

Pontificia Universidad Católica del Perú
Facultad de Ciencias e Ingeniería



**ANÁLISIS DE ALGORITMOS DE PREDICCIÓN Y OPTIMIZACIÓN
PARA LA MEJORA DE LA CADENA DE SUMINISTRO DEL
EFECTIVO**

**Trabajo de investigación para la obtención del grado de
BACHILLER EN CIENCIAS CON MENCIÓN EN INGENIERÍA
INDUSTRIAL**

AUTORES

Alessandra Lily Pinedo Pinedo
Johao Antony Ramirez Guzmán

ASESOR:

Jonatan Edward Rojas Polo

San Miguel, 21 de octubre de 2020

RESUMEN

El presente trabajo de investigación tiene como principal objetivo el análisis y evaluación de los algoritmos que permiten la solución de problemáticas presentes en la cadena de suministro del efectivo, con referencia al proceso de abastecimiento de caudales, donde existe la necesidad de pronosticar los requerimientos de efectivo de clientes, así como definir y diseñar rutas que permitan abastecer y equilibrar los caudales en las sucursales presentes en el nivel final de la cadena de suministro del efectivo (CSC).

Por un lado, se realiza un determinado análisis de métodos de pronóstico y algoritmos predictivos que se encuentran presentes dentro del campo de *Data Mining*. De esta manera, se observa el nivel de diversificación de herramientas para el pronóstico de requerimientos de efectivo de la cadena de suministro y, finalmente, se determina las técnicas más adecuadas para tratar este tipo de información.

Por otro lado, se realiza un análisis de los algoritmos, heurísticas y metaheurísticas presentes en la disciplina de Investigación de Operaciones, que son de gran utilidad para la solución de modelos matemáticos que expresan una problemática organizacional. En este caso, se hace hincapié en las técnicas de solución para problemas de tipo *Traveling Salesman Problem (TSP)* y *Vehicle Routing Problem (VRP)*, que son vitales para formular un problema de enrutamiento de vehículos para suministro de efectivo en la cadena.

Finalmente, se contrasta la teoría analizada gracias al estudio de casos, donde los autores solucionan ciertas problemáticas presentes en la cadena de suministro del efectivo haciendo uso de técnicas y algoritmos específicos, que permiten ajustar el análisis teórico hacia una visión real de cómo se debería abordar una propuesta de mejora en la administración del efectivo.

AGRADECIMIENTOS

Agradezco a mi madre, por apoyarme constantemente durante toda mi etapa universitaria y por impulsarme a alcanzar mis objetivos personales y como futura profesional.

Agradezco a mi compañera de tesis por su responsabilidad y constante esfuerzo para el cumplimiento de nuestro propósito común.



ÍNDICE

CAPÍTULO 1. MARCO TEÓRICO	1
1.1.- Cadena de suministro	1
1.1.1.- Cadena de suministro del efectivo	2
1.1.2.- Inventarios	7
1.1.2.1.- Clasificación de inventarios según el papel que desempeñan	7
1.1.2.2.- Costos pertinentes	7
1.1.2.3.- Sistemas de reposición de inventarios	9
1.1.2.3.1.- Lote económico de compra (EOQ)	9
1.2.- Investigación de operaciones	12
1.2.1.- Teoría de grafos	16
1.2.1.1.- Definición de grafo	17
1.2.1.2.- Características de los grafos	18
1.2.1.3.- Tipología de grafos	19
1.2.1.4.- Ciclo hamiltoniano	20
1.2.2.- Problema del agente viajero (<i>Travelling Salesman Problem</i> , TSP)	20
1.2.2.1.- Formulación del modelo	21
1.2.2.2.- Solución del modelo	22
1.2.3.- Problema de enrutamiento de vehículos (VRP)	24
1.2.3.1.- Variantes del problema de enrutamiento de vehículos	25
1.2.3.2.- Capacitated Vehicle Routing Problem (CVRP)	27
1.2.3.2.1.- Formulación del modelo	27
1.2.3.2.2.- Solución del modelo	29
1.3.- Pronósticos	36
1.3.1.- Tipos de pronósticos (causales, cualitativos, cuantitativos).	36
1.3.2.- Error de pronósticos	43
1.4.- Data mining	46
1.4.1. Recopilación de datos	47
1.4.2. Preprocesamiento de información	47
1.4.3. Proceso de análisis y algoritmos	49
1.5.- Redes Neuronales Artificiales (RNA)	51
CAPÍTULO 2. ESTUDIO DE CASOS	53
2.1.- <i>Caso 1: Optimization of ATM and Branch Cash Operations using an integrated cash requirement forecasting and cash optimization model</i>	53
2.1.1.- Descripción de la problemática	53
2.1.2.- Procedimiento	54
2.1.3.- Aplicación de herramientas de predicción y optimización	56
2.1.3.1.- Aplicación de herramientas de predicción	56
2.1.3.2.- Aplicación de herramientas de optimización	56
2.1.4.- Resultados obtenidos	57
2.1.5.- Conclusiones	59

2.2.- <i>Caso 2: An Inventory-Routing Problem with Pickups and Deliveries Arising in the Replenishment of Automated Teller Machines</i>	59
2.2.1.- Descripción de la problemática	59
2.2.2.- Procedimiento	60
2.2.3.- Aplicación de técnicas de optimización	61
2.2.3.1- Formulación del modelo	61
2.2.3.2- Aplicación del algoritmo de ramificación y corte	64
2.2.3.3- Aplicación de heurística de clasificación	64
2.2.4.- Solución de la problemática y resultados obtenidos	64
2.2.5.- Conclusiones	69
2.3.- <i>Caso 3: Cash demand forecasting in ATMs by clustering and neural networks</i>	69
2.3.1.- Descripción de la problemática	69
2.3.2.- Procedimiento	69
2.3.3.- Aplicación de herramientas de predicción	71
2.3.3.1.- Aplicación del modelo de series de tiempo multiplicativo con discretización de la estacionalidad	71
2.3.3.2.- Aplicación de modelo de clustering utilizando el método de secuencia alineada o distancia SAM	72
2.3.3.3.- Aplicación de redes neuronales en el pronóstico de la demanda	73
2.3.4.- Solución de la problemática y resultados obtenidos	74
2.3.5.- Conclusiones	75
2.4.- <i>Caso 4: A new game-theoretical multi-objective evolutionary approach for cash-in-transit vehicle routing problem with time windows (A Real life Case)</i>	75
2.4.1.- Descripción de la problemática	75
2.4.2.- Procedimiento	77
2.4.3.- Aplicación de herramientas	79
2.4.3.1.- Teoría de juegos	79
2.4.3.2.- Vehicle Routing Problem with Time Window (VRPTW)	80
2.4.3.3.- Algoritmo genético	80
2.4.4.- Solución de la problemática y resultados obtenidos	82
2.4.5.- Conclusiones	83
2.5.- <i>Caso 5: Improvement of Demand Forecasting Models with Special Days</i>	84
2.5.1.- Descripción de la problemática	84
2.5.2.- Procedimiento	84
2.5.3.- Aplicación de herramientas de pronóstico	85
2.3.4.- Solución de la problemática y resultados obtenidos	86
2.5.5.- Conclusiones	87
CAPÍTULO 3. DIAGNÓSTICO DEL SECTOR BANCARIO	87
3.1.- Análisis de la situación actual del sector bancario	87
3.2.- Análisis de indicadores financieros de la Banca Comercial	88
3.3.- Análisis de <i>stakeholders</i> en la Banca Comercial	94
CAPÍTULO 4. CONCLUSIONES	95
CAPÍTULO 5. BIBLIOGRAFÍA	97

ÍNDICE DE FIGURAS

Figura 1. Red de suministro.	2
Figura 2. Ciclo de vida del efectivo.	3
Figura 3. Vista esquemática de la cadena de suministro del efectivo.	6
Figura 4. Diagrama del nivel de inventario como	10
Figura 5. Costos logísticos en la orden de un pedido.	11
Figura 6. Representación del modelo Q.	12
Figura 7. Variables de decisión de un modelo (con n=6).	14
Figura 8. Representación del problema de los siete puentes de Königsberg.	17
Figura 9. Variable de decisión del modelo TSP.	21
Figura 10. Aplicación de la heurística de barrido en un grafo de enrutamiento de vehículos con capacidad limitada.	31
Figura 11. Componentes de la demanda.	38
Figura 12. Representación de una regresión.	43
Figura 13. Línea de procesamiento de datos.	47
Figura 14. Estructura de una red neuronal.	52
Figura 15. Flujograma del proceso de solución de la problemática.	55
Figura 16. Línea de procesamiento de datos.	60
Figura 17. Diagrama de bloques del procedimiento.	71
Figura 18. Índices de estacionalidad por cluster.	75
Figura 19. Localización geográfica de Teherán y los bancos localizados en el sector.	76
Figura 20. Solución acotada de distancia más corta.	76
Figura 21. Representación típica de un cromosoma.	78
Figura 22. Procedimiento para configurar la ruta óptima.	79
Figura 23. Procedimiento para calcular el riesgo por robos.	80
Figura 24. Algoritmo de fusión de un solo punto.	82
Figura 25. Resultado de aplicación de algoritmo VRPTW	83
Figura 26. Resultados de aplicación del método GLM para el clúster 1.	87
Figura 27. Indicador de solvencia: Ratio de capital global 2019-2020.	89
Figura 28. Indicador de rentabilidad: Rentabilidad económica de las entidades bancarias en el Perú 2019-2020.	90
Figura 29. Indicador de rentabilidad: Rentabilidad de capital de las entidades bancarias en el Perú 2019-2020.	91
Figura 30. Indicador de liquidez de las entidades bancarias en el Perú 2019-2020.	92
Figura 31. Indicador de créditos atrasados de las entidades bancarias en el Perú 2019-2020.	93
Figura 32. Índice de morosidad promedio durante el año 2020.	93
Figura 33. Índice de morosidad promedio por segmento, durante el año 2020.	94

ÍNDICE DE TABLAS

Tabla 1. Estadísticas de resumen de las oficinas en Estambul y los cajeros automáticos en Turquía de Kuveyt Turk antes de la optimización.	54
Tabla 2. Errores porcentuales absolutos medios mensuales para la medición de la precisión de las predicciones de requerimientos de efectivo.	58
Tabla 3. Estadísticas de resumen de las oficinas en Estambul y los cajeros automáticos en Turquía de Kuveyt Turk antes de la optimización.	58
Tabla 4. Brecha entre límites para la solución aplicando heurística de clasificación.	65
Tabla 5. Mejora de resultados al aplicar heurística de clasificación.	67
Tabla 6. Cuartiles usados para la discretización del índice estacional.	72
Tabla 7. Resultados de la aplicación de redes neuronales en los clusters.	74



CAPÍTULO 1. MARCO TEÓRICO

En el presente capítulo, se explican las definiciones teóricas relevantes para llevar a cabo la solución de las problemáticas a las que se enfrenta la cadena de suministro del efectivo. El marco teórico se subdivide en cinco partes: Cadena de suministro, Investigación de operaciones, Pronósticos, *Data mining* y Redes neuronales.

1.1.- Cadena de suministro

De acuerdo con Chopra (2016), una cadena de suministro se define como un conjunto de eslabones que se encuentran relacionados, ya sea directa o indirectamente, con la finalidad de satisfacer las necesidades del cliente, siendo este un elemento integral en la cadena. Además, tal como menciona Carreño (2018), los eslabones mencionados refieren a empresas que mantienen constante colaboración y coordinación, los cuales son atributos fundamentales para garantizar la gestión integral de los flujos de información, material y fondos que tienen lugar a lo largo de toda la cadena de suministro.

Asimismo, Chopra (2016) afirma que una cadena de suministro normalmente se encuentra conformada por las siguientes organizaciones: empresas proveedoras de componentes o materias primas, fabricantes o productores, mayoristas o distribuidores, minoristas y, por último, los clientes finales. Es importante resaltar que el término puede comprenderse como una cadena lineal con un único actor involucrado en cada nivel de la cadena, sin embargo, se trata de una red de suministro que implica mayor complejidad debido a que múltiples actores o empresas pueden pertenecer a cada nivel de la cadena, formando así una red interconectada a través de ciertos flujos bidireccionales. La figura 1 presenta una cadena de suministro donde se evidencia la multiplicidad de eslabones por etapa, donde ciertos productos fluyen de proveedores a productores, distribuidores, mayoristas, minoristas y, finalmente, consumidores (Chopra, 2016). Las líneas de la figura expresan las verdaderas conexiones entre los actores, ya que, en la realidad, un fabricante puede adquirir materia prima e insumos de diversos proveedores e incluso tiene la capacidad de suministrar sus productos tanto a mayoristas como a minoristas.

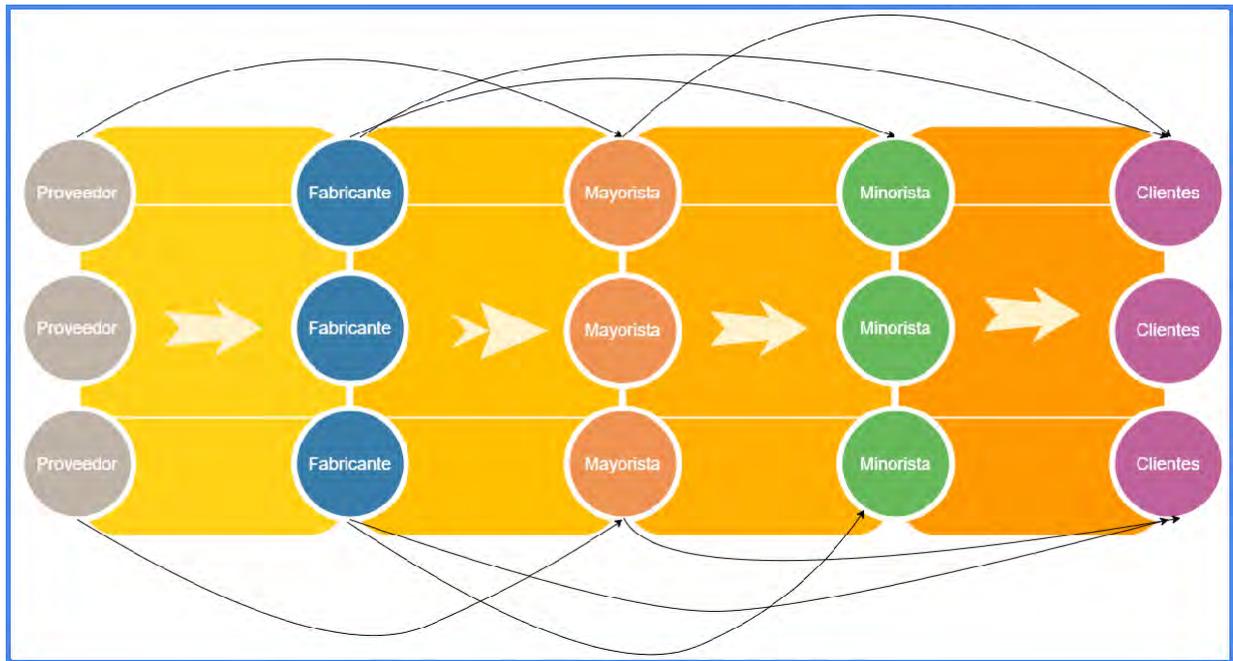


Figura 1. Red de suministro.

Fuente: Chopra, S. (2016)

Además, Carreño (2018) resalta que la cadena de suministro no solo se encuentra conectada a través de flujos de productos o materiales, sino que también existe información y fondos que deben ser administrados. El flujo de productos hace referencia a los materiales de la cadena, desde que son materias primas, pasando por un proceso de transformación, hasta que se convierten en productos terminados, listos para ser adquiridos por los clientes finales. Por otro lado, el flujo de información es relevante ya que diversos datos sobre nuevos productos o avisos de llegada de pedidos y, de manera inversa, información de órdenes de compra o cambios en el mercado deben ser gestionados para garantizar la efectividad de la cadena. Por último, el flujo de fondos se encuentra particularmente relacionado con el dinero, ya que incluye tanto a las transacciones realizadas entre los eslabones de la cadena, como el único flujo positivo, que es el pago de los clientes por el producto otorgado.

