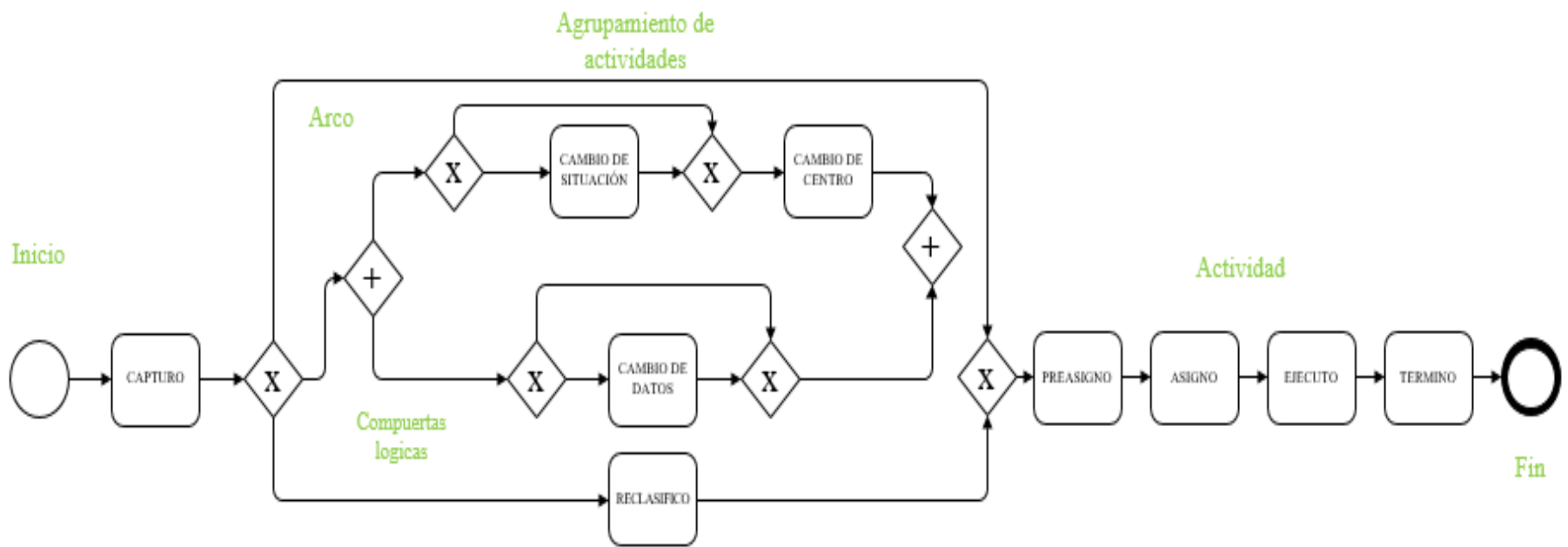


Figura 96. Modelo BPMN.

El algoritmo minado heurístico trabaja con las frecuencias de los eventos y las secuencias de los mismos, puede eliminar las actividades poco frecuentes y da como resultado un modelo red causal (figura 98), está formado por los arcos que son los que unen las actividades, para mostrar el flujo del proceso indicando la frecuencia y la dependencia que hay entre ellas, además muestra aquellas actividades que se repiten, haciéndolo de él un modelo sólido.

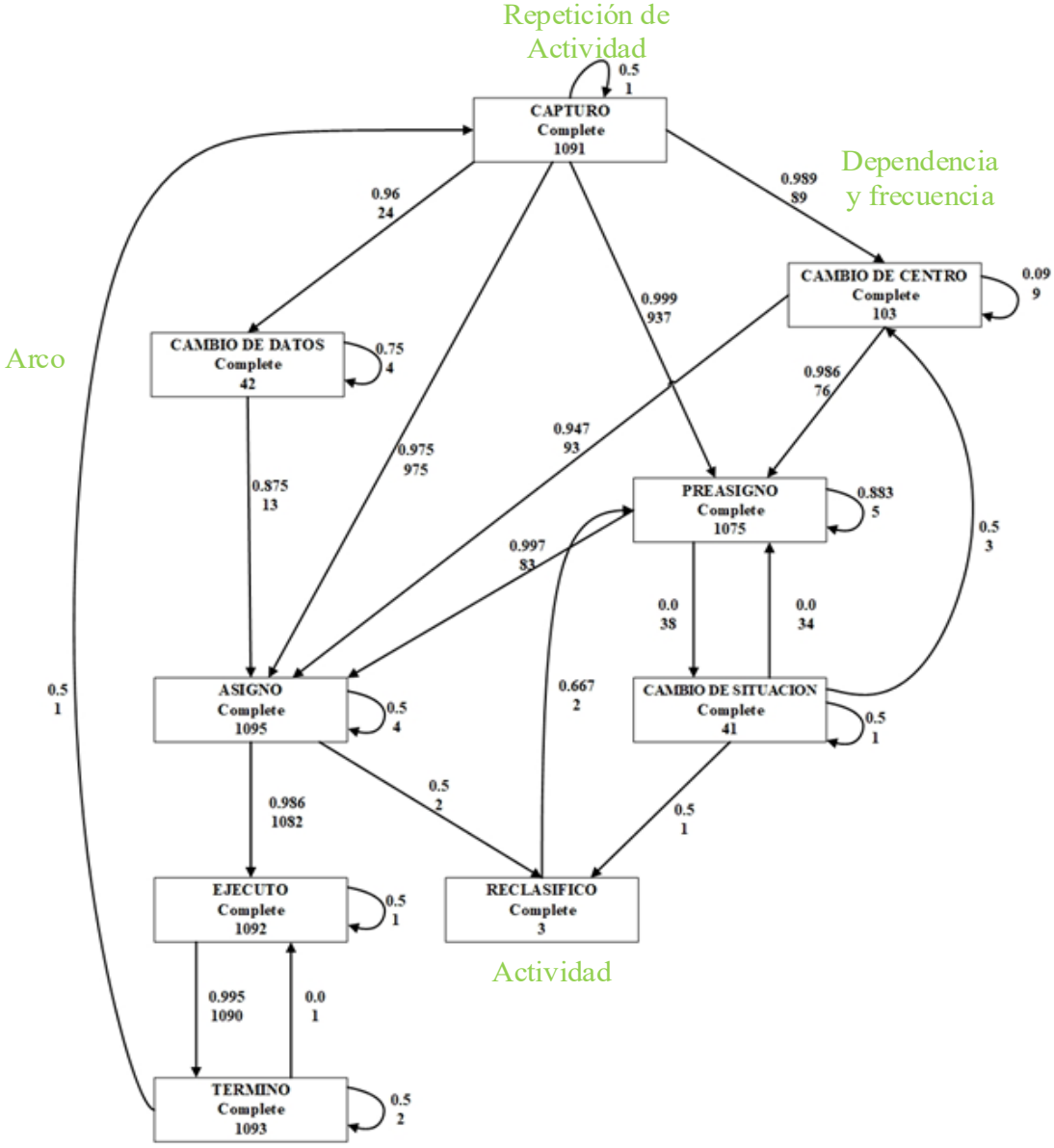


Figura 97. Modelo red causal.

Como se puede observar en los diferentes modelos antes presentados (figura 95, 96,97 y 98) el modelo red causal es el más robusto debido a que muestra el flujo del proceso indicando las frecuencias y la dependencia que hay entre las actividades, la secuencia de las mismas, aquellas actividades que se repiten, las que son poco frecuentes, las cuales pueden ser eliminadas para obtener un modelo de procesos más sólido respecto al de los demás modelos y por tanto se utilizará para los siguientes análisis debido a que proporciona mayor información.

Para poder conocer el flujo completo del proceso se realizó el análisis de minería de procesos aplicando el algoritmo minado heurístico (figura 99), se trabajó con los umbrales bajos para mostrar el modelo completo utilizando los siguientes parámetros:

- umbral de dependencia= 0.60
- umbral de observaciones positivas= 1
- umbral de bucle de longitud uno= 0
- umbral de bucle de longitud dos=0
- umbral de larga distancia= 0;

Como resultado en el modelo se descubrieron nueve actividades de las cuales: cambio de datos, cambio de centro, cambio de situación y reclasifico fueron integradas por alguna necesidad o motivos extraordinarios y por tal motivo no se encuentran documentadas y se tiene una desviación con el proceso real de la empresa.

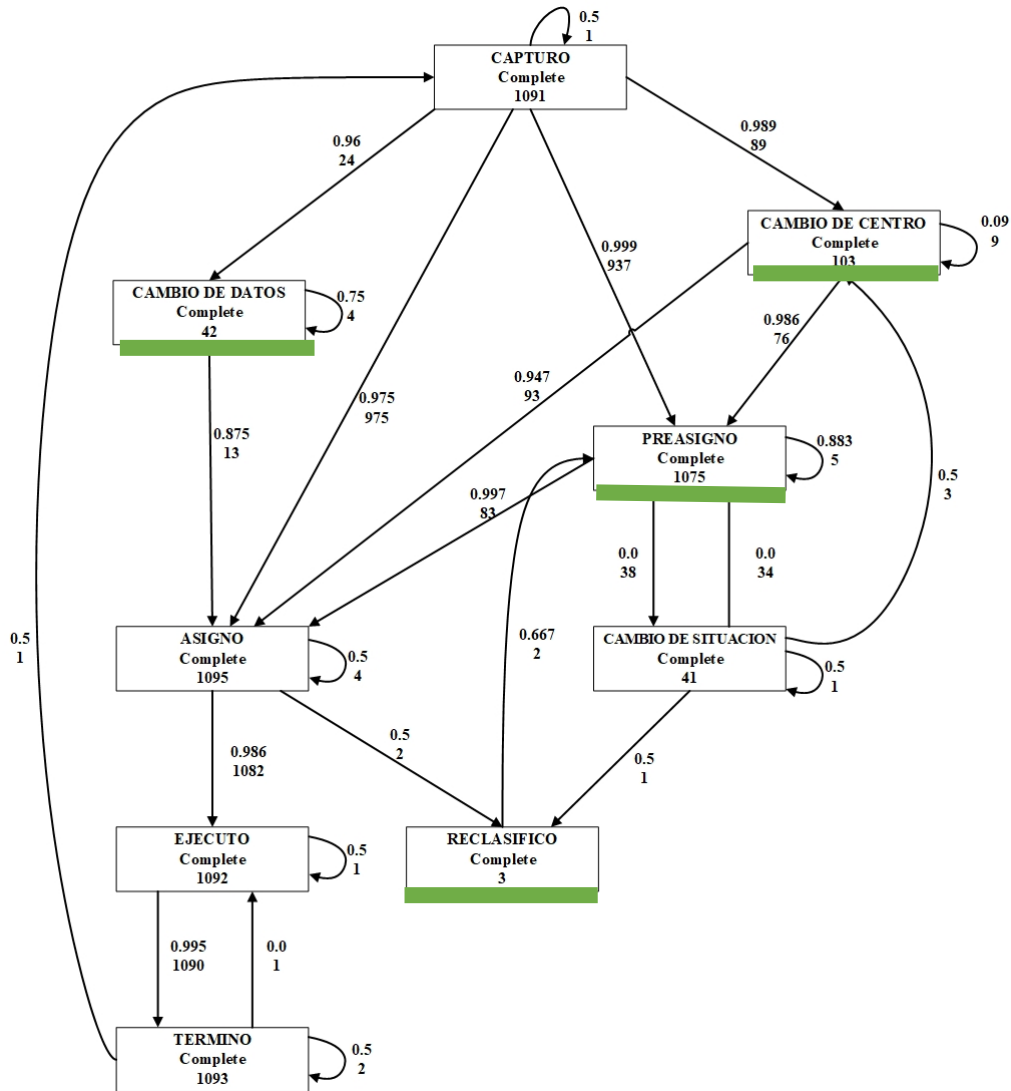


Figura 98. Modelo descubierto.

Para poder ajustar el comportamiento del modelo descubierto en la figura 99 y asemejarlo con el modelo de la empresa se tuvieron que cambiar los parámetros de los umbrales quedando de la siguiente manera:

- relativo al mejor umbral = 0.05
- umbral de dependencia = 0.90
- umbral de bucle longitud uno y dos = 0.90
- umbral de larga distancia = 0.90

Como resultado se obtuvieron las cinco principales actividades con las que cuenta la empresa (figura 100 y 101) y como se mencionó con anterioridad, entre más cerca se encuentre la dependencia a 1 mayor relación entre las actividades habrá.