1.1.1.- Cadena de suministro del efectivo

Van Anholt (2014) define la cadena de suministro del efectivo como el conjunto de organizaciones y entes individuales involucrados en los flujos ascendentes y descendentes de efectivo, objetos de valor, finanzas y otros, que puedan ser intercambiados por un valor monetario en tiempo presente. Por su parte, Rajamani (2005) identifica a la cadena de suministro del efectivo como una cadena de comportamiento puro, puesto que no existe consumo real en ninguno de sus eslabones y en caso de presentar una calidad inadecuada, el efectivo será

inmediatamente desechado del sistema por un ente regulador. Además, en el caso de una cadena multinacional, el autor afirma que esta tiene un comportamiento de circuito cerrado, pues el dinero se mantiene fluyendo en el sistema en todo su ciclo de vida y vuelve a circular múltiples veces hasta ser sacado del sistema por su deterioro. En el caso de una cadena nacional o local, el dinero puede ser exportado o importado de fuentes externas, formando parte de una cadena mucho más grande.

Según Rajamani (2005), las reservas federales de cada país clasifican al efectivo en cuatro categorías básicas (Ver Figura 2). *New Cash* refiere al dinero proveniente del flujo inicial de salida explicado en el párrafo anterior, el cual, una vez dentro del sistema, pertenece a una de las tres categorías restantes. La primera es denominada *Fit Cash* o efectivo apto, el cual tiene una calidad suficiente para ser ofrecido al cliente, por medio de un cajero automático, oficina o abono directo. La segunda categoría refiere al *Non-ATM Fit Cash* o efectivo no apto para cajeros automáticos, el cual es adecuado para casi todas las operaciones exceptuando los retiros en cajeros, debido a que el Banco Central promueve la salida de dinero proveniente de emisión primaria, es decir, dinero impreso recientemente. La tercera categoría se denomina *Unfit Cash* o efectivo deteriorado, el cual se encuentra sucio, roto o no sigue los parámetros impuestos por la Reserva Federal o Banco Central. El efectivo perteneciente a esta categoría es inaceptable para la circulación y debe ser destruido por el Banco Central, para luego ser repuesto a las cuentas poseedoras de este efectivo en forma de dinero no físico, es decir, a través de un depósito a la cuenta del dueño del caudal. Cabe resaltar que la clasificación anterior fue definida solo para billetes, sin embargo, las monedas tienen la misma clasificación, pero varían en el tratamiento de deteriorados.



Figura 2. Ciclo de vida del efectivo.
Fuente: Rajamani (2005).

La cadena inicia con la producción de efectivo a partir de que ciertas empresas especializadas en acuñación e impresión utilizan materias primas para producir billetes a nombre del Banco Central. Un Banco Central, banco de reserva o autoridad monetaria es la institución que administra el flujo inicial de salida del efectivo, su suministro y las tasas de interés, con la finalidad de conseguir la estabilidad de precios y minimizar las fluctuaciones económicas a gran escala. Con el objetivo de garantizar el suministro de efectivo a cada uno de los demandantes, se contratan proveedores de servicios logísticos; conocidos como empresas transportadoras de efectivo o transportistas blindados, los cuales transportan, almacenan y realizan actividades de *cross-docking* que involucren efectivo entre entidades de la cadena. Según Nan y Long (2009), estas actividades contribuyen con la eficiencia de la cadena, por lo que estos actores representan uno de los eslabones más importantes.

Continuando con el flujo de efectivo, las llamadas "*Brink's Vaults*", "*Bank Vaults*" o bóvedas, a cargo de las entidades bancarias, reciben el dinero de las empresas transportadoras para su posterior distribución a los eslabones en etapas posteriores del ciclo de vida del efectivo. La bóveda funciona como un almacén de efectivo a gran escala, la cual recibe el caudal en diferentes clasificaciones y lo distribuye según se requiera. Cabe resaltar que estas entidades se encargan de recontar el efectivo recibido por el Banco Central y ordenarlo por denominaciones. Este puede seguir su curso hacia dos destinos, de acuerdo con Rajamani (2005), los cuales son:

- Cajeros Automáticos: máquinas dispensadoras o receptoras de efectivo en múltiples divisas, en las cuales se puede retirar o depositar efectivo, respectivamente. Los cajeros automáticos pueden encontrarse dentro de las sucursales bancarias o fuera de ellas. Funcionan como almacenes de efectivo, por lo que, dependiendo de su funcionalidad, deben ser abastecidos o desabastecidos con regular frecuencia.
- Oficinas: sucursales bancarias que prestan servicios a clientes grandes y pequeños. Funcionan como almacenes de efectivo y requieren ser abastecidos o vaciados con cierta frecuencia, actividad que encargan a las empresas transportadoras de valores.

Finalmente, el efectivo de alguna de estas fuentes se dirige al consumidor final, en forma de efectivo apto y a su vez los consumidores devuelven el efectivo hacia

alguno de estos canales, generando una recirculación pura a lo largo de la cadena. A continuación, se muestra el flujo descrito en la *Figura 3*.



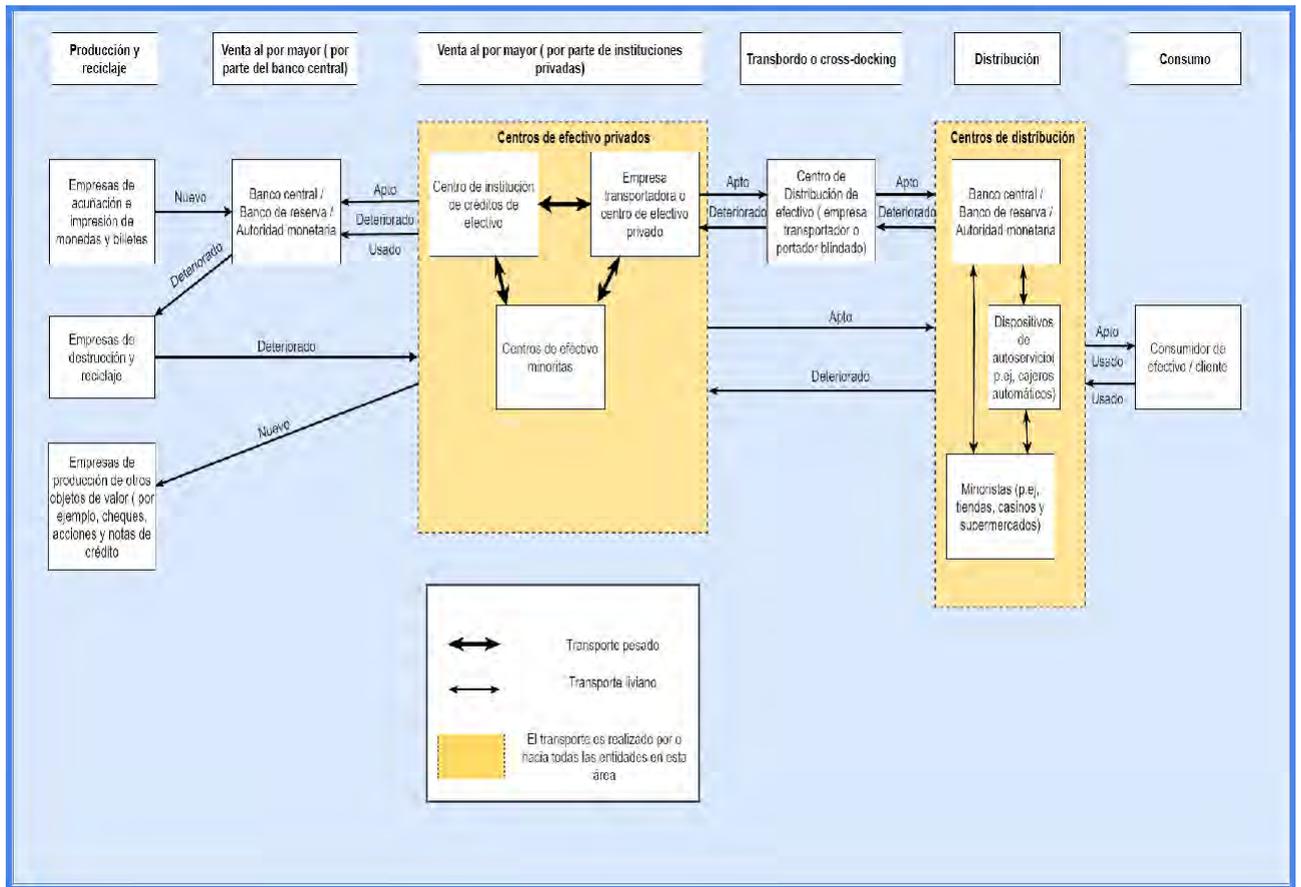


Figura 3. Vista esquemática de la cadena de suministro del efectivo.
 Fuente: Van, R. (2014).

1.1.2.- Inventarios

De acuerdo con Ballou (2004), los inventarios o *stocks* son acumulaciones de material que se encuentran en diversas ubicaciones de la cadena, visualizándose como materias primas, repuestos, componentes o productos terminados. Además, Carreño (2018) menciona que dichas existencias se suelen encontrar guardadas en almacenes, pero también pueden estar en tránsito a través de un medio de transporte o, incluso, ubicadas en establecimientos o tiendas, listas para su adquisición por el cliente final.

La necesidad de mantener inventarios en una cadena radica en la gran dificultad de conocer con exactitud la demanda de los productos, así como de suministrarlos de manera instantánea (Carreño, 2018). Tal como afirma Ballou (2004), aunque mantener inventarios significa incurrir en costos de almacenamiento, también permite que una empresa reduzca sus costos a través de economías de escala presentes en la compra, producción y transportación. Adicionalmente, contar con inventarios permite que una empresa mejore su nivel de servicio al cliente, ya que será capaz de responder con mayor celeridad a las constantes solicitudes de los clientes.

1.1.2.1.- Clasificación de inventarios según el papel que desempeñan

Según la clasificación realizada por Carreño (2018), los *stocks* pueden ser de tipo normal o de seguridad, de acuerdo con el papel que desempeñan. El inventario normal es aquel que está destinado a satisfacer la demanda promedio de los procesos de la cadena, mientras que el inventario de seguridad hace alusión a la cantidad de existencias presentes con la finalidad de afrontar variaciones de la demanda de los productos, además de posibles alteraciones en los tiempos de entrega de los proveedores. Ballou (2004) comenta que dicho *stock* de seguridad se considera como una cantidad adicional al *stock* normal, el que es calculado a partir de conceptos estadísticos que permiten cuantificar el nivel de confianza de la cantidad de inventario de este tipo y el grado de variabilidad tanto de los requerimientos de productos como del tiempo de reaprovisionamiento.

1.1.2.2.- Costos pertinentes

Ballou (2004) afirma que los costos logísticos suponen un conflicto entre sí, lo que genera que se deba buscar un balance de estos, para así minimizar el costo total. En particular, los costos asociados a la gestión de inventarios, que

también tienen el mismo comportamiento de conflicto, serán necesarios para definir la política de reposición de inventarios y son los siguientes:

- A. **Costos de compra:** Hace referencia al costo por la adquisición de los artículos o materiales, que proviene del producto del precio unitario por la cantidad a solicitar del mismo.
- B. **Costos de adquisición:** Son todos aquellos costos que se encuentran involucrados en la emisión de un pedido, específicamente con los procesos que realiza el área de compras y que involucra el procesamiento, ejecución, transmisión, manejo, compra y seguimiento de un pedido, por lo que dicho costo permanece constante o fijo por pedido (Ballou, 2004).
- C. **Costos de posesión de inventarios:** Los costos de poseer inventarios se suscitan como producto de mantener artículos durante un periodo de tiempo (Ballou, 2004). Según Carreño (2018), estos costos se pueden clasificar en costos financieros, de almacenamiento, seguros y de riesgos de inventario. Los costos financieros pueden ser intereses o costos de oportunidad, de acuerdo a la procedencia del dinero, mientras que los costos de almacenamiento hacen referencia a los egresos debido a la ocupación de un espacio determinado, denominado almacén, en el que se guardan los inventarios. Además, los costos provenientes de seguros forman parte de mantener inventarios, ya que estos se contratan con el objetivo de proteger los *stocks* frente a pérdidas por incendios, robos u otros daños. Por último, los costos de riesgos de inventario refieren a mermas, obsolescencia y pérdidas a los que las existencias se encuentran expuestas durante su almacenamiento.
- D. **Costos por falta de existencias:** Estos costos se presentan cuando se efectúan las roturas de *stock*, los cuales pueden ser de dos tipos: costos por pérdida de venta y costos de pedido pendiente (Ballou, 2004).

Por un lado, tal como menciona Carreño (2018), los costos asociados a las pérdidas de venta están relacionados a un costo de oportunidad por el retiro de un pedido por parte del cliente, es decir, se trata del dinero que la empresa deja de percibir debido al desabastecimiento de productos. En

este apartado, también se deben considerar los efectos negativos que genera la falta de respuesta ante una solicitud por parte del cliente, ya que este puede optar por cambiar de proveedor, alterando el nivel de ventas futuras debido a la pérdida absoluta de dicho cliente.

Por otro lado, los costos de pedido pendiente hacen referencia a los desembolsos por la utilización de medios de transporte más eficientes, pero menos económicos, así como costos de mano de obra extra para la realización del procesamiento del pedido, a causa de la problemática de desabastecimiento. Es preciso mencionar que también se encuentra involucrado, en este tipo de costo, el impacto negativo generado sobre las ventas, debido a las demoras o tiempos adicionales de espera por parte del cliente.

1.1.2.3.- Sistemas de reposición de inventarios

En este apartado, se optará por definir las dos políticas de reposición de inventarios más utilizadas:

1.1.2.3.1.- Lote económico de compra (EOQ)

Hillier (2010) define que el objetivo de hallar el lote económico de compra se basa en determinar la frecuencia y la cantidad a colocar en el inventario de modo que el total de costos pertinentes explicados en la sección anterior sea mínimo. Según Carreño (2018), para que esta política sea aplicable, deben cumplirse tres premisas: la demanda y el lead time del proveedor son conocidos y constantes, no existen descuentos por uso de economías de escala por parte del proveedor y el lote se entrega una única vez con la totalidad del pedido. Hillier (2010) considera esta política como un “*Modelo EOQ básico*” y posteriormente define variaciones a esta, que serán descritas en los próximos apartados. Carreño (2018) afirma que para este modelo la necesidad de stocks de seguridad es nula y el nivel de inventario promedio corresponde a la mitad del lote ordenado. En la *Figura 4* se describe el patrón de los niveles de inventario con un tamaño de lote denominado “Q”.

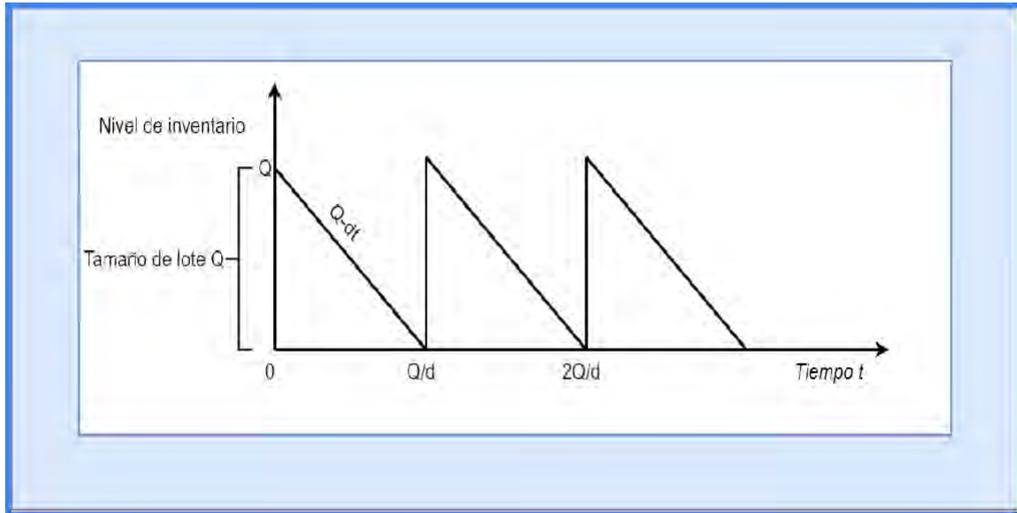


Figura 4. Diagrama del nivel de inventario como función del tiempo del modelo EOQ básico.
Fuente: Hillier y Lieberman (2010).