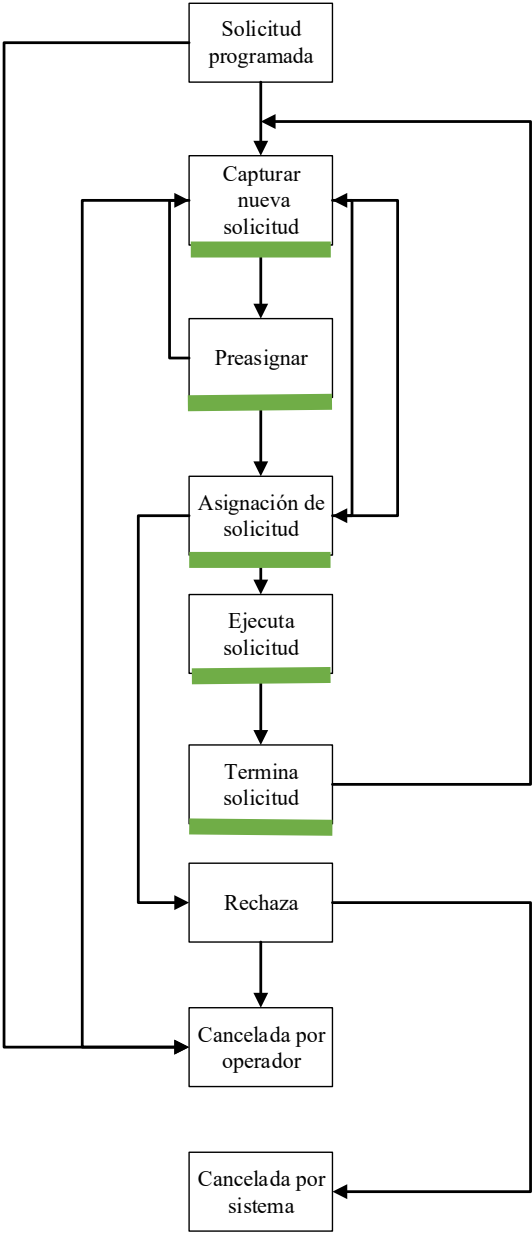


Figura 99. Modelo de la empresa.

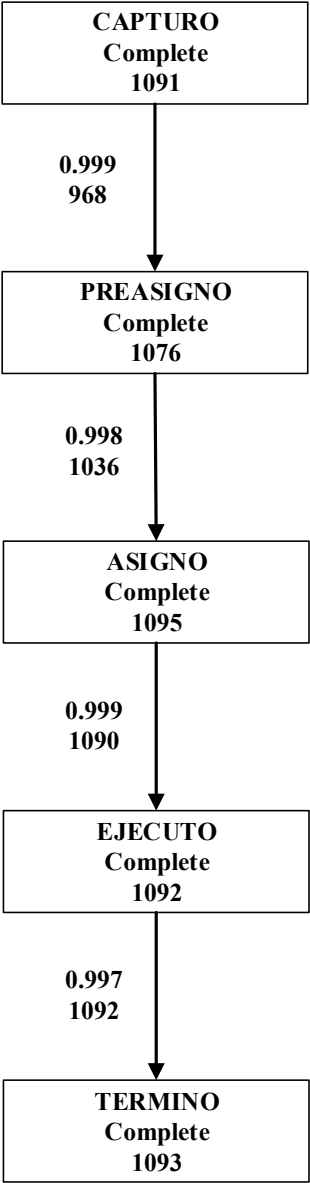


Figura 100. Modelo descubierto ajustado.

## 5.6. Etapa 4 “Análisis de rendimiento”

En esta etapa se analizó el desempeño del proceso; en base a los modelos descubiertos en la etapa 3 se muestran los tiempos de rendimiento del proceso principal y del proceso general; además de identificar en que actividades del flujo del proceso se presentan los cuellos de botella, estos resultados brindaran información valiosa para la toma de decisiones.

En el modelo del proceso principal se detectó que tarda en promedio 9.2 días en pasar de la actividad “capturo” a “Preasigno” (figura 102) y se puede observar que entre las actividades principales (capturo, preasigno, ejecuto y termino) existe una dependencia casi de 1 (entre 997 y 999) lo que significa que existe mucha dependencia entre las actividades.

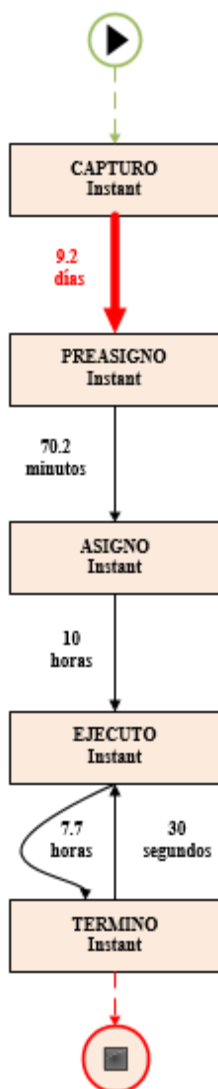


Figura 101. Tiempos de rendimiento del proceso principal.

Como se puede observar en el modelo del proceso general (figura 103) las actividades pueden pasar de un evento a otro de manera instantánea, pueden durar minutos, horas e incluso días; existen tres actividades que tardan mayor tiempo en su ejecución: de la actividad “cambio de centro” a “preasigno” con un tiempo promedio de 52.7 días y de la actividad “cambio de datos” a “preasigno” con un promedio de tiempo alrededor de los 67.7 días, lo cual da como resultado la creación de los cuellos de botella.

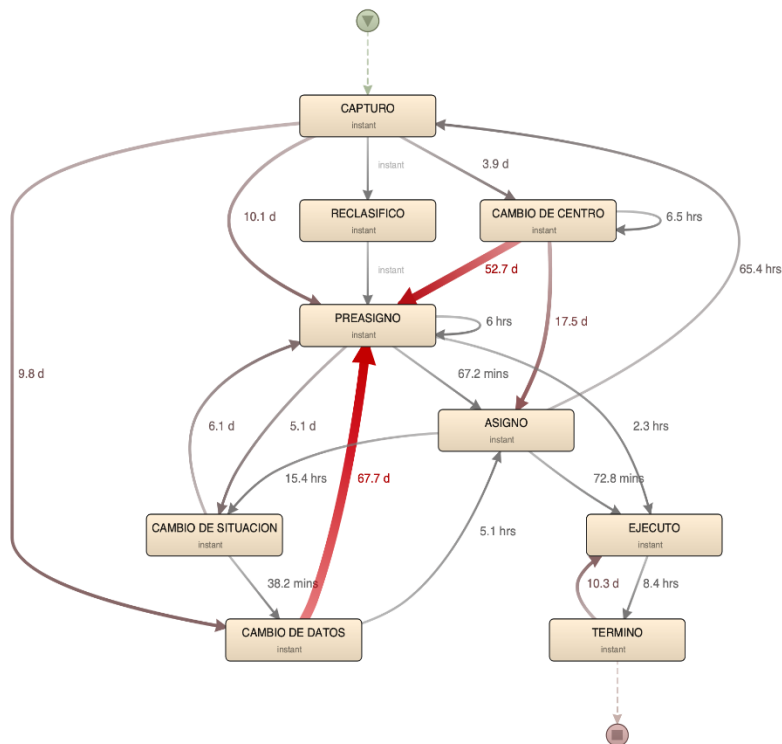


Figura 102. Tiempos de rendimiento del proceso general.

Al determinar los tiempos de rendimiento (figura 92) en la figura 104 se puede observar los cuellos de botella de “Capturo” a “Preasigno” y de las actividades “Preasigno” a “cambio de datos” que son las que mayor tiempo toman para su ejecución.

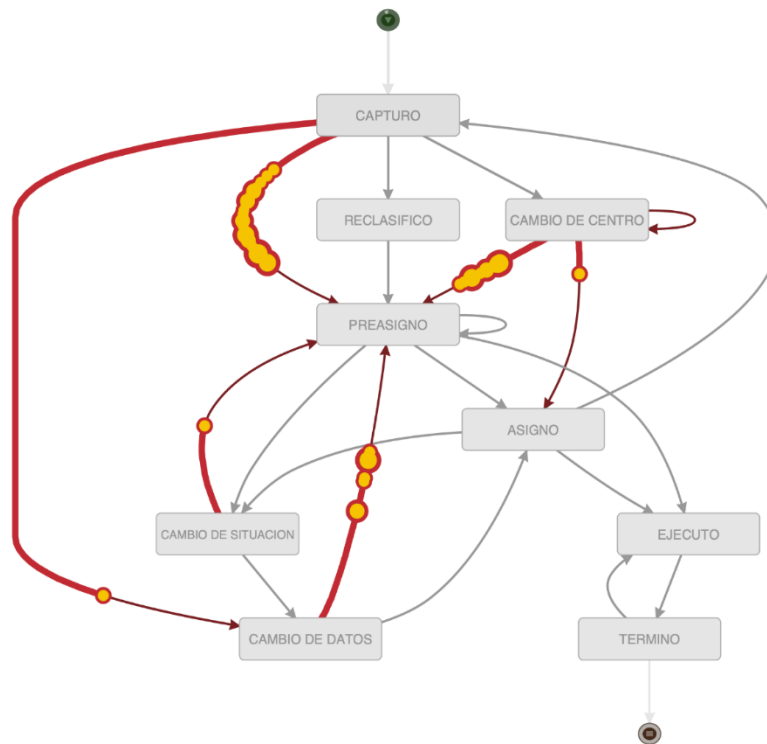


Figura 103. Cuellos de botella dentro del proceso.

Dentro del sistema de información se identificaron 39 variantes, es decir, existen 39 posibles rutas por las que circulan los datos como se puede observar en la tabla 13, estas variantes presentan tiempos promedio que van desde los 4 minutos hasta los 242 días con 19 horas, entre ellas destaca una variante que cuenta con el mayor número de casos, la cual pasa por las 5 actividades principales y con una duración que va de los 4 minutos hasta los 10 días con 19 horas, por otra parte, se identifican 23 rutas que cuentan con 1 caso cada una, pasan por las cinco actividades principales, además, incluyen las cuatro actividades no documentadas, y en algunos casos repiten o saltan alguna actividad.

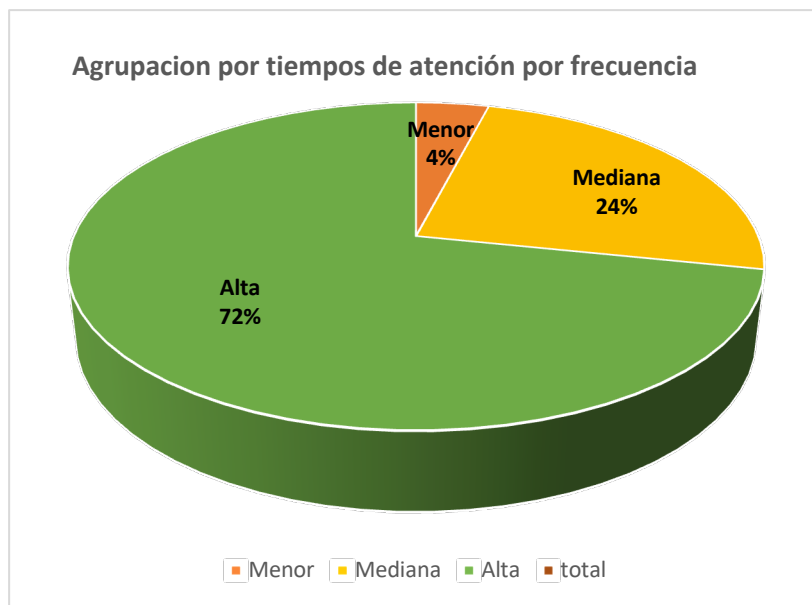
Tabla 11. Variantes.

<b>Variante</b>	<b>Casos</b>	<b>Eventos</b>	<b>Duración mediana</b>	<b>Duración media</b>
<b>Variante 1</b>	799	5	4 min	10 días, 19 horas
<b>Variante 2</b>	46	6	14 días, 4 horas	66 días, 7 horas
<b>Variante 3</b>	37	4	19 min	6 días, 1 hora
<b>Variante 4</b>	23	7	40 min	17 días, 12 horas
<b>Variante 5</b>	18	5	9 días, 21 horas	20 días, 7 horas
<b>Variante 6</b>	14	6	23 días, 2 horas	84 días, 17 horas
<b>Variante 7</b>	5	6	2 horas, 13 min	3 horas, 9 min
<b>Variante 8</b>	5	5	44 min	1 hora, 19 min
<b>Variante 9</b>	5	7	60 días, 3 horas	56 días, 58 min
<b>Variante 10</b>	3	7	3 días, 23 horas	75 días, 16 horas
<b>Variante 11</b>	2	9	4 días, 15 horas	4 días, 15 horas
<b>Variante 12</b>	2	6	1 hora, 19 min	1 hora, 19 min
<b>Variante 13</b>	2	7	77 días, 12 horas	77 días, 12 horas
<b>Variante 14</b>	2	8	8 días, 4 horas	8 días, 4 horas
<b>Variante 15</b>	2	7	9 días, 22 horas	9 días, 22 horas
<b>Variante 16</b>	2	5	2 días, 18 horas	2 días, 18 horas
<b>Variante 17</b>	1	8	216 días, 21 horas	216 días, 21 horas
<b>Variante 18</b>	1	6	242 días, 19 horas	242 días, 19 horas
<b>Variante 19</b>	1	9	180 días, 20 horas	180 días, 20 horas
<b>Variante 20</b>	1	9	124 días, 23 horas	124 días, 23 horas
<b>Variante 21</b>	1	5	8 horas, 41 min	8 horas, 41 min
<b>Variante 22</b>	1	9	22 días, 23 horas	22 días, 23 horas
<b>Variante 23</b>	1	6	31 días, 1 min	31 días, 1 min
<b>Variante 24</b>	1	6	5 min	5 min
<b>Variante 25</b>	1	5	91 días	91 días
<b>Variante 26</b>	1	6	23 min	23 min
<b>Variante 27</b>	1	8	8 horas, 36 min	8 horas, 36 min
<b>Variante 28</b>	1	2	0 milis	0 milis
<b>Variante 29</b>	1	9	6 minutos	6 minutos
<b>Variante 30</b>	1	6	5 días, 2 minutos	5 días, 2 minutos
<b>Variante 31</b>	1	5	15 horas, 24 min	15 horas, 24 min
<b>Variante 32</b>	1	9	1 día, 14 horas	1 día, 14 horas
<b>Variante 33</b>	1	7	213 días, 22 min	213 días, 22 min
<b>Variante 34</b>	1	8	8 días, 12 horas	8 días, 12 horas
<b>Variante 35</b>	1	7	21 horas, 59 min	21 horas, 59 min