Se determina el función *Costo Total de Compra* y se escoge el lote que minimice los costos de posición de inventarios, el cual se obtiene derivando la función e igualando a cero (*Ver Figura 5*). Chase(2009) define a la posición de inventario como la cantidad disponible más la orden de pedido menos los pedidos acumulados hasta el momento del reabastecimiento, por lo que para optimizarlo se debe minimizar la función descrita, la cual se presenta a continuación:

$$TC = D * C + (D/Q) * S + (Q/2) * H \quad (1)$$

Donde:

TC = Costo anual total

D = Demanda (anual)

C = Costo por unidad

Q = Cantidad a pedir

S = Costo de preparación o costo de hacer un pedido

H = Costo anual de mantenimiento y almacenamiento por unidad de inventario promedio

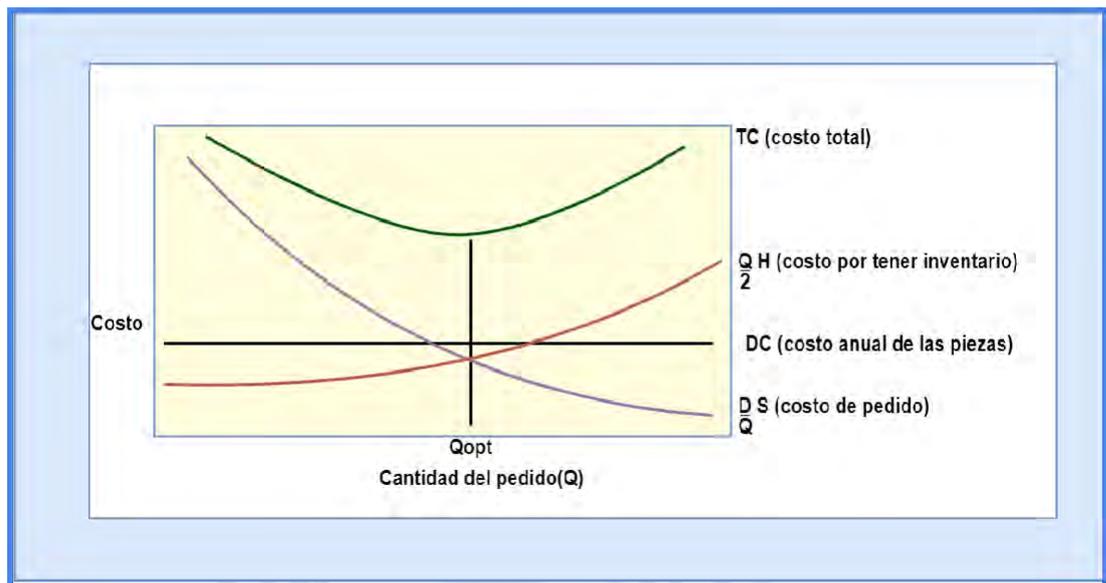


Figura 5. Costos logísticos en la orden de un pedido.

Fuente: Chase (2009).

Como resultado de la optimización, se obtiene la siguiente expresión para calcular la cantidad óptima de pedido o EOQ.

$$Q_{opt} = \sqrt{\frac{2DS}{H}} \quad (2)$$

Fuente: Chase (2009).

Cabe resaltar que también es importante determinar el mejor momento para realizar la orden optimizada, lo que se define como punto de reorden y estará enlazada directamente al lead time del proveedor (Carreño 2018).

1.1.2.3.2.- Punto reposición (Sistema Q)

Carreño (2018) y Chase (2009) establecen la necesidad de estudiar un caso en el que la demanda no sea constante, limitación del modelo EOQ; por lo que, desarrollan la lógica del punto de reorden, sistema Q o sistema de revisión continua. Carreño (2018) afirma que la demanda no es conocida ni es constante en este modelo, sino que, por el contrario, se asume aleatoria y que se rige bajo una distribución normal con media μ y desviación estándar de σ^2 conocidas, lo que aporta una visión real y consecuente con situaciones prácticas. Chase (2009) complementa que, dada esta problemática, es necesario contar con inventario de seguridad para ofrecer un nivel de

La primera etapa requiere de un estudio detallado acerca del sistema relevante o proceso involucrado en la problemática a solucionar. Taha (2012) afirma que la definición de la problemática requiere responder a las siguientes preguntas: ¿Cuáles son las posibles alternativas?, ¿En base a qué restricciones se debe tomar la decisión? y ¿Cuál es la meta para evaluar las posibles alternativas?. Hillier y Lieberman (2010) complementan que también se debe tomar en cuenta las relaciones entre áreas de la organización con el área o proceso de estudio. Esto se debe a que los objetivos apropiados, que suelen ser establecidos a partir de los criterios de los tomadores de decisiones en la organización, deben estar alineados a objetivos globales de la organización, pues no se trata de encontrar soluciones únicamente óptimas para el proceso de estudio, sino de encontrar soluciones óptimas a nivel de toda la compañía (Hillier y Lieberman, 2010).

Asimismo, es necesario recolectar información relevante del sistema, ya que servirá posteriormente para definir los parámetros o constantes del modelo a definir en la siguiente etapa (Hillier y Lieberman, 2010).

B. Formulación de un modelo científico

La siguiente etapa del proceso consiste en la formulación de la problemática definida bajo la elaboración de un modelo matemático, el cual permitirá representar idealmente la realidad por la que la organización se encuentra atravesando (Hillier y Lieberman, 2010). Las interrogantes definidas por Taha (2010) en la primera etapa responden a los principales componentes de un modelo de la disciplina de investigación de operaciones. Tal como expresa Hillier y Lieberman (2010), un modelo matemático se encuentra conformado tanto por sistemas de ecuaciones como por expresiones matemáticas que permiten describir el problema raíz. Específicamente, el modelo cuenta con los siguientes componentes:

Variables de decisión. Son los elementos del modelo que representan los recursos del problema (Kong, 2010). De este modo, si el modelo debe tomar n -ésimas decisiones, estas se plasmarán a través de variables o incógnitas $x_1, x_2, x_3, \dots, x_n$ (Ver Figura 7), para las que debe calcularse un resultado asociado (Hillier y Lieberman, 2010).

Variables de decisión para un modelo con 6 variables:

X1= número de empleados que empiezan a las 0 horas (primer periodo)
X2= número de empleados que empiezan a las 4 horas (segundo periodo)
X3= número de empleados que empiezan a las 8 horas (tercer periodo)
X4= número de empleados que empiezan a las 12 horas (cuarto periodo)
X5= número de empleados que empiezan a las 16 horas (quinto periodo)
X6= número de empleados que empiezan a las 20 horas (sexto periodo)

Figura 7. Variables de decisión de un modelo (con n=6).

Fuente: Kong, M. (2010)

Función objetivo. Hace alusión a una función $f(x)$, la que puede ser equivalente a una utilidad, costo, cantidad, tiempo, etc. En términos de Hillier y Lieberman (2010), la función objetivo refiere a la medida del desempeño del sistema que se expresa en términos de las variables de decisión y que finalmente se buscará optimizar (maximización o minimización del resultado). Un ejemplo de dicha función se puede visualizar a continuación:

$$n = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \quad (3)$$

Fuente: Kong, M. (2010)

Restricciones. Las condiciones que se imponen sobre las variables de decisión se denominan restricciones del modelo. Estas se visualizan en dicho modelo a través de ecuaciones o desigualdades, que permiten limitar las soluciones del modelo, restringiendo la cantidad de posibles soluciones de la problemática (Hillier y Lieberman, 2010). A continuación, se presentan las restricciones de un modelo de ejemplo:

$$x_6 + x_1 \geq 2; \quad (4)$$

$$x_1 + x_2 \geq 4; \quad (5)$$

$$x_2 + x_3 \geq 8; \quad (6)$$

$$x_3 + x_4 \geq 6; \quad (7)$$

$$x_4 + x_5 \geq 6; \quad (8)$$

$$x_5 + x_6 \geq 4; \quad (9)$$

Fuente: Kong, M. (2010)

Es necesario mencionar que las constantes presentes en las expresiones de la función objetivo y de las restricciones tienen el nombre de parámetros, los

que deberán asignarse a partir de una adecuada recolección de datos relevantes.

C. Solución del modelo científico

Una vez construido el modelo, se debe proceder a resolverlo utilizando diferentes métodos que permitan alcanzar las posibles soluciones (Kong, 2010). Para ello, Taha (2012) menciona que una solución será factible sólo si satisface todas las restricciones definidas en el modelo. Además, será óptima si corresponde al mejor valor de la función objetivo, pues no solo se trata de encontrar una solución a tal problemática, sino que la investigación de operaciones tiene un enfoque basado en encontrar una mejor solución, también conocida como solución óptima (Hillier y Lieberman, 2010).

Con respecto a los métodos utilizados, estos pueden ser múltiples, por lo que el método más apropiado para la solución dependerá del tipo y nivel de complejidad del modelo matemático previamente formulado. Una de las técnicas más empleadas en la disciplina de investigación operativa es la programación lineal, la cual está diseñada para el caso en que la función objetivo y restricciones son ecuaciones o desigualdades lineales. También existen otros métodos de solución como las técnicas de programación entera, donde se buscan variables enteras, así como las técnicas de programación por objetivos, programación de red y programación no lineal (Hillier y Lieberman, 2010).

La mayoría de las técnicas descritas poseen una particularidad y es que se determinan mediante algoritmos o reglas que describen cómo obtener los valores de las variables de decisión a través de un proceso repetitivo de cálculos, donde cada repetición corresponde a una iteración o escalón para llegar al resultado óptimo final (Kong, 2010).

De acuerdo con Hillier y Liberman (2010), en ocasiones, los equipos de investigación operativa tienden a utilizar procedimientos heurísticos que, si bien no garantizan una solución óptima, permiten encontrar una solución subóptima que corresponde a una buena solución. Esto se debe a que algunos modelos matemáticos pueden ser imposibles de resolver debido a la complejidad matemática y, en la actualidad, no existe algún algoritmo de optimización apropiado para utilizarlo en la solución del modelo (Taha, 2012).

D. Prueba del modelo y mejora continua

Esta etapa de prueba y mejoramiento se conoce como validación del modelo. En primera instancia, un modelo matemático tendrá muchas fallas, por lo que se deben incorporar arreglos que permitan considerar nuevos factores, realizar estimaciones más precisas de los parámetros o tomar en cuenta interrelaciones de las variables. Por ello, el modelo irá cambiando de acuerdo con la adición de nuevas consideraciones, las que permitirán obtener una nueva versión de este, el cual producirá resultados válidos y aceptables (Hillier y Lieberman, 2010).

E. Preparación para aplicación del modelo

Una vez que el modelo haya sido validado y debidamente modificado, se deberá optar por la implementación de un sistema de soporte para tomar decisiones, de tal manera que este pueda ser utilizado múltiples veces por el personal administrativo, a manera de apoyo en la solución de problemáticas de negocio (Hillier y Lieberman, 2010). Este sistema debe estar integrado con bases de datos que proporcionen la información necesaria para alimentar el modelo, además de integrarse con un programa que genere informes con los resultados del modelo y las variables de salida, presentando un análisis detallado de la información e incluyendo los impactos de dichos resultados (Hillier y Lieberman, 2010).

F. Implementación

Como etapa final del proceso, el sistema previamente desarrollado deberá ser implementado según los lineamientos establecidos por la gerencia. Una vez implementado, el sistema debe estar en constante mejora y retroalimentación, ya que, por un lado, los supuestos pueden modificarse o, por otro lado, se puede requerir de ciertos cambios provenientes del funcionamiento del sistema (Hillier y Lieberman, 2010).

1.2.1.- Teoría de grafos

En este apartado, se desarrollarán los principales conceptos de la teoría de grafos. Debido a su gran aplicación en la vida real, esta proporcionará nociones básicas que permitirán darle solución a problemas de optimización de ruteo de vehículos.

De acuerdo con Jungnickel (2008), la teoría de grafos se inició con Leonhard Euler, matemático y físico suizo, quien solucionó el problema de los siete puentes de Königsberg (hoy Kaliningrado). Según Álvarez y Parra (2013), Euler se planteó el siguiente problema: ¿Es posible realizar una caminata a lo largo de la ciudad, de tal forma que cruce los siete puentes una sola vez antes de regresar al punto de partida?. Cabe resaltar que la ciudad se encuentra conformada por una Isla Mayor, además de tres territorios, uno ubicado al norte, otro al este y el último al sur de la ciudad. Además, estos territorios se comunicaban a través de los siete puentes con los que contaba la ciudad portuaria. Entonces, Euler se percató que la solución dependía únicamente de las propiedades de conexión de los territorios y no de sus formas precisas. Por ello, representó a los cuatro territorios a través de puntos o vértices, mientras que los puentes que unían estos territorios se reemplazaron por curvas. Gracias a dicha representación, se obtuvo el primer grafo que permitió modelar la realidad, el que se puede visualizar en la Figura 8 (Jungnickel, 2008). A partir de dicha simplificación Euler pudo probar que no era factible cruzar todos los puentes sin tener que volver a pasar nuevamente por uno de ellos, siempre que haya tres o más vértices en los que convergen un número impar de caminos. En este caso, tres vértices contaban con tres curvas que llegaban a su punto, mientras que al vértice de la Isla Mayor llegaban cinco caminos (Álvarez y Parra, 2013).

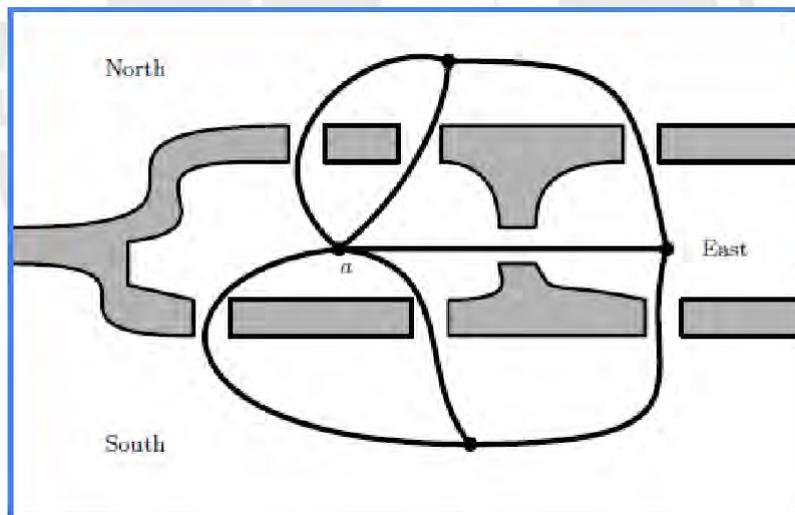


Figura 8. Representación del problema de los siete puentes de Königsberg.
Fuente: Jungnickel, D. (2008)

1.2.1.1- Definición de grafo

Según Caicedo, Wagner y Méndez (2010), un grafo se refiere a una estructura de datos no lineal, que se utiliza para modelar una gran diversidad de aplicaciones. Se encuentra formado por un conjunto de vértices V y un conjunto de arcos E , los cuales

hacen referencia a segmentos que unen dos vértices. La notación de un grafo, de acuerdo con Jungnickel (2008), se representa a continuación:

$$G = (V, E) \quad (10)$$

Asimismo, un arco e , elemento del conjunto E , se representa a partir de los vértices a y b que une dicho arco:

$$e = \{a, b\} \quad (11)$$

1.2.1.2.- Características de los grafos

A. Orden y talla de un grafo

A continuación, se empezará definiendo el orden de un grafo. Caicedo, Wagner y Méndez (2010) mencionan que el orden hace referencia al número de nodos totales del grafo. De acuerdo con la notación de los autores, el orden se representa de la siguiente manera:

$$ord(g) = |N(G)| = N \quad (12)$$

Por otro lado, la talla de un grafo representa el número de arcos del grafo y se denota como:

$$talla\ de\ G = |A(g)| = A \quad (13)$$

Entonces, un grafo será de orden N y de talla A .