Como se puede observar en la tabla 14, el 78% de los casos (779) pasan por las actividades principales que están documentadas por la empresa, por el contrario, solo el 2% de las variantes repiten o saltan alguna actividad y el 20% restante además de pasar por las actividades principales pasan por las cuatro actividades descubiertas (grafica 2).

Tabla 12. Frecuencia de tiempos.

Agrupación por tiempos de atención		
Frecuencia	Casos	Porcentaje
Menor	50	2%
Mediana	300	20%
Alta	889	78%
<b>Total</b>	<b>1092</b>	<b>100%</b>



Grafica 2. Tiempo de atención.

## 5.7. Etapa 5 “Análisis de roles”

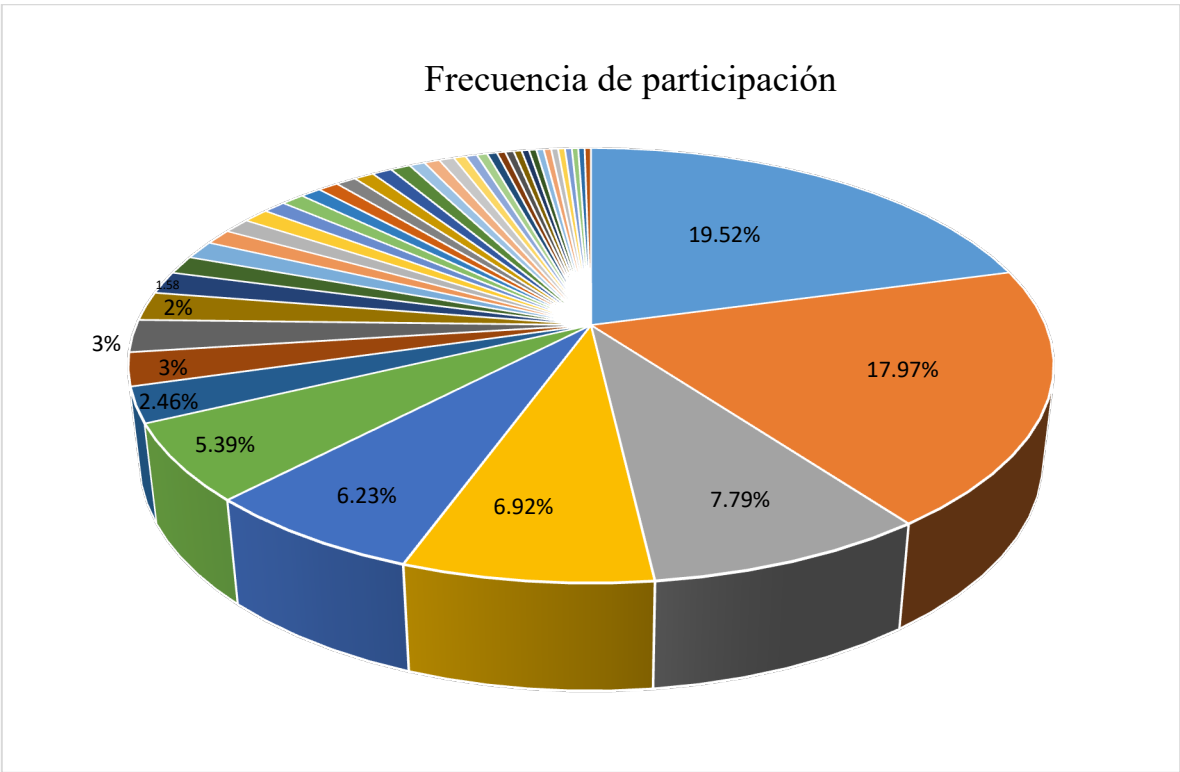
En esta fase se analizaron los datos disponibles sobre las diferentes personas que están involucradas en el proceso; dentro de la documentación obtenida de CFE se manejan cinco tipos de roles: operador, agente, supervisor, oficinista y zonal, sin embargo, no se sabe que rol desempeña cada uno de los 105 usuarios (tabla 15).

Tabla 13. Frecuencias de los recursos dentro del proceso.

Recursos dentro del proceso			Recursos dentro del proceso		
Recurso	Frecuencia	Frecuencia relativa %	Recurso	Frecuencia	Frecuencia relativa %
persona 6	999	19.52	persona 42	40	0.78
persona 8	920	17.97	persona 60	40	0.78
persona 4	399	7.79	persona 62	31	0.61
persona 80	354	6.92	persona 70	31	0.61
persona 14	319	6.23	persona 84	30	0.59
persona 15	276	5.39	persona 45	24	0.47
persona 2	126	2.46	persona 18	23	0.45
persona 10	119	2.32	persona 78	21	0.41
persona 99	116	2.27	persona 81	20	0.39
persona 105	103	2.01	persona 55	17	0.33
persona 17	81	1.58	persona 83	17	0.33
persona 11	66	1.29	persona 46	16	0.31
persona 31	66	1.29	persona 23	15	0.29
persona 35	56	1.09	persona 59	15	0.29
persona 43	55	1.07	persona 69	15	0.29
persona 32	52	1.02	persona 85	15	0.29
persona 25	46	0.9	persona 9	14	0.27
persona 38	46	0.9	persona 21	14	0.27
persona 22	42	0.82	persona 49	14	0.27
persona 3	41	0.8	persona 19	13	0.25
persona 34	41	0.8	persona 26	13	0.25
persona 20	40	0.78	persona 39	13	0.25

De acuerdo a lo antes mencionado, se analizaron únicamente las frecuencias (tabla 15), los resultados muestran que existen 2 personas que tienen una mayor participación dentro del proceso (990 y 920 eventos en los que se involucraron), además, se identificaron 8 personas con una participación que va desde los 103 a los 399 eventos, 24 personas participaron en un rango que oscila de 16 a los 81 eventos y el resto del personal (86 usuarios) se involucró en algún evento solamente 1 vez (grafica 3),

A partir de esos resultados se llega a la conclusión que 2 usuarios tienen una carga de trabajo excesiva por lo que es necesario conocer cuál es el rol de cada uno de ellos, así mismo las actividades y tareas que realizan, para evitar la sobrecarga de trabajo o en su caso realizar una recomendación para la reestructuración adecuada de roles y de actividades delegadas correspondientes, dando como resultado la distribución equitativa de trabajo y así mismo asegurar que cada uno de los usuarios realice únicamente las tareas pertinentes y que estén relacionadas directamente con el proceso.