B. Grado de un vértice

De acuerdo con Jungnickel (2008), el grado de un vértice es el número de arcos o aristas que convergen en este, el cual se denomina x . La notación, según Caicedo, Wagner y Méndez (2010), es la siguiente:

$$grad(x) \quad (14)$$

C. Trayectoria de un grafo

Una trayectoria, denominada T , se define como una secuencia de nodos, con notación v_1, v_2, \dots, v_n , tal que los pares ordenados $(v_1, v_2), (v_2, v_3), \dots, (v_{n-1}, v_n)$ representan a los arcos que componen la trayectoria, los cuales pertenecen al grafo (Caicedo, Wagner y Méndez, 2010).

De acuerdo con Álvarez y Parra (2013), un circuito o ciclo es un tipo de trayectoria que se caracteriza por volver a su punto de origen, es decir, se trata de una trayectoria T que inicia y finaliza en el mismo vértice. Además, existen los bucles o lazos, los cuales forman una conexión con el mismo nodo, es decir, salen del vértice y convergen en el mismo, sin pasar por otro de ellos (Caicedo, Wagner y Méndez, 2010).

1.2.1.3.- Tipología de grafos

En primer lugar, los grafos se pueden clasificar de acuerdo con las características de sus arcos. Si un grafo cuenta con sus arcos direccionados, se convierte en un grafo dirigido u orientado, mientras que uno con aristas que no cuentan con una dirección lleva el nombre de grafo no dirigido y el orden de sus elementos en la notación no altera la interpretación, debido a que representa al mismo arco (Caicedo, Wagner y Méndez, 2010). En el caso de los grafos dirigidos, cada uno de sus arcos, de acuerdo con su notación, se compone por un par ordenado (a,b) , donde “a” representa el vértice de origen, mientras que “b” representa el vértice de destino, por lo que el orden de los elementos de la arista si será relevante en su notación.

En segundo lugar, los grafos pueden clasificarse bajo la cantidad de arcos que unen sus vértices. Tal como afirma Caicedo, Wagner y Méndez (2010), un grafo se denomina como simple si es que, en primer lugar, no cuenta con lazos y tampoco existe más de un arco que unos dos vértices de G . En cambio, los que sí tienen al menos un par de nodos con más de un arco uniéndolos se denomina grafo múltiple o multigrafo.

Por otro lado, si todos los vértices de un grafo G tienen el mismo grado, el grafo es regular (Jungnickel, 2008). Un caso particular de esta tipología es el grafo completo, el cual se caracteriza por poseer nodos que se encuentran unidos a todos los otros, lo que significa que cada vértice tendrá un grado equivalente a $n - 1$, donde n representa el número de vértices del grafo G . A partir de ello, se demuestra que un grafo completo contiene exactamente $\frac{n(n-1)}{2}$ arcos (Caicedo, Wagner y Méndez, 2010).

Además, los grafos se pueden clasificar como cíclicos o acíclicos. Un grafo cíclico es aquel que contiene algún ciclo o circuito en su interior, mientras que un grafo acíclico es aquel que no cuenta con estos (Caicedo, Wagner y Méndez, 2010). Tal como

expresa Jungnickel (2008), un grafo acíclico con n vértices tiene como máximo $n - 1$ arcos.

Por último, un grafo es conexo, si y sólo si, para cualquier par de vértices, existe como mínimo un camino posible que los una, sin importar que se deba pasar a través de otros vértices o que el par de vértices pueden estar unido por un camino con más de un arco.

1.2.1.4.- Ciclo hamiltoniano

Tal como menciona Caicedo, Wagner y Méndez (2010), es común evaluar la posibilidad de pasar por todos los vértices de un grafo conexo no dirigido, sin repetir ningún nodo y de tal manera que, al final, se retorne al punto de partida. A tal trayectoria se le denomina ciclo hamiltoniano y su nombre deriva del matemático William Rowan Hamilton, quien inventó el juego icosiano, el cual constaba de un dodecaedro regular, el cual tenía las esquinas etiquetadas con nombres de distintas ciudades y el objetivo del juego era encontrar un camino a lo largo de los bordes del dodecaedro, que permita visitar las ciudades una sola vez, llegando, finalmente, a la ciudad de partida (Jungnickel, 2008).

Asimismo, una trayectoria que se componga de $n - 1$ arcos y recorra todos los vértices de un grafo, se denomina hamiltoniana, mientras que un grafo $G = (N, A)$ que contenga un ciclo hamiltoniano es llamado grafo hamiltoniano.

Tal como menciona Jungnickel (2008), existen dos teoremas que permiten reconocer que un grafo sea hamiltoniano. Primero, el teorema de Ore expresa que, si un grafo G es conexo y sin lazos, con tres o más vértices y si la suma de los grados de dos vértices u y v es mayor a su cantidad de vértices, considerando ello para cada par de vértices, tal que $u \neq v$, entonces G se considera como hamiltoniano. Segundo, el teorema de Dirac menciona que si un grafo G cuenta con tres o más vértices y cada vértice de G tiene como mínimo un grado de $\frac{n}{2}$, es decir, al menos tiene un grado equivalente a la mitad de la cantidad de vértices, entonces G es hamiltoniano.

1.2.2.- Problema del agente viajero (*Travelling Salesman Problem, TSP*)

Una vez definida la teoría de grafos, si G tiene asociado a sus arcos ciertos valores de distancias d_{ij} o costos asociados al recorrido, resulta de interés conocer cuál es el ciclo que recorren todos los vértices una sola vez y que, adicionalmente, tenga asociado el mínimo

recorrido o costo de transporte. Esta casuística se conoce como el problema del agente viajero (Caicedo, Wagner y Méndez, 2010). Este se caracteriza por ser uno de los problemas fundamentales en la optimización y es modelado por un grafo en el cual los vértices se asemejan a n ciudades y arcos todos los posibles caminos entre las ciudades. La solución del problema corresponde a un camino circular o ciclo hamiltoniano, por el que el viajero visitará todas las ciudades y, finalmente, estará de regreso a su ciudad de origen al final del viaje, tomando en cuenta que se trata del ciclo óptimo, donde los costos o distancias se minimizan (Jungnickel, 2008).

Tal como comenta Taha (2012), el modelo TSP está definido bajo dos tipos de datos. Es necesario conocer el número de ciudades n , así como las distancias d_{ij} entre las ciudades, asignándose una distancia infinita o muy elevada si las ciudades no se pueden comunicar. Dado que se trata de un recorrido, no es posible conectar una ciudad consigo misma, por lo que también se colocará una distancia muy elevada a las distancias con $i = j$ dentro de la matriz.

Adicionalmente, un modelo TSP matemático es considerado como simétrico si las distancias entre dos ciudades de ida y vuelta son iguales, es decir, todos los d_{ij} son equivalentes a los d_{ji} , en cambio, si existe diferencia entre las distancias de ida y vuelta de dos ciudades, entonces el modelo es de carácter asimétrico (Taha, 2012).

1.2.2.1.- Formulación del modelo

La variable de decisión de un problema del agente viajero corresponde a una de tipo binaria, que expresa que el recorrido incluirá al desplazamiento del viajero desde la ciudad i hacia la ciudad j , cuando x_{ij} toma el valor de 1:

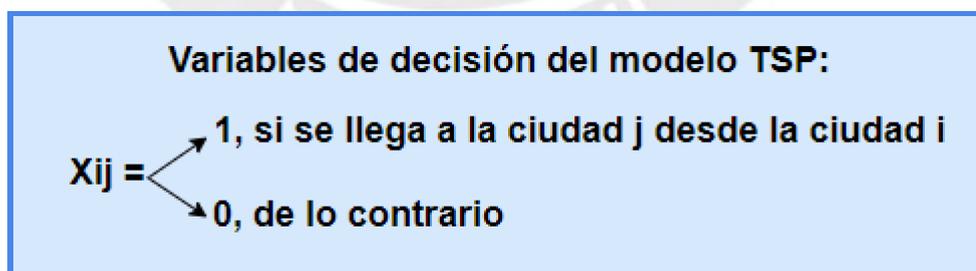


Figura 9. Variable de decisión del modelo TSP.

Fuente: Taha, H. (2012)

La función objetivo busca optimizar la distancia que recorre el viajero, la que se visualiza a continuación:

$$\text{Minimizar } z = \sum_{i=1}^n \sum_{j=1}^n d_{ij} x_{ij}, \text{ para } d_{ij} = \infty \text{ para todos los } i = j \quad (14)$$

Fuente: Taha, H. (2012)

Finalmente, las restricciones son las siguientes:

$$\sum_{j=1}^n x_{ij} = 1, i = 1, 2, \dots, n \quad (15)$$

$$\sum_{i=1}^n x_{ij} = 1, i = 1, 2, \dots, n \quad (16)$$

$$x_{ij} = (0, 1) \quad (17)$$

La solución forma un viaje redondo por las ciudades (18)

Fuente: Taha, H. (2012)

Las restricciones (15) y (16) definen un modelo de asignación, donde solo un arco dirigido debe salir de cada ciudad i y solo un arco dirigido debe llegar a cada ciudad j , respectivamente. La restricción (17) define que la variable de decisión sea de carácter binaria, mientras que la última restricción del modelo (18) exige que la solución forme un ciclo (Taha, 2012).

1.2.2.2.- Solución del modelo

Para obtener una solución del modelo se dispone de ciertos algoritmos que pueden ubicarse dentro de las siguientes categorías (Caicedo, Wagner y Méndez, 2010):

- Algoritmos que obtienen una solución óptima pero que requiere de un procedimiento exhaustivo.
- Algoritmos que obtienen como resultado un ciclo hamiltoniano, pero que no siempre es el óptimo. Este tipo de algoritmos requiere de un procedimiento con un número considerable de operaciones.

Taha (2012) especifica que existen dos algoritmos de programación entera exactos, que permiten solucionar este modelo: el algoritmo de ramificación y acotamiento y el algoritmo de plano de corte. Sin embargo, el procedimiento propuesto de solución en ellos no lo hace computacionalmente confiable, por lo que también se suelen utilizar las heurísticas, que permiten obtener soluciones razonables del problema, pero no necesariamente óptimas.

A. Algoritmos TSP exactos

- Algoritmo de ramificación y acotamiento (*Branch and bound algorithm*)

Este propone iniciar con la solución de un modelo de asignación, es decir, tomando en cuenta la restricción (15), (16) y (17), lo que puede ser resuelto a través de cualquier software como AMPL o Excel. Si dicha solución se trata de un ciclo hamiltoniano, entonces la solución óptima se ha encontrado. Sin embargo, es común encontrarse con subrecorridos, donde se deberá iniciar con una ramificación de los posibles resultados adicionando nuevas restricciones al modelo, basadas en hacer que alguna de las variables de decisión x_{ij} del sub recorrido más pequeño sean igual a 0. De esta manera, se llegará a la solución óptima final, cuando se terminen de analizar todos los nodos y en el que se obtenga el menor valor de la función objetivo para un conjunto de variables de asignación que formen un ciclo y no varios subrecorridos (Taha, 2012).

- Algoritmo del plano de corte

Este algoritmo propone solucionar el modelo a través de la adición de un conjunto de restricciones que permitan obtener un ciclo o recorrido y excluir soluciones con subrecorridos. Se define una variable continua u_i y otra u_j , ambas mayores o iguales a cero y para las ciudades $i, j = 2, 3, \dots, n$ (Taha, 2012). A partir de esta definición, se agregan las siguientes restricciones de cortes al modelo:

$$u_i - u_j + nx_{ij} \leq n - 1, \quad i = 2, 3, \dots, n; \quad j = 2, 3, \dots, n; \quad i \neq j \quad (19)$$

Sin embargo, tal como menciona Taha (2012), el modelo, utilizando este algoritmo, crece exponencialmente al incrementar la cantidad de ciudades, convirtiéndose en computacionalmente insoluble.

B. Heurísticas de búsqueda local

Las heurísticas de tipo local proporcionan una solución óptima local, sin embargo, este no siempre resulta ser el verdadero ciclo óptimo.

- Heurística del vecino más cercano

Esta heurística propone una solución para el problema del agente viajero a partir de un procedimiento que inicia con la elección aleatoria de una de todas las ciudades del modelo, para luego conectarla con la ciudad más cercana a ella. La ciudad previamente conectada se deberá conectar con el vértice más cercano que aún no se encuentre conectado. Este proceso debe continuar hasta que se haya definido el recorrido (Taha, 2012). Durante la realización del proceso se debe ir modificando el valor z , o distancia total del recorrido, a partir de la siguiente ecuación definida por Caicedo, Wagner y Méndez (2010):

$$z = z + dij \quad (20)$$

donde dij hace referencia a la nueva distancia entre las ciudades i y j que ha sido conectada a la red. Cabe mencionar que el valor de z empieza siendo 0, al momento de comenzar con el procedimiento.

- Heurística de inversión

Tal como comenta Taha (2012), este tipo de heurística consta de la aplicación de mejoras a recorridos actuales obtenidos. Tal como su nombre lo expresa, se trata de mejorar el recorrido actual a partir de la aplicación de una inversión en el orden de los vértices intermedios de un subrecorrido abierto, el que se caracteriza por no contar con un arco para llegar a ser cerrado. Asimismo, se establecen ciertas limitantes, como que el número mínimo de elementos del subrecorrido debe ser igual a 2; el número máximo, equivalente a $n - 2$ para modelos simétricos y equivalente a $n - 1$ cuando son asimétricos.

1.2.3.- Problema de enrutamiento de vehículos (VRP)

Según Baldacci, Battarra y Vigo (2008), el problema de enrutamiento de vehículos o Vehicle Routing Problem refiere básicamente al diseño de rutas óptimas que permitan distribuir y entregar productos hacia un conjunto dado de consumidores que lo demandan, haciendo uso de una flota de vehículos o medios de transporte terrestre determinados. Irnich, Toth y Vigo (2014) añaden que estas rutas a determinar, para el posterior recorrido por la flota de vehículos proporcionada, deben ser factibles de ejecutar y, además, deben permitir obtener el mínimo costo de transporte.

Tal como afirma Irnich, Toth y Vigo (2014), Dantzig y Ramser fueron los pioneros en la definición del problema de enrutamiento de vehículos, aplicándolo a la entrega de gasolina a

diferentes estaciones en el año 1959. Los autores propusieron la primera formulación matemática, así como la solución del problema mediante un enfoque algorítmico. Tiempo después, específicamente en el año 1964, Clarke y Wright propusieron un método de solución efectivo, a partir de la aplicación de una heurística que permita obtener resultados del problema de enrutamiento de vehículos aproximados al óptimo.

Por otro lado, este tipo de problema corresponde a una generalización o extensión del problema del agente viajero definido previamente, siendo el TSP de gran ayuda para la formulación de otros problemas combinatorios más complejos, pero de gran relevancia práctica, como lo es el VRP, que permite resolver problemas de distribución en la cadena de suministros (Daza, Montoya y Narducci, 2009).

Realizando una analogía, el TSP corresponde a la variante más básica del VRP, ya que se trata de diseñar una sola ruta que minimice los costos de distribución y permita abastecer la demanda de los clientes utilizando un único vehículo. Sin embargo, en la realidad logística, se presentan diversas variantes de un enrutamiento de vehículos, debido a que se presentan distintas demandas, que deben ser abastecidas por una flota de vehículos con cierta capacidad, además de considerar que muchas veces se perciben limitaciones de tiempo para la realización de las entregas (Daza, Montoya y Narducci, 2009). Por ello, se especifican las variantes del *Vehicle Routing Problem* en el siguiente subíndice.

1.2.3.1.- Variantes del problema de enrutamiento de vehículos

Vehicle Routing Problem alberga un conjunto de variantes que se han formulado a lo largo de las décadas. Las versiones más reconocidas debido a su gran aplicación en el proceso de distribución física se detallan a continuación:

A. *The Capacitated Vehicle Routing Problem (CVRP)*

Esta variante corresponde a la versión más conocida de VRP y consiste en la distribución de productos desde un único almacén hacia distintos puntos de demanda, los cuales solicitan determinada cantidad de productos. Asimismo, se cuenta con cierto número de vehículos con los mismos costos de transporte asociados e igual capacidad, que se encuentran disponibles para efectuar ciertas rutas, de tal manera, que toda la demanda se satisfaga y la flota retorne, finalmente, al almacén de partida (Irnich, Toth y Vigo, 2014). Este tipo de problema se conoce como ruteo de vehículos con flota homogénea, debido a que los vehículos son idénticos, a nivel de capacidad y de costos de transporte (Oliveira, 2004).

B. *Fleet Size and Mix Vehicle Routing Problem (FSMVRP)*

En este tipo de problema, los vehículos con los que se dispone para la distribución tienen distintos tamaños, capacidades y costos asociados, lo que se conoce como flota heterogénea. Según Oliveira (2004), existe un conjunto T de tipos de vehículos, donde $k \in T$ y la capacidad de cada tipo de vehículo k es equivalente a q^k . Asimismo, los costos de cada tipo de vehículo se presentan como c_{ij}^k y los tiempos de viaje son equivalentes a t_{ij}^k , para cada tipo. Este caso conlleva a no solo diseñar las rutas de viaje para suministrar la demanda, sino que también se debe definir la composición o mix de vehículos a utilizar para los viajes (Oliveira, 2004).

C. *Vehicle Routing Problem with Backhauls (VRPB)*

Si bien la mayoría de los problemas de enrutamiento de vehículos aluden a una entrega o distribución de bienes hacia distintas ubicaciones de demanda, también conocido como *one-to-many VRP*, existen problemáticas que refieren a la contraparte de ello, es decir, donde se desea realizar el recojo de materiales desde ciertas ubicaciones para posteriormente ser transportadas al almacén. Este tipo de casuística es común al final de la cadena de suministro, donde existen compañías que aplican la logística inversa, donde se colectan devoluciones de clientes o se recogen envases u otros materiales a ser reciclados. Los problemas referidos al recojo de bienes se conocen como *many-to-one VRP* (Irnich, Toth y Vigo, 2014).

Entonces, si se combinan tanto recojos como entregas se obtienen los problemas de tipo *pickup-and-delivery (PDPs)*, donde la variante más generalizada de este subgrupo es la de *VRP with Backhauls*. Ahora, se cuenta con dos tipos de consumidores, unos que requieren de entregas de productos del almacén y otros que tienen productos que deben ser recogidos y trasladados al almacén de la compañía de distribución. Asimismo, se establece una relación de precedencia en el servicio hacia los dos tipos de clientes, es decir, los clientes que requieren de entregas recibirán su demanda antes que se proceda a realizar el recojo de material del otro tipo de clientes (Irnich, Toth y Vigo, 2014).

D. *Split Delivery Vehicle Routing Problem (SDVRP)*

En el marco de la distribución física y la logística, se puede llevar a cabo un requerimiento donde la demanda supere la capacidad de un vehículo, por lo que será necesario efectuar un servicio con más de una visita. Este es el caso de este tipo de problema, que permite que el nivel de requerimientos de un cliente pueda ser dividido en demandas más pequeñas, tal que puedan ser suministradas por diferentes vehículos. Asimismo, existen casos en los que atender la demanda en cantidades más pequeñas produce la disminución de costos logísticos, por lo que el método de división de pedidos resulta más eficiente en la distribución (Irnich, Toth y Vigo, 2014).

E. *Vehicle Routing Problem with Time Window (VRPTW)*

Tal como menciona Oliveira (2004), en este tipo de problema, cada cliente solicita su entrega en una ventana de tiempo, es decir, define un rango o periodo de tiempo $[a,b]$ permitido en el que el vehículo debe arribar y suministrar el requerimiento. Esto significa que el tiempo en el que la flota llegue al punto de ubicación no debe superar el límite superior “b” de tiempo y si llega antes del límite inferior “a”, entonces deberá esperar a que el cliente abra.

F. *Periodic Vehicle Routing Problem (PVRP)*

Este problema se da cuando el horizonte de tiempo para realizar las entregas de los pedidos es de varios días, por lo que, en este caso, se deberán asignar el o los días para el suministro de los pedidos de cada cliente, además de agrupar clientes y diseñar rutas que permitan abastecer la demanda cada día. En este caso, debe existir un enrutamiento de vehículos por día, tal que permita distribuir los pedidos en el periodo de tiempo asignado.

1.2.3.2.- Capacitated Vehicle Routing Problem (CVRP)

Dado que CVRP es la variante del VRP más utilizada y estudiada, se procede a profundizar en la formulación de este tipo de problema de enrutamiento de vehículos, así como a explicar ciertos procedimientos de solución del mismo.

1.2.3.2.1.- Formulación del modelo

Según Irnich, Toth y Vigo (2014), antes de la formulación del modelo, se realizará una breve explicación de la notación matemática. En primer lugar, en un modelo de CVRP, el depósito de origen es representado por el punto 0, el que debe abastecer un conjunto de puntos n que representan a los

consumidores, siendo $N = \{1, 2, \dots, n\}$ el número total de estos. Además, existe una cantidad específica de demanda por parte de cada consumidor $i \in N$, la que se representa por un escalar p_i . Entonces, la flota de M vehículos que suministrará esta demanda, al ser homogénea, posee una capacidad equivalente a P_k (donde $1 \leq k \leq M$) y opera bajo costos iguales, los que se representan como c_{ij} y refieren al costo de transporte asociado a la movilización desde el punto i hasta el punto j , por lo que se trata de una matriz $N \times N$. Durante una ruta, un vehículo parte del depósito común hacia cada uno de los consumidores $S \subseteq N$ que debe abastecer en este viaje.

Estructurando la información, el punto 0 del depósito $\forall y$ el conjunto N conforman los vértices del grafo dirigido, mientras que los arcos del gráfico son las líneas que conectan a los nodos. Es importante mencionar que la demanda asociada al nodo del depósito es equivalente a 0 (Irnich, Toth y Vigo, 2014).

Tal como menciona Daza, Montoya y Narducci (2009), el problema tiene como variable de decisión a x_{ijk} , que es una matriz de tamaño $N \times N \times M$ que alberga al conjunto de binarios que expresan la asignación de los arcos (i, j) a utilizar en la distribución realizada por el vehículo k .

A continuación, se presenta la formulación del modelo CVRP, bajo la notación realizada por Daza, Montoya y Narducci (2009).

$$\min \sum_{i=0}^M \sum_{j=0}^N \sum_{k=1}^M c_{ij} x_{ijk} \quad (21)$$

Sujeto a:

$$\sum_{k=1}^M \sum_{j=1}^N x_{ijk} \leq M, \quad i = 0 \quad (22)$$

$$\sum_{k=1}^M \sum_{j=0}^N x_{ijk} = 1, \quad \forall i \in [1, N] \quad (23)$$

$$\sum_{j=1}^N x_{ijk} = \sum_{i=1}^N x_{ijk}, \quad \forall k \in \{1, M\}, \quad i = 0 \quad (24)$$

$$\sum_{i=0}^N \sum_{j=0}^N p_i x_{ijk} \leq P_k, \quad \forall i \in [1, N] \quad (25)$$

$$\sum_{i=1}^N \sum_{j=1}^N x_{ijk} \leq |S| - 1, \quad \forall S \subseteq (V - \{0\}), |S| \geq 2, \forall k \in \{1, M\} \quad (26)$$

$$x_{ijk} \in \{0, 1\}, \forall i, j \in [1, N], \forall k \in \{1, M\} \quad (27)$$

Fuente: Daza, Montoya y Narducci (2009).

Previamente, se presenta la función objetivo en (21), que busca minimizar los costos asociados al transporte total de los requerimientos de los N consumidores desde el depósito de partida. Las siguientes corresponden a las restricciones del modelo, donde (22) hace referencia a la capacidad en cuanto a cantidad de vehículos, pues solo podrán partir como máximo M vehículos desde el punto de partida. Las siguientes dos restricciones garantizan que solo un arco ingrese y salga de cada nodo, es decir, que un cliente sea abastecido una única vez. La restricción (25) consigue que la capacidad de un vehículo no sea excedida por la demanda de los clientes asignados a dicho medio de transporte. La siguiente restricción establece que no deben existir rutas o viajes que no sean del tipo hamiltonianas, mientras que la última restricción (27) exige que la variable de decisión sea binaria (Daza, Montoya y Narducci, 2009).

1.2.3.2.2.- Solución del modelo

De acuerdo con Oliveira (2004), la solución de los modelos de VRP pueden tomar dos enfoques distintos. Pueden utilizarse algoritmos exactos, que encuentran una solución óptima, pero que presentan una elevada complejidad cuando se trata de un problema de ruteo con muchos clientes (más de 50). En este tipo de solución, se suele utilizar el esquema de ramificación y acotamiento, así como el método de plano de corte, algoritmos que también se utilizan para la solución de problemas TSP y que pueden consultarse en el apartado 1.2.2.2.

Daza, Montoya y Narducci (2009) agregan que existen otras técnicas de solución, que se dividen en heurísticas y metaheurísticas. La primera alude a métodos de optimización local, mientras que la siguiente se refiere a distintos métodos de optimización local inteligente, sin embargo, estas no presentan una base matemática formal que las sustente y no garantizan la optimalidad global.

En este apartado, se explicarán las heurísticas y metaheurísticas más utilizadas para la solución de problemas de enrutamiento de vehículos,

debido a que estas permiten obtener un resultado en menor tiempo que haciendo uso de algoritmos exactos.

A. Heurísticas de búsqueda local

De acuerdo con Oliveira (2004), las heurísticas son procedimientos que se basan en la exploración en un espacio de búsqueda local para obtener soluciones aceptables en tiempos de cálculo razonables. A continuación, se detallan las técnicas heurísticas más difundidas.

- El método de los ahorros

Clarke y Wright (1964) presentaron un método heurístico que actualmente forma parte de la base para el desarrollo de otros métodos de solución (Gaskell, 1967). Esta heurística lleva ese nombre ya que los autores propusieron el uso de ahorros a nivel de distancias para la solución de un modelo VRP. Oliveira (2004) menciona que el procedimiento indica que dos rutas distintas pueden ser unidas, formando una ruta integral que genere un ahorro como resultado de la combinación, que se calcula de la siguiente manera:

$$s_{ij} = d_{oi} + d_{oj} - d_j \quad (28)$$

El ahorro se calcula de dicha forma ya que en vez de que un vehículo recorra las distancias de los arcos $(0, i)$ y $(0, j)$ por separado, ahora recorrerá la distancia del arco (i, j) (Oliveira, 2004). Además, se considera que es un ahorro si y sólo si s_{ij} es mayor a 0, puesto que si toma un valor nulo será indistinto el camino que se tome y si toma valores negativos, entonces se prefiere que las rutas se mantengan separadas.

El procedimiento inicia desde una solución inicial, generalmente con una ruta por cliente, la cual irá siendo modificada al adicionar combinaciones entre rutas que generen los mayores ahorros. Cada posibilidad de unión será incluida solo si no viola alguna restricción de la formulación del modelo VRP. Finalmente, el método terminará cuando ya no existan más ahorros considerables a incluir en la solución (Oliveira, 2004).

- Heurística de barrido

Es una de las principales heurísticas de la familia de asignar primero, rutear después. En este tipo de técnicas se busca, en primer lugar, obtener *clusters* o grupos de clientes que pertenecerán a la misma ruta o viaje. Luego, se especificará la ruta u orden a seguir para la distribución de materiales a cada *cluster*, asignando en cada viaje a un vehículo y garantizando que no se supere la capacidad de este. Cada ruta se definirá a partir de un TSP ya que se trata de un único vehículo que debe abastecer a cierta cantidad de clientes (Oliveira, 2004).

Específicamente, esta heurística propone que los *clusters* se generen girando una semirrecta que parte desde el vértice “depósito”, de tal manera que sean incluidos todos los clientes que sean barridos por la semirrecta. La conformación de un cluster terminará cuando se llegue al límite, tal que no se supere la capacidad de un vehículo por grupo. Para ello, es necesario que los vértices de clientes se definan a partir de coordenadas polares (p_i, θ_i) , sabiendo que el origen, con coordenadas $(0,0)$, es el vértice “depósito” y de este parten las semirrectas que definirán los *clusters* (Oliveira, 2004). Se presenta la Figura 10, la cual muestra los 4 *clusters* obtenidos aplicando esta heurística, donde se visualiza que no se suelen superponer las rutas.

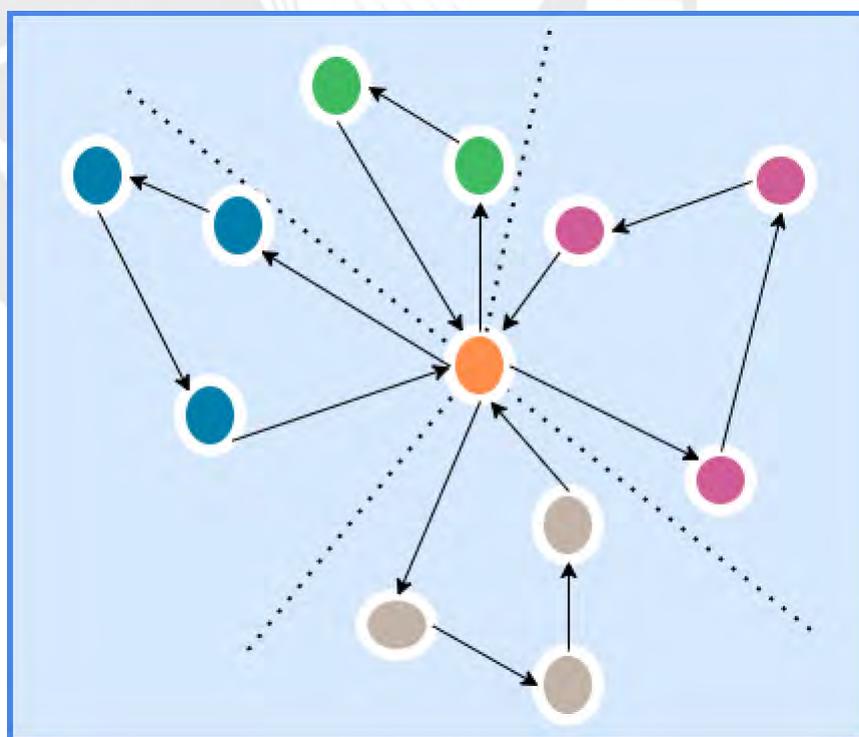


Figura 10. Aplicación de la heurística de barrido en un grafo de enrutamiento de vehículos con capacidad limitada.

Fuente: Oliveira, A. (2004).

El orden en que se seleccionan a los clientes en cada cluster parte de ordenarlos ascendentemente de acuerdo a θ y luego seleccionar aleatoriamente a cierto cliente para partir con la aplicación del procedimiento del barrido hasta superar la restricción de capacidad (Oliveira, 2009). Gráficamente, el barrido se lleva a cabo de manera horaria, ya que se irá eligiendo a cada cliente, dentro de un *cluster*, de menor a mayor ángulo.

B. Metaheurísticas de búsqueda local

Taha (2012) afirma que las metaheurísticas permiten minimizar el entrapamiento en una solución óptima local, lo que no sucede con las heurísticas. Oliveira (2004) agrega que las metaheurísticas refieren a procedimientos que realizan una exploración general del espacio de posibles soluciones, permitiendo obtener mejores resultados en comparación a los obtenidos haciendo uso de las heurísticas.

En este apartado, se realizará una breve explicación de las metaheurísticas más significativas para el enrutamiento de vehículos.

- Metaheurística de búsqueda tabú

Según Taha (2012), la búsqueda tabú se caracteriza por no caer en una solución óptima local debido a que prohíbe, a través de una lista tabú, volver a analizar ciertas soluciones previamente elegidas. Sin embargo, esta lista no es permanente, ya que mantiene a los movimientos de búsqueda previa durante cierto tiempo, denominado periodo de tenencia, por lo que al completar este periodo, se volverá a habilitar el movimiento para futuras búsquedas. De acuerdo con Oliveira (2004), un movimiento se refiere a la actividad que se aplica a una solución actual s_t para llegar a una nueva solución vecina s_{t+1} y los movimientos que se deben evitar, debido a que obtienen las soluciones tabúes, se denominan movidas tabú.