Grafica 3. Frecuencias de participación de los recursos.

## 5.8. Etapa 6 “Transferencia de resultados

En esta fase se elaboró un informe que presenta los resultados obtenidos los cuales se muestran a continuación.

Hallazgos y resultados obtenidos del caso de estudio:

- La información obtenida de la empresa sirvió para realizar el análisis, sin embargo, se tuvo que hacer uso de internet para obtener información disponible y publica para poder conocer la estructura organizacional, la contextualización, los servicios, los procesos y procedimientos de la empresa.
- **Los datos** obtenidos (1 año 8 meses 605 días) fueron **representativos** para el análisis de minería de procesos.
- La documentación proporcionada por la empresa no cuenta con un dato específico respecto a **el tiempo definido** para la ejecución del proceso, por lo que se trabajó con el tiempo promedio de los casos, el cual equivale a **15.9 días**.
- Al analizar el registro de eventos, se descubrieron **9 actividades** dentro del proceso, **990 casos**, **5119 eventos** y **112 personas** involucradas, además, no se encontraron **casos incompletos, es decir, todos los casos se inician, ejecutan y terminan**.
- Al realizar el modelado de los datos se detectaron **4 actividades** (cambio de datos, cambio de centro, cambio de situación y reclasifico) que no se encuentran documentadas por la empresa.
- Al aplicar el algoritmo minado heurístico se detecta que existen bucles dentro del modelo (**casos que repiten o casos que saltan** alguna actividad), porque quizá se hizo una **mala asignación o existe algún error en los datos** dentro del sistema de información, lo que puede provocar **retrabajo, mayor tiempo en alguna** actividad y por tanto  **cuellos de botella, dando como resultado** desviación o distorsión en el modelo de proceso documentado por la empresa.
- **Al analizar el flujo** del modelo se descubrió que el 14% de los casos recorren las actividades del modelo principal y además por las cuatro actividades que no están

documentadas (cambio de datos, cambio de centro, cambio de situación, reclasifico).

- Los modelos creados a partir de los algoritmos (Alpha, minado heurístico, Fuzzy miner y minería inductiva) pasan por las nueve actividades y proporcionan información (precisión, generalización y simplicidad), sin embargo, se trabajó con el algoritmo minado inductivo que da como resultado una **red causal** y proporciona información referente a las **frecuencias y dependencias entre las actividades del proceso**.
- Se descubrió que existen **inconsistencias en la captura de las actividades** debido a que existen casos en el que las actividades se realizan de manera instantánea lo cual indica que todas las actividades no se pueden realizar al mismo tiempo o por el contrario tardan demasiado tiempo en ejecutarse el paso de una actividad a otra (días), en este último caso provoca **cuellos de botella** en ciertas actividades (cambio de centro” a “preasigno” y de la actividad “cambio de datos” a “preasigno).
- **Se detecto que los datos proporcionados del sistema de información de la empresa cuenta con errores ortográficos** (en los nombres del personal, en las observaciones), **en algunos casos maneja la misma fecha y hora** (puede ser que todas las actividades de un caso se realizan al mismo tiempo), **se duplica alguna actividad que da como resultado la** creación de bucles, direccionamiento de actividades fuera del flujo principal que dan como resultado un modelo tipo espagueti en otras palabras un modelo con falta de precisión, complejo y por tanto difícil de interpretar.
- La documentación obtenida de la empresa no se tienen definido **el rol que fue determinado a** cada persona involucrada en el proceso (112), por lo que solo se analizaron las **frecuencias**, existen **2 personas** que tienen mayor participación dentro del proceso (990 y 920 eventos en los que se involucraron), existen 8 personas con una participación que va desde los 103 a los 399 eventos, 24 personas que van de los 16 a los 81 eventos y el resto del personal (86) solamente 1 vez se involucra en algún evento.

- La minería de procesos ayudo a identificar el proceso real de la empresa en base a los datos analizados.

## Conclusiones

Los hallazgos y resultados obtenidos de la minería de procesos fueron:

- En la actualidad se cuenta con herramientas de inteligencia de negocios que ayudan a las empresas a mejorar sus procesos, a tomar decisiones más acertadas, en menor tiempo y con mayor exactitud, **la minería de procesos** es una de esas herramientas ayuda a las empresas a conocer lo que realmente está sucediendo en el o los procesos de la empresa y si van acorde a lo documentado por la misma.
- Para un análisis más apropiado de los datos, se debe tener especial cuidado en el **formato del *event log*** (fechas, nombres, tiempos, espacios, errores ortográficos, duplicidad de letras, uso de caracteres, etc.).
- Para realizar un **análisis más integral**, es necesario tener contacto directo con el personal adecuado y contar con la documentación necesaria de la empresa o departamento, esto permitirá tener una contextualización de la misma (la estructura organizacional, a sus objetivos y funciones, procesos, etc.) para conocer y trabajar en base a las necesidades y objetivos a cubrir de la organización.

- Es recomendable trabajar con **proyectos** en los que se realizan **preguntas o hipótesis** a responder o en su caso **tener objetivos** para mejorar el desempeño del proceso puesto que los proyectos basados en datos solo pretenden descubrir hallazgos o patrones.
- En base a las herramientas informáticas de la minería de procesos **DISCO** es una herramienta que cuenta con un entorno amigable para el usuario, lo que la hace más fácil de usar para analizar, modelar e interpretar los resultados. Por su parte, **PROM** es una de herramienta de software libre que cuenta con un conjunto de plugin (funciones) que la hace una herramienta más robusta, completa y a la vez complicada, por lo cual se debe tener un conocimiento más amplio sobre cada una de las funcionalidades con las que cuenta para su adecuada aplicación.
- La **minería de procesos** es relativamente nueva (surge en el año 2009 a partir de la minería de datos), por lo que las empresas desconocen acerca de los beneficios de la misma.
- De acuerdo con la literatura encontrada se concluye que las empresas **no han adoptado la cultura** para analizar los datos que se encuentran almacenados en los sistemas de información.
- Se comprobó que la minería de procesos genera nuevo conocimiento, resultado del análisis de los datos que se encuentran en los sistemas de información, el cual es relevante y sobre todo es una estrategia para la toma de decisiones.
- La minería de procesos **es una herramienta fundamental** para las organizaciones, el tener que gestionar grandes **cantidades de procesos** y poder **mejorarlos** supone una gran **ventaja competitiva** frente al resto de las organizaciones, puesto que la minería de procesos **no solo sirve para representar** los procesos de negocio, sino que también **ayuda a mejorar los modelos y a comprobar su efectividad**.