- Algoritmos genéticos para *Vehicle Routing Problem*

De acuerdo con Taha (2012), los algoritmos genéticos (GA) replican el proceso evolutivo biológico para la solución de problemas de optimización. Análogamente, un cromosoma es equivalente a una solución factible, la que contiene una cantidad de genes, los cuales poseen cierta codificación para representar la solución en el algoritmo. Dicha codificación puede ser binaria o numérica, sin embargo, se optará por definir una codificación numérica para representar soluciones de rutas (Taha, 2012).

Además, Oliveira (2004) afirma que un algoritmo genético (GA) trabaja con una población P de posibles soluciones o cromosomas i . Adicionalmente, cada cromosoma se mide bajo su nivel de aptitud, que se representa como una función objetivo, generalmente la función objetivo del modelo del problema a resolver, que en el caso de un Vehicle Routing Problem corresponde a los costos de transporte o a la distancia recorrida y el cromosoma más apto será el que proporcione el menor valor de la función (Taha, 2012).

Es necesario hacer hincapié en que existen tanto operadores evolutivos de selección como de cruzamiento para llevar a cabo el algoritmo. En primer lugar, en la fase de selección, el operador escoge a los miembros de la población. Entonces, dicho operador genera una población reducida, denominada intermedia, que suele priorizar a los cromosomas más aptos, bajo factores probabilísticos. Por ejemplo, existe un operador denominado Selección por Torneo, donde se determina cierta cantidad de cromosomas de manera aleatoria y, luego, se seleccionan a los de menor función objetivo para ser parte de la población intermedia (Oliveira, 2004).

El algoritmo se lleva a cabo a partir de iteraciones, donde para cada una de estas se utilizan operadores evolutivos de cruzamiento que alteran a ciertos miembros de la población, generando una población intermedia nueva o renovada (Oliveira, 2004). Tal como explica Taha (2012), los operadores evolutivos de cruzamiento permiten seleccionar a dos padres P_1 y P_2 de la población, los cuales se reproducirán cruzando sus genes para obtener una descendencia (dos hijos C_1 y C_2), la que finalmente reemplazará a los miembros menos aptos

para obtener la siguiente población, lo que se conoce como algoritmo genético de estado estacionario.

De acuerdo con Oliveira (2004), los operadores de cruzamiento clásico se basan en la realización de n cortes, de manera aleatoria, en la codificación de cada padre para hacer un intercambio de segmentos intermedios y generar a los dos hijos.

Oliveira (2004) considera que este proceso sigue tres fases: selección de la población intermedia, cruzamiento de padres y mutación en los genes. Este procedimiento se repetirá hasta finalizar el algoritmo. Específicamente, la aplicación de este algoritmo para un problema TSP, que es la base de un VRP, sigue los mismos pasos del esquema mencionado por Oliveira, sin embargo, se comentarán algunas técnicas específicas a usar en este modelo:

- Selección de la población intermedia: Para definir la primera solución o ruta a incluir en la población, se debe iniciar desde el nodo de partida y agregar, secuencialmente, un vértice al azar que salga del nodo previamente definido y que no sea redundante, es decir, que cumpla con la condición de generación de una ruta. Si no es posible continuar con el procedimiento, se deberá iniciar con el proceso nuevamente (Taha, 2012).
- Cruzamiento de padres: Para realizar el cruzamiento, es necesaria la elección de dos padres de la población intermedia. Estos dos padres serán los cromosomas más aptos. Luego, para el intercambio de genes, se puede utilizar el procedimiento conocido como cruzamiento ordenado, en el cual se debe seleccionar dos puntos de cruce c_1 y c_2 (donde $c_1 < c_2 \leq n$), de tal manera que se puedan intercambiar los nodos de orden $c_1, c_1 + 1, \dots, c_2$ de los padres P_1 y P_2 para formar parte de la ruta de los hijos C_1 y C_2 , respectivamente. Este tipo de cruzamiento expresa que luego se deberán crear las listas intermedias L_1 y L_2 utilizando los nodos de P_1 y P_2 en el siguiente orden: $c_2 + 1, c_2 + 2, \dots,$

$n, 1, \dots, c_2$. Como siguiente paso, se crearán listas intermedias modificadas L_1' y L_2' , donde L_1' es equivalente a L_1 (nodos de orden $c_1, c_1 + 1, \dots, c_2$ del padre P_2) y de la misma manera para L_2' , solo que con L_2 y el padre P_1 . Finalmente, para obtener C_1 y C_2 se deben asignar los elementos de L_1' y L_2' en las posiciones faltantes, en el siguiente orden: $c_2 + 1, c_2 + 2, \dots, n, 1, \dots, c_1 - 1$. En la posición $c_1, c_1 + 1, \dots, c_2$ se colocan los elementos del padre inverso (si se trata de C_1 , entonces P_2) (Taha, 2012).

- Mutación en los genes: La mutación alude a la modificación de los genes del hijo, que puede ocurrir con cierta probabilidad, donde se deben intercambiar dos vértices aleatorios de la ruta, sin tomar en consideración el nodo de partida (Taha, 2012).

Con respecto a algoritmos genéticos específicamente usados para la solución de problemas de enrutamiento de vehículos, Oliveira (2004) define los siguientes:

- *Genetic Vehicle Representation* (GVR): Una vez que se cuenta con las posibles soluciones del VRP, se establece un método de cruce de P_1 y P_2 donde se toma una subruta r de P_1 , que, utilizando la notación previa, puede ser $(c_1, c_1 + 1, \dots, c_k)$. Posteriormente, se elegirá al cliente w más cercano a c_1 que no se encuentre en la subruta. Si se da que w pertenece a otra ruta en la solución P_2 , por ejemplo $(0, w_1, \dots, w_j, \dots, 0)$, entonces se debe modificar por una nueva ruta fusionada, que se representa como $(0, w_1, \dots, w_j, c_1, c_1 + 1, \dots, c_k, w_{j+1}, \dots, 0)$. Si la ruta no es factible, entonces se debe particionar a dicha ruta, de tal manera que se logre la factibilidad. Con dicho procedimiento, se obtiene el hijo C_1 , mientras que el hijo C_2 es equivalente a P_1 (Oliveira, 2004).

- Algoritmo de Baker y Ayechew: Este algoritmo se basa en la heurística de barrido, con ciertas alteraciones. En primer lugar, se asignan los clientes a cierto vehículo realizando una modificación de la heurística mencionada, donde se asignan índices a los clientes de tal manera que los índices estén cercanos la distancia entre los dos clientes es pequeña. Luego de la clusterización, se debe asignar una ruta para cada grupo de clientes, utilizando un algoritmo de búsqueda local simple 2-opt y luego otro algoritmo denominado 3-opt, los que eliminan conexiones para generar recorridos. Finalmente, partiendo de las soluciones, se debe utilizar el algoritmo de estado estacionario y ciertos operadores tanto de selección, como de cruzamiento y mutación para hallar el resultado final (Oliveira, 2004).

1.3.- Pronósticos

Heizer y Render (2009) definen que realizar pronósticos refieren a la predicción de eventos futuros en base al uso de información histórica. Esto puede implicar una proyección futura mediante algún tipo de modelamiento matemático, la que puede ser intuitiva dependiendo del buen juicio del experto. Ballou (2004) afirma que para tener una correcta planeación y control de la cadena de suministro se requieren tener estimados precisos que se presentan en forma de pronósticos y predicciones. Krajewski (2016) añade que las herramientas para elaborar estos pronósticos pueden surgir desde juicios de experiencia y opiniones de personas conocedoras.

Chase (2008) propone que se deben tomar algunas decisiones importantes desde incluso antes de obtener una predicción: (1) la variable a pronosticar; (2) la técnica de pronóstico, y (3) el software a utilizar. En el próximo apartado se explicarán las técnicas de pronósticos por tipo.

1.3.1.- Tipos de pronósticos (causales, cualitativos, cuantitativos).

Chase (2008) tipifica los pronósticos según las siguientes clasificaciones: cualitativo, análisis de series de tiempo, relaciones causales y simulación. Los métodos cualitativos son más connotativos y se construyen a partir de los juicios de experto y la experiencia de personas conocedoras del proceso. Con respecto al análisis de series de tiempo, este se nutre con información histórica que permite predecir la demanda en el futuro. Asimismo, este análisis puede tomar en consideración distintos factores como las demandas cíclicas, estacionalidad

en el tiempo e influencia de tendencias. El modelo de relación causal, que se resume en el uso de regresión lineal, asume que la demanda tiene una alta correlación con alguna variable independiente del medio. Por último, los modelos de simulación permiten manejar una variedad de asunciones acerca de los parámetros del pronóstico, los cuales también se conocen como *drivers*.

A. Técnicas cualitativas de pronóstico:

Heizer y Render (2009) consideran las siguientes técnicas cualitativas:

1. **Jurado de opinión ejecutiva:** En este caso, las opiniones de un grupo de gurús, sumado al uso de ciertos modelos, se utilizan para realizar una estimación final de la variable de pronóstico.

2. **Método Delphi:** Este método consiste en que ciertos entrevistados formen grupos, que estén ubicados en distintos sitios. Luego, sus juicios son valorados bajo preguntas definidas por un coordinador. Finalmente, se elabora un resumen de respuestas y se realizan reuniones hasta llegar a un consenso en la proyección (Chase, 2008).

3. **Composición de la fuerza de ventas:** A través de dicho enfoque, cada vendedor, en base a su experiencia y conocimiento previo, estima las ventas que tendrá su región o localización. Luego, se revisan las proyecciones para verificar que sean realizadas y posteriormente se agrupan por niveles para tener un pronóstico global .

4. **Encuesta en el mercado de consumo:** Consiste en solicitar información a posibles clientes con la finalidad de conocer sus planes de compra en un futuro cercano. Chase (2008) afirma que este método es útil para determinar la disposición de compra o uso del servicio por medio de encuestas que recopilen datos estructurados. Además, Heizer y Render (2009) añaden que puede ayudar a mejorar el planeamiento de nuevos productos y su diseño.

Otros autores como Chase (2008) añaden a estos métodos otras técnicas como la analogía histórica o las técnicas acumulativas, las cuales no serán discutidas en este trabajo de investigación.

B. Análisis de series de tiempo:

Heizer y Render (2009) definen una serie de tiempo como una secuencia de datos medidos en periodos de igual magnitud como pueden ser días, años o trimestres. Chase (2008) añade que estos modelos se basan en establecer resultados del futuro con base en información histórica. El primer paso para realizar un pronóstico con este método es descomponer la demanda e identificar ciertos patrones que ayuden a seleccionar el método adecuado para pronosticar. Krajewski (2016) establece que existen ciertos patrones básicos en la mayoría de las series de tiempo aplicables al pronóstico de la demanda, los cuales son:

1. **Horizontal:** Los datos varían alrededor de una media constante (Krajewski, 2016).
2. **Tendencia:** Heizer y Render (2009) la definen como el movimiento gradual de los datos en el tiempo, los cuales pueden ser positivos o negativos. Krajewski (2016) añade que este movimiento debe ser sistemático y sostenido a lo largo del tiempo.
3. **Estacional:** Krajewski (2016) denota que es un patrón incremental o decremental que se repite en la demanda dependiendo del periodo en el que se encuentre. Según Heizer y Render (2009) existen seis patrones de estacionalidad: días de la semana, semana del mes, día del mes, trimestre, mes y la semana a la que pertenece el año.
4. **Cíclico:** Heizer y Render (2009) definen a los ciclos como patrones que ocurren cada cierta cantidad de periodos, los cuales son muy importantes para el análisis y el planeamiento a corto plazo, pues están sujetos al ciclo comercial.
5. **Aleatorio:** Heizer, Render y Krajewski se alinean a que son las variaciones imprevisibles de la demanda, los cuales podrían o no tener una causa asignable. No siguen ningún patrón que se pueda identificar y por consiguiente, no pueden ser predecibles.

A continuación, se presenta la representación gráfica de estos patrones, bajo la nomenclatura de Sapag (2014), el cual considera al componente aleatorio como “Componente no sistemático” y no hace alusión al componente horizontal.

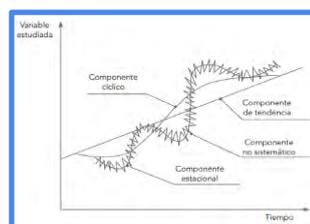


Figura 11. Componentes de la demanda.
Fuente: Sapag (2014).

Krajewski (2016) clasifica los métodos de estimación por series de tiempo en dos grandes grupos, los cuales son: Pronóstico empírico y Métodos con estimación del promedio.

1. Pronóstico empírico:

Krajewski (2016) lo define como un método en el cual la proyección de la demanda para un siguiente periodo es equivalente a la demanda actual del presente periodo. Las ventajas que tiene este método de pronóstico son la simplicidad, costo mínimo y que puede seguir de cerca las tendencias.

2. Métodos con estimación del promedio:

- a. Promedio móvil simple: Chase (2008) afirma que un promedio móvil resulta de mucha utilidad para eliminar el factor aleatorio en el pronóstico, puesto que considera periodos pasados para la estimación. Krajewski (2016) añade que este método resulta de mayor utilidad cuando la demanda no presenta tendencias pronunciadas ni patrón estacional. Para el cálculo del pronóstico, se debe seleccionar el periodo de análisis, ante esto, Chase (2008) denota que esta decisión depende mucho del comportamiento de los datos y existen varios efectos conflictivos. Mientras más grande sea el periodo, habrá mayor uniformidad en los datos pronosticados; sin embargo, esto elimina la posibilidad de retrasar las tendencias. La fórmula por utilizar para el cálculo del promedio móvil simple, siguiendo la notación de Chase (2008), es la siguiente:

$$F_t = \frac{A_{t-1} + A_{t-2} + A_{t-3} + \dots + A_{t-n}}{n} \quad (29)$$

Donde:

F_t = Pronóstico para el siguiente periodo

n = Número de periodos para promediar

A_{t-1} = Ocurrencia real en el periodo pasado

A_{t-2}, A_{t-3} y A_{t-n} = Ocurrencias reales hace dos periodos, tres periodos y hasta n periodos.

- b. Promedio móvil ponderado: Este método permite asignar importancias relativas a cada uno de los elementos, de tal manera que la suma de factores de importancia sea uno (Chase, 2008). Heizer y Render (2009) afirman que puede utilizarse pesos diferenciales para dar mayor importancia a los valores más recientes, puesto que cuando se presenta una tendencia o un patrón observable, es importante acercarse lo más posible a la tendencia en el corto plazo. La selección del periodo se basa básicamente en la explicación anterior, pero cobra mayor importancia en este método, porque existirían un mayor número de ponderaciones que asignar. Se muestra a continuación la fórmula para calcular un pronóstico bajo este método, utilizando la notación de Chase (2008):

$$F_t = w_1 * A_{t-1} + w_2 * A_{t-2} + w_3 * A_{t-3} + \dots + w_n * A_{t-n} \quad (30)$$

Donde:

F_t = Pronóstico para el siguiente periodo

n = Número de periodos para promediar

A_{t-1} = Ocurrencia real en el periodo pasado

A_{t-2} , A_{t-3} y A_{t-n} = Ocurrencias reales hace dos periodos, tres periodos y hasta n periodos.

w_1, w_2, w_3, w_n = Ponderación dada a la ocurrencia real para el periodo

- c. Suavizamiento exponencial: Krajewski (2016) lo define como un modelo ajustado con base en el método de promedio ponderado que permite estimar asignando mayor ponderación a demandas recientes que a demandas en periodos anteriores. Según Chase (2008), lo más resaltante de este modelo es que no se eliminan observaciones anteriores al agregar nuevas piezas de datos y a su vez se toma mayor importancia a las ocurrencias más recientes, lo cual resulta lógico para el autor.

Para realizar un suavizado exponencial se necesita de tres variables importantes: el pronóstico del periodo pasado, la demanda en aquel instante y el parámetro de suavizado o constante de uniformidad (α).

Por otro lado, para poner en marcha el método en cuestión, se requiere un punto de partida o pronóstico base. Krajewski (2016) propone dos maneras de calcular ese pronóstico inicial o base: la primera consiste en utilizar la demanda del periodo inmediatamente anterior, la segunda, solo si se cuenta con información histórica confiable, consiste en calcular el promedio de distintos periodos de demanda en los que se cuenta con información. Se muestra a continuación la ecuación para pronosticar un registro con uniformidad exponencial, según la notación de Chase (2008):

$$F_t = F(t - 1) + \alpha(A(t - 1) - F(t - 1)) \quad (31)$$

Donde:

F_t = El pronóstico para el periodo t

F_{t-1} = El pronóstico hallado el periodo anterior

A_{t-1} = La demanda real del periodo anterior

α = La constante de uniformidad

Los autores concuerdan que este modelo puede ser ajustado para tomar en cuenta algunos de los patrones básicos discutidos líneas arriba, específicamente los de tendencia y estacionalidad. Se presenta un breve alcance de estas variantes a continuación.

c.1. **Inclusión de Tendencia:** Krajewski (2016) afirma que, en este enfoque, se enfoca el pronóstico en el promedio y la tendencia, por lo que se requieren dos constantes de suavizado distintas para realizar el pronóstico. Realizar esta distinción genera que la constante de tendencia afecte directamente a los valores cercanos y este varíe en tanto se presenten tendencias crecientes o decrecientes.

c.2. **Inclusión de Estacionalidad:** Krajewski (2016) denota que se puede incluir la estacionalidad solo limitando los datos tomados para la proyección a periodos de la misma estación, es decir, hacer modelos específicos para cada estación. Otra manera de realizar los pronósticos tomando en cuenta la estacionalidad es aplicando el método estacional multiplicativo, método que consiste en hallar factores estacionales para obtener pronósticos a partir de demandas promedio por estación.

C. Relaciones causales:

Estos métodos deben emplearse cuando se cuenta con información histórica y se puede determinar la relación entre la variable que se intenta pronosticar y otros factores externos o internos (Krajewski, 2016). Chase (2008) los nombra modelos de pronóstico asociativo, ya que afirma que casi siempre se toma en cuenta varias variables relacionadas con la que se quiere pronosticar. Heizer y Render (2009) añaden que, debido al modelo estadístico detrás del modelo asociativo, este enfoque es más potente que los métodos de series de tiempo, ya que estos solo toman en cuenta la información histórica en la variable de pronóstico.

La herramienta más desarrollada de esta tipología es la regresión lineal, en la que una variable que se desea pronosticar está asociada directamente con una o incluso más variables independientes mediante una ecuación de regresión de primer grado (Krajewski, 2016). Chase (2008) añade que antes de realizar el análisis, es necesario graficar una dispersión de variables para verificar visualmente si los datos tienen una relación lineal. Se presenta a continuación la recta de regresión, a partir de la cual se realizará el pronóstico:

$$\hat{y} = a + bx \quad (32)$$

Donde:

\hat{y} = valor de la variable a explicar

a = intersección con el eje

b = pendiente de la recta

x = variable explicativa

Asimismo, Krajewski (2016) afirma que se deben buscar los números de a y b que minimizan la suma de las desviaciones cuadráticas entre los puntos pronosticados por la recta de regresión y valores reales. Se muestra la gráfica de la recta de regresión y sus componentes en la *Figura 12*.

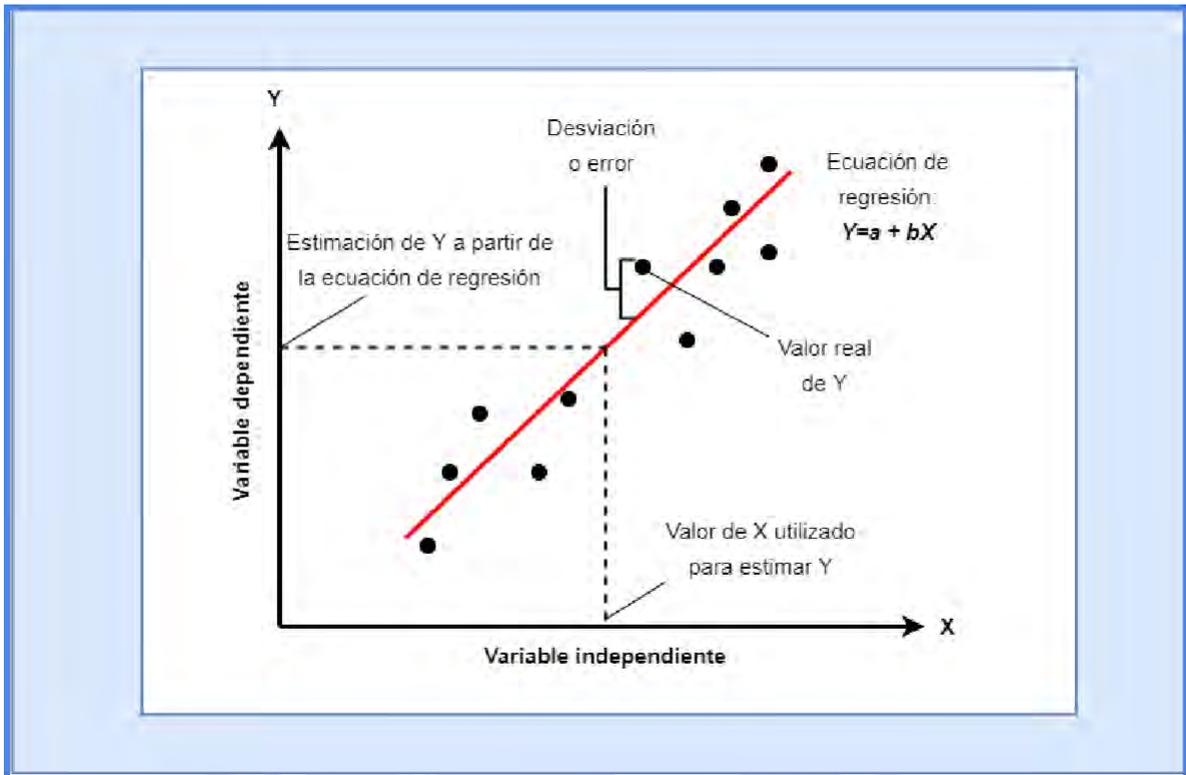


Figura 12. Representación de una regresión.
Fuente: Krajewski (2016).

D. Modelos de simulación:

Chase (2008) los define como modelos dinámicos, soportados por software computacional que permiten hacer asunciones con respecto a las variables internas y las variables externas que pueden afectar al modelo. Se pueden iterar cambios en cualquiera de las variables y verificar el impacto en el pronóstico. Otros autores no consideran esta técnica en su literatura.

1.3.2.- Error de pronósticos

Chase (2008) describe el término error como la diferencia entre el valor pronosticado y el valor real. Además, hace analogía a los residuales analizados en estadística. Krajewski (2016) y Chase (2008) clasifican a los errores de pronóstico en: errores de sesgo o errores aleatorios.

A. Errores de sesgo: Los errores sesgados ocurren como resultado de errores sistemáticos o consistentes, como, por ejemplo, la selección incorrecta de variables, las relaciones erróneas o inexistentes entre variables, la inclusión de un comportamiento de tendencia en una situación incorrecta, entre otros.

- B. Errores aleatorios: Son todos aquellos errores que el modelo no puede explicar, es decir, que son el resultado de factores impredecibles e imprevisibles que generan desviación en los pronósticos realizados.

Medición de errores:

Heizer y Render (2009) denotan que comúnmente se utilizan muchas medidas para calcular el error global en el pronóstico en la industria. Estas mediciones pueden utilizarse tanto como para comparar modelos como para dar seguimiento a los pronósticos, asegurando su correcto funcionamiento. Las tres medidas más comunes son la desviación absoluta media o MAD, el error cuadrático medio o MSE y el error porcentual absoluto medio o MAPE. Krajewski (2016) añade a estas mediciones el cálculo de la suma acumulada de errores de pronóstico o CFE, el sesgo promedio, desviación estándar y la señal de rastreo. Se detalla cada una de estas mediciones a continuación:

- A. **Desviación absoluta media (MAD):** Se define como la suma de valores absolutos de los errores de pronóstico calculados de manera individual divididos entre el número de periodos con información histórica disponible (Heizer y Render, 2009). Chase (2008) afirma que este indicador había sido reemplazado por el cálculo de la desviación estándar pero actualmente es utilizado debido a su eficacia al obtener la señal de rastreo. Se presenta la fórmula utilizada para el cálculo, bajo las notaciones de Chase(2008):

$$MAD = \sum_{i=1}^n \frac{|A_t - F_t|}{n} \quad (33)$$

Donde:

A = Demanda real para el periodo analizado

F = Demanda pronosticada para el periodo de análisis

t = Número del periodo analizado

n = Número total de periodos a analizar

Error cuadrático medio (MSE): Se calcula como el promedio de los cuadrados de las diferencias encontradas entre los valores pronosticados y los observados (Heizer y Render, 2009). Se presenta la fórmula utilizada para el cálculo, bajo las notaciones de Krajewski (2016):

$$MSE = \frac{\sum E_t^2}{n} \quad (34)$$

Donde:

E = Diferencia entre valor pronosticado y observado

t = Número del periodo analizado
n = Número total de periodos a analizar

B. Error porcentual absoluto medio (MAPE): Se define como el promedio de los errores y se expresa como un porcentaje de la suma real. Este indicador elimina el factor de magnitud del elemento que se pronostica, lo que puede ser útil para pronósticos con valores elevados. Se muestra la fórmula con la notación utilizada por Krajewski (2016):

$$MAPE = \frac{(\sum |Et|/Dt)(100)}{n} \quad (35)$$

Donde:

E = Diferencia entre valor pronosticado y observado

D = Demanda pronosticada para el periodo de análisis

t = Número del periodo analizado

n = Número total de periodos a analizar

C. Suma acumulada de errores de pronóstico (CFE): Krajewski (2016) añade este indicador como punto de partida en la medición de errores y lo recomienda en pronósticos con un periodo relativamente largo, pues mide el error total de un pronóstico. Se muestra continuación la fórmula con la notación utilizada por este autor:

$$CFE = \sum Et \quad (36)$$

Donde:

E = Diferencia entre valor pronosticado y observado

t = Número del periodo analizado

D. Error de pronóstico o sesgo promedios (\underline{E}): Krajewski (2016) hace diferencia entre este y el indicador anterior en sentido de comparar proyecciones con periodos variables. Se calcula el indicador de la siguiente manera:

$$\underline{E} = CFE/n \quad (37)$$

Donde:

CFE= Suma acumulada de errores de pronóstico

n = Número total de periodos a analizar

E. Desviación estándar (σ): Se utiliza para medir la dispersión del pronóstico, es decir, la dispersión entre el valor observado y el valor esperado. Se muestra la fórmula a continuación, utilizando la notación de Krajewski(2016):

$$\sigma = \sqrt{\frac{\sum(Et - \underline{E})^2}{n-1}} \quad (38)$$

Donde:

\underline{E} = Error de pronóstico promedio

E = Diferencia entre valor pronosticado y observado

t = Número del periodo analizado

n = Número total de periodos a analizar

F. Señal de rastreo: Chase (2008) afirma que la señal de rastreo es una medida que indica si el promedio calculado cambia de manera sistemática, específicamente indica el número de desviaciones absolutas medias por encima o debajo del valor real en el que se encuentra el valor pronosticado. Krajewski (2016) añade que este indicador mide si un método está previendo con precisión estos cambios mencionados. Se muestra la fórmula a continuación, utilizando la notación de Krajewski (2008):

$$TS = CFE/MAD \quad (39)$$

Donde:

CFE= Suma acumulada de errores de pronóstico

MAD= Desviación absoluta media

1.4.- Data mining

Zaki y Wagner (2020) afirman que la minería de datos y el aprendizaje automático permiten obtener conocimientos fundamentales y conocimiento a partir de datos, así como el desarrollo de novedosos patrones y modelos descriptivos y predictivos que permiten tomar decisiones gerenciales en las compañías. En este apartado, se desarrollarán todos los puntos necesarios para trabajar con datos a gran escala, basados en la teoría de la minería de datos. Aggarwal (2015) añade a la concepción de *Data Mining* los conceptos de preprocesamiento, limpieza y recolección de datos agrupándolos dentro del *Data Mining Process*, el cual se detalla en la *Figura 13*.

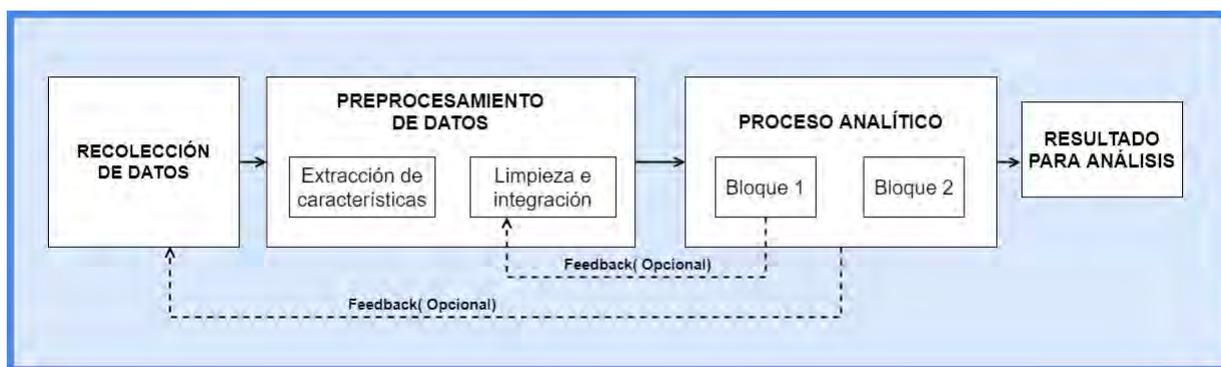


Figura 13. Línea de procesamiento de datos.

Fuente: Aggarwal (2015).

1.4.1. Recopilación de datos

La recolección de datos se puede realizar utilizando un hardware específico, como por ejemplo las redes de sensores especializados o, por otro lado, empleando trabajo manual mediante la recopilación de información como resultado de la investigación primaria, sea realizando encuestas, entrevistas o convocando *focus groups*. Cabe resaltar que esta etapa es altamente específica y suele estar fuera del proceso de minería de datos; sin embargo, es de vital importancia porque impacta en los procesos posteriores del *Data Mining*.

1.4.2. Preprocesamiento de información

Aggarwal (2015) afirma que esta etapa es la más crucial dentro del proceso de minería de datos, sin embargo, es poco valorada, dado que muchos autores se centran en el proceso analítico. Él la separa en tres etapas:

- A. Extracción de características: Por lo general se cuenta con una gran cantidad de información que no siempre es útil en su totalidad, por ello, Aggarwal (2015) propone extraer las características que se consideren relevantes para la aplicación que se desee implementar. Max Bramer (2020) propone una serie de pasos necesarios para la extracción:
 - a. Dimensionar la ganancia de información en cada atributo en la base original.
 - b. Descartar a todos los atributos que no siguen el criterio específico definido
 - c. Usar el conjunto de datos específico definido al algoritmo establecido.

- B. Limpieza de datos: El proceso de limpieza de datos surge debido a errores asociados a las fuentes de datos, dado que estos pueden tener registros erróneos o faltantes, que dificultan el análisis posterior. Bramer (2020) afirma que, en situaciones reales, se pueden registrar datos incorrectos debido a juicios subjetivos o mal uso de

equipos de medición. Estos valores erróneos pueden o no ser valores posibles para el *atributo* que se desee medir, por lo que, podría o no ser tomado en cuenta para el análisis posterior. Aggarwal (2015) clasifica la limpieza de datos en tres aspectos importantes:

a. **Manejo de data faltante:** Muchos registros en los datos pueden permanecer sin especificar, ya sea por errores en la recopilación o por la naturaleza inherente de los datos. El proceso de estimar los datos faltantes es llamado *imputación* (Bramer, 2020).

b. **Manejo de data incorrecta:** Aggarwal (2015) denota que frente a data que aparenta ser incorrecta se puede optar por eliminar registros o corregirlos según tres factores claves. El primer factor refiere al reconocimiento de inconsistencias, por ejemplo, ciertos campos pueden permanecer sin información debido a que existen otros campos que contienen la misma información duplicada. Ante esto, se opta simplemente por duplicar el campo duplicado en el registro faltante o la eliminación de un campo por completo por contener información que no agrega valor para el análisis.