## **Glosario**

**BPM** (Business System Management) Sistema de Gestión por Procesos software empresarial que permite modelar, implementar y ejecutar actividades interrelacionadas procesos de cualquier tipo que pueden incluir clientes, proveedores y otros agentes como participantes en las áreas de los procesos.

**ERM** (Enterprise Risk Management) Gestión de Riesgos Empresariales métodos o herramientas con las que cuentan algunas empresas para gestionar los riesgos con los que se involucra a la vez que intenta lograr los objetivos que se plantearon, esta herramienta intenta prevé lo que va a suceder, anticiparse a los eventos.

**EVENT LOG:** conjunto de información de los registros el tipo de dato.

**LOG BUG:** Es el registro de una actividad es como una bitácora de todas las actividades que se van realizando.

**Sistema PAISS** Process Aware Information System incluyen softwares que permite tener una visión de procesos y actividades aisladas son como bitácoras que registran información sobre las actividades de una organización que afectan la satisfacción del cliente.

**KPIs** Indicador clave o medidor de desempeño de un proceso, estos sistemas permiten tener una visión general del proceso

**Sistema SAP** Systems Applications Products in Data Processing) Sistemas Aplicaciones Productos) sistema que le permite a la empresa administrar os recursos humanos, financieros, contables, productivos, logísticos, etc.

**Sistema ERP** Enterprise Resource Management sistema de planificación de recursos empresariales se hace cargo de operaciones internas de una empresa como la producción, distribución, recursos humanos.

**Utf-8 Unicode** 8 bit Unicode Transformation Format (Unicódigo de transformación de formato de 8 bits) utiliza símbolos de longitud variable, incluye especificación ASCII

## Referencias

Aalst, W. van der. (2012). Process mining: Overview and opportunities. *ACM Transactions on Management Information Systems*, 3(2), 1–17. <https://doi.org/10.1145/2229156.2229157>

Aalst, W. van der. (2013). Service mining: Using process mining to discover, check, and improve service behavior. *IEEE Transactions on Services Computing*, 6(4), 525–535. <https://doi.org/10.1109/TSC.2012.25>

Aalst, W. van der. (2016). *Process mining: Data science in action* (2nd ed.). Springer.

Aalst, W. van der, Adriansyah, A., Alves de Medeiros, A. K., Arcieri, F., Baier, T., Blickle, T., Bose, J. C., van den Brand, P. C. W., Brandtjen, R., Buijs, J. C. A. M., Burattin, A., Carmona, J., Castellanos, M., Claes, J., Cook, J., Costantini, N., Curbera, F., Damiani, E., de Leoni, M., & Wynn, M. T. (2011). Process mining manifesto. En F. K. Barkaoui & S. Dustdar (Eds.), *Business Process Management Workshops (BPM 2011)* (pp. 169–194). Springer. [https://doi.org/10.1007/978-3-642-28108-2\\_19](https://doi.org/10.1007/978-3-642-28108-2_19)

Aalst, W. van der. (2011). *Process mining: Data science in action*. Springer.

Acosta, L., López, O., Vázquez, Y., & Velastegui, E. (2018). Más que minería de procesos y predicción de comportamientos. Un nuevo enfoque [More than mining processes and prediction of behaviors: A new approach]. *Ciencia Digital*, 2(2), 8–28. <https://doi.org/10.33262/cienciadigital.v2i2.70>

Aguirre, H., & Rincón, R. (2015). Minería de procesos: Desarrollo, aplicaciones y factores críticos [Process mining: Development, applications, and critical factors]. *Cuadernos de Administración*, 28(50), 137–157. <https://doi.org/10.11144/Javeriana.cao28-50.mpda>

Arias, M., & Rojas, E. (2018). Guía para gestionar procesos de negocio a través de minería de procesos [Guide to managing business processes through process mining]. *Intersedes*, 17(36), 1–28. <https://doi.org/10.15517/isucr.v17i36.26799>

Bogarín, V. (2018). *Mejora en el descubrimiento de modelos de minería de procesos en educación mediante agrupación de datos de interacción con la plataforma Moodle* [Tesis de doctorado, Universidad de Córdoba]. UCOPress.

Bose, R., & Aalst, W. van der. (2010). Trace alignment in process mining: Opportunities for process diagnostics. *Department of Mathematics and Computer Science, University of Technology, Eindhoven, The Netherlands*, 227–242. [https://www.researchgate.net/publication/202481415\\_Trace\\_Alignment\\_in\\_Process\\_Mining\\_Opportunities\\_for\\_Process\\_Diagnostics](https://www.researchgate.net/publication/202481415_Trace_Alignment_in_Process_Mining_Opportunities_for_Process_Diagnostics)

Bozcaya, M., Gabriels, J., & Van der Werf, J. (2009). Process diagnostics: A method based on process mining. *International Conference on Information, Process, and Knowledge Management* (pp. 22–27). Cancún, México. <https://doi.org/10.1109/eKNOW.2009.29>

CELONIS. (2024). Cuando los procesos funcionan, la IA funciona [When processes work, AI works]. <https://www.celonis.com/es/>

CFE. (2022). Manuales [Manuals]. <https://www.cfe.mx/nuestraempresa/Pages/manuales.aspx>

CFE. (2022). Comisión Federal de Electricidad: Estructura Orgánica Básica [Federal Electricity Commission: Basic Organizational Structure]. <https://www.cfe.mx/estructura/Pages/default.aspx>

Cognizant. (2024). Cognizant España: Análisis de procesos [Cognizant Spain: Process analysis]. <https://www.cognizant.com/es/es/glossary/process-analysis>

Coursera. (2021). *Process mining: Data science in action* [Video]. <https://www.coursera.org/learn/process-mining/home/welcome>

Diez, A. (2016). *Aplicación de la minería de procesos al análisis y modelado de misiones multi-robot* [Trabajo de fin de grado, Universidad Politécnica de Madrid]. Repositorio Institucional UPM.

EcuRed. (2018). Minería de procesos [Process mining]. [https://www.ecured.cu/index.php?title=Miner%C3%ADa\\_de\\_procesos&oldid=3142490](https://www.ecured.cu/index.php?title=Miner%C3%ADa_de_procesos&oldid=3142490)

Fluxicón. (2020). *Libro de minería de procesos: Requerimientos de datos* [Process mining book: Data requirements]. <https://fluxicon.com/book/read/dataext/>

Fulgueira, C., & Reyes, B. (2016). Paralelización del algoritmo fuzzy miner con Open CL [Parallelization of the fuzzy miner algorithm with Open CL]. *Revista Cubana de Ciencias Informáticas*, 10(4), 1–14. ISSN: 2227-1899

G2. (2024). QPR ProcessAnalyzer Reviews & Product Details. <https://www.g2.com/products/qpr-processanalyzer/reviews>

Gastón, F. (2007). *Técnicas de minería de procesos de negocio distribuidos con Bonita OS* [Tesina de licenciatura, Universidad Nacional de La Plata]. Repositorio Institucional UNLP.

González, C. (2016). *Minería de procesos en ambientes sensorizados* [Trabajo de fin de grado, Universidad Carlos III de Madrid]. Repositorio Institucional UC3M.

González, A., Leal, L., Martínez, D., & Morales, D. (2019). Herramientas para la gestión de procesos [Tools for process management]. *Cuadernos Latinoamericanos de Administración*, XV (28), 1–13. <http://www.redalyc.org/articulo.oa?id=409659500003>

Grupo de Ciencia de Procesos y Datos, Universidad RWTH de Aquisgrán. (2024). Minería de Procesos: XESame [Process mining: XESame]. <http://processmining.org/old-version/xesame.html>

Hernández, P. (2012). *Aplicación de técnicas de minería de procesos (PM) para el control y mejoramiento del proceso de compras nacionales e internacionales de bienes y servicios para proyectos de investigación de la Pontificia Universidad Javeriana* [Tesis de maestría, Pontificia Universidad Javeriana]. Repositorio Institucional PUJ.

Loaiza, G. (2018). *Aplicación de una técnica de minería de procesos para el modelamiento del proceso de ventas en una empresa textil en Arequipa* [Tesis de maestría, Universidad Nacional de San Agustín de Arequipa]. Repositorio Institucional UNSA.

Lucid Software Inc. (2024). ¿Qué es la notación de modelado de procesos de negocio? [What is business process modeling notation?]. Lucidchart. <https://www.lucidchart.com/pages/es/que-es-la-notacion-de-modelado-de-procesos-de-negocio>

Morales, A., Martínez, C. I., de Jesús Amador, J., Hidalgo, C. & Zolá, R. (2022). Minería de procesos aplicada a un sistema de solicitudes de servicios al cliente: Un caso de estudio basado en datos. *RISTI: Revista Ibérica de Sistemas e Tecnologías de Informação*, (52), 117-132.

NEXTECH. (s.f.). ¿Qué es BPMN y para qué sirve? [What is BPMN and what is it for?]. Nextech. <https://nextech.pe/que-es-bpmn-y-para-que-sirve/>

Nuritha, I. (2017). Structural similarity measurement of business process model to compare heuristic and inductive miner algorithms performance in dealing with noise. *Procedia Computer Science*, 124, 255–263. <https://doi.org/10.1016/j.procs.2017.12.154>

Pérez, D., Yzquierdo, R., Pupo, E., & Orellana, A. (2018). Minería de proceso para la comprensión del proceso en la etapa de diagnóstico [Process mining for process comprehension on diagnosis stage]. *Ingeniería Industrial*, XXXIV(2), 146–159. <https://dialnet.unirioja.es/servlet/articulo?codigo=6673492>

Pérez, S. (2015). *Minería de procesos* [Tesis de maestría, Universidad Pontificia Bolivariana]. Repositorio Institucional UPB.

Pohl, T. (2019). *An inductive miner implementation for the PM4PY framework* [Bachelor's thesis, RWTH Aachen University]. Repositorio Institucional RWTH.

ProcessMaker. (2020, 15 de septiembre). Uso de herramientas y modelos de procesos empresariales [Use of tools and business process models]. <https://www.processmaker.com/es/blog/business-process-modeling/>

Raeburn, A. (2024). Diagramas de flujo de trabajo: Qué es y cómo hacerlo con ejemplos [Workflow diagrams: What they are and how to make them with examples]. Asana. <https://asana.com/es/resources/workflow-diagram>

Rojas, D., Chavarro, P., & Moreno, L. (s.f.). Técnicas de lógica difusa aplicadas a la minería de datos [Fuzzy logic techniques applied to data mining]. *Universidad Tecnológica de Pereira*. ISSN: 0122-1701

SAS, (2025). Minería de datos ¿Qué es y porque es importante?. [https://www.sas.com/es\\_es/insights/analytics/data-mining.html](https://www.sas.com/es_es/insights/analytics/data-mining.html)

Silva Osses, A. (2017). *Metodología para el análisis de proceso de negocio basada en minería de procesos y datos* [Memoria de titulación, Universidad Técnica Federico Santa María de Chile]. Repositorio Institucional UTFSM. <https://repositorio.usm.cl/bitstream/handle/11673/14074/3560902038214UTFSM.pdf>

Turdibayeva, K. (2024). ¿Qué son los registros de sucesos en la minería de procesos? [What are event logs in process mining?]. ProcessMaker. <https://www.processmaker.com/es/blog/what-are-event-logs-in-process-mining/>

Velázquez, L. (2017). *Método de minería de procesos para incrementar la productividad para proyectos de investigación de las cajas municipales* [Tesis de maestría, Universidad Nacional de San Agustín de Arequipa]. Repositorio Institucional UNSA.

Verbeek, H. M. W., Buijs, J. C. A. M., van Dongen, B. F., & van der Aalst, W. M. P. (2010). XES, XESame, and ProM 6. En *Proceedings of the 2010 International Conference on Business Process Management* (pp. 60–75). Springer. [https://doi.org/10.1007/978-3-642-17722-4\\_5](https://doi.org/10.1007/978-3-642-17722-4_5)

Weijters, A. J. M. M., van der Aalst, W. M. P., & Medeiros, A. K. A. (2006). Process mining with the heuristics miner-algorithm. *Technische Universiteit Eindhoven, 166*, 1–35. [https://www.researchgate.net/publication/229124308\\_Process\\_Mining\\_with\\_the\\_Heuristics\\_Miner-algorithm](https://www.researchgate.net/publication/229124308_Process_Mining_with_the_Heuristics_Miner-algorithm)

Windgassen, C. (2024). Introducción a la minería de procesos: ¿Qué es la minería de procesos? [Introduction to process mining: What is process mining?]. GBTEC Software AG. <https://www.gbtec.com/es/recursos/process-mining/>