Con respecto al segundo factor clave, Aggarwal (2015) resalta que es importante tener un conocimiento de los dominios que se está manejando, dado que puede haber inconsistencias que invaliden registros por completo. Por ejemplo, campos que refieran a capitales deben corresponder al país en el que se encuentran.

Finalmente, el tercer factor se refiere a los métodos centrados en datos, es decir, al uso de estadística para detectar registros que no representen el comportamiento de los datos. Esto es denominado como *datos anómalos* y su detección representa una técnica analítica muy estudiada por autores en esta rama. Bramer (2020) hace hincapié en la importancia de la detección de anomalías, debido a que considera que, muchas veces, eliminar estos datos resulta peligroso, al restarle conocimiento útil al sistema subyacente. Este análisis será tocado con mayor profundidad en apartados posteriores.

c. **Escalado y normalización:** En muchos escenarios, diferentes características hacen alusión a escalas diferentes de medición, por lo que compararlas resultaría imposible y podría ignorar características con las

menores escalas. Ante esto, Aggarwal (2015) propone dos tipos de normalizaciones; la primera refiere a la estandarización de la distribución normal, restando la media al valor a normalizar y dividiéndolo entre la desviación estándar. El autor hace especial énfasis en que este método de normalización parte bajo el precepto de una distribución normal de los datos y a que los valores normalizados bajo este método suelen no ser representativos en el rango de [-3,3].

Por otro lado, el segundo tipo de normalización se realiza utilizando la *escala mínima-máxima*, es decir, escalar los valores según qué tan desviados están de la diferencia entre el mínimo y máximo de los datos. Este método no es efectivo cuando los valores mínimos y máximos son valores extremos o valores atípicos, dado que minimiza el rango de escalamiento.

- C. Selección de características y transformación: Aggarwal (2015) afirma que cuando los datos tienen dimensiones muy altas, muchos de los algoritmos de minería de datos no funcionan o no son efectivos, por lo que, se debe transformar el conjunto de datos para adaptarlos a un espacio más susceptible de análisis. Por otro lado, Bramer (2020), añade que ciertos atributos o campos particulares pueden transformarse de tal manera que provean de un análisis diferente agregue mayor valor. Por ejemplo, el autor afirma que campos como la edad pueden separarse en rangos específicos para conveniencia analítica.

1.4.3. Proceso de análisis y algoritmos

Esta etapa corresponde a la etapa final en el proceso de minería de datos, la cual consiste en analizar la data procesada mediante algoritmos que permitan solucionar problemas o crear conocimiento innovador (Aggarwal, 2015). Bramer (2020) afirma que las aplicaciones o algoritmos de *data mining* están siempre relacionados con al menos uno de los cuatro *superproblemas* en la minería de datos: patrones de asociación, *clustering*, clasificación y análisis de *outliers* o datos anómalos. En esta sección se detallarán cada uno de los algoritmos siguiendo esta clasificación.

A. Patrones de asociación

Los problemas de *Association Pattern Mining* o modelos de aprendizaje no supervisado, bajo la denominación de Bramer (2020), nacen con la necesidad de encontrar relaciones entre diversos valores de variables, generalmente en forma de reglas conocidas como *reglas de asociación*. Aggarwal (2015) denota que, por lo

general, para aplicar este modelo se trabajan con matrices de datos binarias, de tal manera que la relación entre valores de éxito y fracaso de cada una de las variables pueda generar patrones de recurrencia que sirvan para generalizar el caso a un conjunto más grande de datos, utilizando una medida conocida como la *confianza*.

La *confianza* de una regla de asociación no es más que la probabilidad condicional de que, en una regla $X \Leftrightarrow Y$, un conjunto de datos contenga la característica Y dado una característica X conocida (Aggarwal, 2015).

B. Clustering

Zaki (2020) define al *clustering* como la tarea de dividir datos puntuales y agruparlos en grupos naturales llamados *clusters*, de modo que los puntos dentro de estos grupos reúnan características similares y puntos de diferentes *clusters* sean los más diferenciados posibles. Aggarwal (2015) añade que este método ayuda a resumir la gran cantidad de información con la que se cuenta y permite comprenderla de mejor manera. Cabe resaltar que Bramer (2020) también agrupa a este tipo de algoritmos dentro de los modelos de aprendizaje no supervisado.

Los métodos más conocidos para realizar esta tarea son el algoritmo de k-medias, k-medias de Kernel y algoritmo de *Expectation-Maximization (EM)*, afirma Zaki (2020).

C. Clasificación

Aggarwal (2015) explica que muchos de los problemas en la minería de datos se dirigen hacia un objetivo especializado representado por una característica en particular, denominada etiqueta de clase. Se requiere asignar etiquetas correctas a conjuntos de puntos, lo que es denominado conjunto de entrenamiento o *training set*. El modelo aprendido se conoce como modelo de entrenamiento y los puntos de datos nunca antes vistos que necesitan ser clasificados se denominan *conjunto de datos de prueba*. El algoritmo que crea el modelo de entrenamiento para la predicción también es a veces referido como *el aprendiz*.

Por lo tanto, Zaki (2020) lo denomina aprendizaje supervisado, dado que un grupo de datos se utiliza para aprender la estructura de los grupos más grandes, asemejándose al caso de un maestro supervisando a su estudiante. Entonces, dado un conjunto de datos de prueba, un modelo de clasificación intentará reflejar el número y estructura de los grupos disponibles en el conjunto de datos dentro del conjunto de entrenamiento.

Aggarwal (2015) afirma que estos algoritmos tienen usualmente dos fases marcadas:

- Fase de entrenamiento: En esta fase, el modelo de entrenamiento es construido en base al conjunto de datos de prueba. Esto puede entenderse como un modelo matemático resumido de los grupos etiquetados en el conjunto de entrenamiento.
- Fase de testeo: En esta etapa, el modelo de entrenamiento es utilizado para determinar el identificador de grupo o etiqueta de clase de una o más etapas de la prueba invisibles, es decir, el modelo determinará qué característica o atributo es más adecuado para cada uno de los registros según el aprendizaje obtenido en la fase de entrenamiento.

Existen muchos tipos de metodologías de clasificación, tales como los árboles de decisión, clasificadores basados en reglas, clasificadores probabilísticos, máquinas de vectores de soporte, redes neuronales y aprendizaje basado en instancias; sin embargo, en este apartado solo se hará especial énfasis a las redes neuronales, las cuales serán explicadas en un apartado específico posterior.

D. Análisis de outliers

Aggarwal (2015) afirma que la mayoría de los métodos de detección de *outliers* crean un modelo de patrones normales, es decir, valores que representen un comportamiento normal de los datos.

1.5.- Redes Neuronales Artificiales (RNA)

Estas son algoritmos de clasificación basados en la estructura de una red de neuronas biológicas, puesto que el comportamiento de aprendizaje del algoritmo es similar al del proceso cognitivo humano. Aggarwal (2015) afirma que el estímulo externo en una red neuronal artificial es dado un una data de entrenamiento, o conjunto de datos de entrenamiento. La clave para la efectividad de una red neuronal radica en su arquitectura, desde con una sola capa hasta estructuras más complejas con muchas capas.

Según Zaki (2020) una neurona artificial actual como unidad de procesamiento, la cual recibe señales de entrada mediante una función que suma ponderadamente ciertos factores para generar una variable de salida. Además, existe una neurona binaria la cual controla que la suma ponderada no supere un máximo valor, por lo que, en el caso de ser superada, le

asigna valor de cero y por el contrario, se le asigna el valor de uno. Se muestra a continuación la esquematización de una neurona perteneciente a una red neuronal.

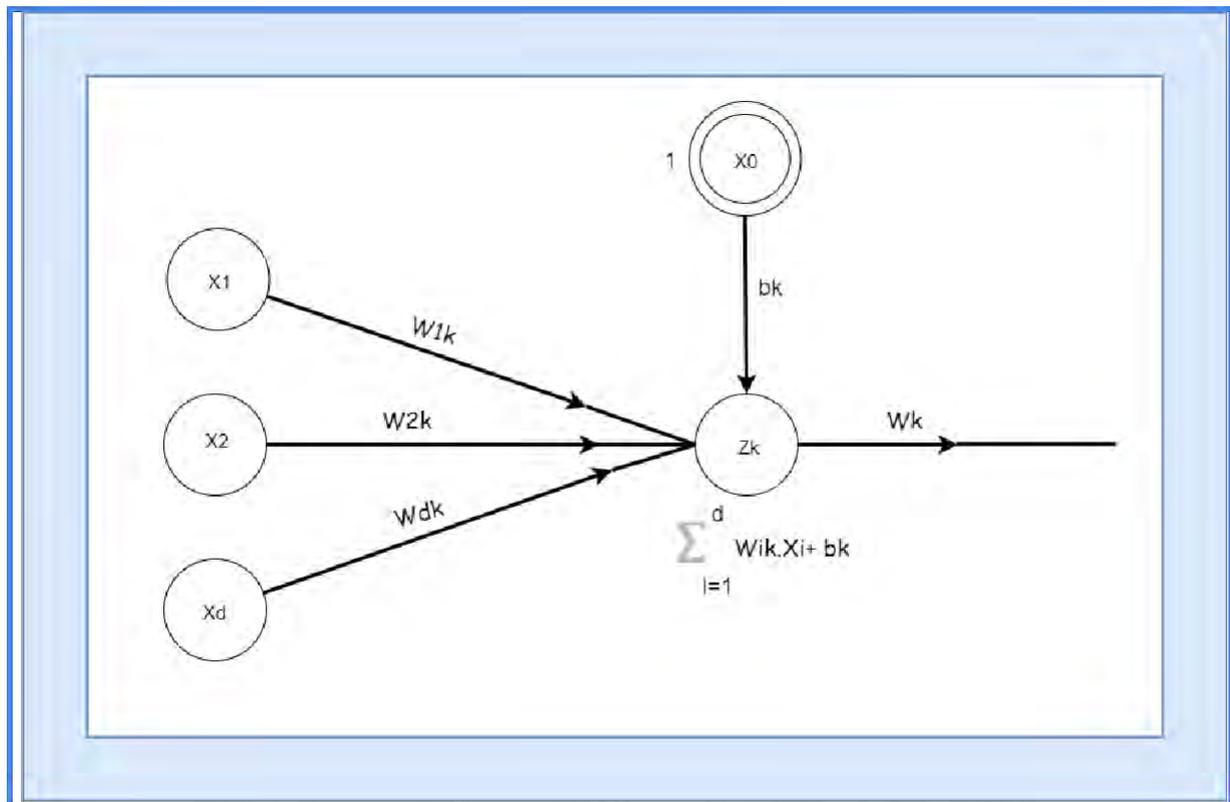


Figura 14. Estructura de una red neuronal.
Fuente: Aggarwal (2015).

El valor de salida se obtiene luego de aplicar una *función de activación* z_k , a la suma ponderada neta, denominada net_k , la cual está compuesta por la suma ponderada de factores más el valor discreto de la neurona de sesgo. La fórmula del net_k se denota de la siguiente manera, bajo los términos de Zaki (2020):

$$net_k = b_k + \sum_{i=1}^d w_{ik} * x_i = b_k + w^T x$$

(40)

Fuente: Aggarwal (2015).

El valor z_k es pasado a otras neuronas, en las cuales funcionará como valor de entrada y será sometida a otras funciones de activación. Zaki define muchos tipos de funciones de activación, las cuales serán usadas en situaciones específicas. Entre las más importantes destacan la función identidad, función por tramos, función lineal rectificadora, función sigmoidea y la función de tangente hiperbólica.

CAPÍTULO 2. ESTUDIO DE CASOS

En este capítulo, se detallarán cinco casos, que permitirán profundizar en los algoritmos y técnicas para la implementación de mejoras en la cadena de suministro del efectivo. Se ha optado por la elección de los siguientes casos:

- Caso 1: El primer caso pone énfasis en la necesidad de presentar modelos integrados de predicción y de optimización de costos, para obtener mejoras más en la administración del efectivo de un banco.
- Caso 2: El segundo caso se enfoca principalmente en presentar una solución real con respecto al diseño óptimo de rutas para el suministro y recojo de efectivo de los cajeros automáticos. En este, se utiliza tanto un algoritmo exacto de ramificación y corte, pero también se propone el uso de heurísticas de clasificación antes de emplear el algoritmo exacto.
- Caso 3: El tercer caso se enfoca en presentar una solución al problema de abastecimiento de cajeros automáticos diseñando una herramienta de predicción utilizando herramientas como las series de tiempo multiplicativas, *clustering* y redes neuronales.
- Caso 4: El cuarto caso pone énfasis en la inclusión del riesgo por robos como factor a tomar en cuenta dentro del proceso de abastecimiento de cajeros automáticos. Se hace uso de teoría de juegos para estimar el riesgo de robo en cada zona y un algoritmo genético híbrido para optimizar las rutas, teniendo en cuenta además periodos de atención y minimización de rutas repetidas.
- Caso 5: El último caso se basa en demostrar que se puede solucionar un problema de predicción utilizando el método de suavizado exponencial y, así, obtener iguales o mejores niveles de precisión que al utilizar métodos computacionales avanzados en el rubro de predicción del efectivo para cajeros automáticos.

2.1.- Caso 1: Optimization of ATM and Branch Cash Operations using an integrated cash requirement forecasting and cash optimization model

2.1.1.- Descripción de la problemática

Este caso se sitúa específicamente en Turquía, sin embargo, la problemática que se describe trasciende los límites nacionales. En particular, el banco Kuveyt Turk, una entidad de mediana escala logró incrementar su nivel de participación en el mercado debido al éxito financiero producto de grandes esfuerzos operacionales y su constante innovación en productos y servicios. Como consecuencia de ello, triplicó la cantidad de oficinas y cajeros automáticos presentes en Turquía, tal que le permita atender a todo el conjunto de clientes

fidelizados y nuevos. Al año 2015, el banco ya contaba con 307 oficinas y 378 cajeros automáticos.

Aunque Kuveyt Turk era uno de los bancos con mayor avance y respaldo del país, aún no sabía cómo resolver la gran problemática que atraviesan muchas cadenas de suministro del efectivo a nivel global. La empresa mantenía grandes cantidades de efectivo almacenadas, es decir, tanto en cajeros automáticos como oficinas, incurriendo en costos debido a la inactividad del efectivo e imposibilidad de utilización para inversiones. En ese entonces, la empresa consideraba únicamente un factor para el abastecimiento de sus sucursales y cajeros automáticos, que es básicamente mantener el alto nivel de servicio al cliente, sin tomar en cuenta la minimización de los costos operativos. Los hechos se demuestran a continuación, a partir de las métricas de Kuveyt Turk antes de la optimización de su gestión del efectivo (Ver Tabla 1):

Tabla 1. Estadísticas de resumen de las oficinas en Estambul y los cajeros automáticos en Turquía de Kuveyt Turk antes de la optimización.

Número de oficinas en Estambul	Número de ATMs y XTMs	Índice promedio de dinero inactivo en oficinas (%)*	Dinero retenido promedio en ATMs (TL)	Ratio promedio de dinero en ATMs (%)**	Costos de transferencia de efectivo (TL por cajero automático)
133	378	6.68%	102.894	1.74%	7.699
*: Calculado como el ratio de la cantidad promedio de dinero inactivo sobre la cantidad total de cuentas corrientes en las oficinas de Estambul					
**: Calculado como el ratio del total de dinero retenido en ATMs sobre la cantidad total de cuentas corrientes					

Fuente: Bilir, C. y Doseyen, A. (2018).

Las estadísticas demuestran que Kuveyt Turk efectivamente se encontraba realizando actividades operativas ineficientes con respecto a la gestión de caudales. Esto se debe a que el índice de dinero inactivo en las oficinas de Estambul era equivalente al 6.68%, mientras que la industria percibía índices menores al 3%, en promedio. Por otro lado, el ratio de dinero almacenado en los cajeros automáticos del país con respecto a la cantidad total de cuentas corrientes ascendía a 1.74%, mientras que la industria, en promedio, alcanzaba índices inferiores al 1%.

2.1.2.- Procedimiento

Para darle solución a tal problemática, se llevó a cabo un procedimiento que se detalla en la Figura 15.

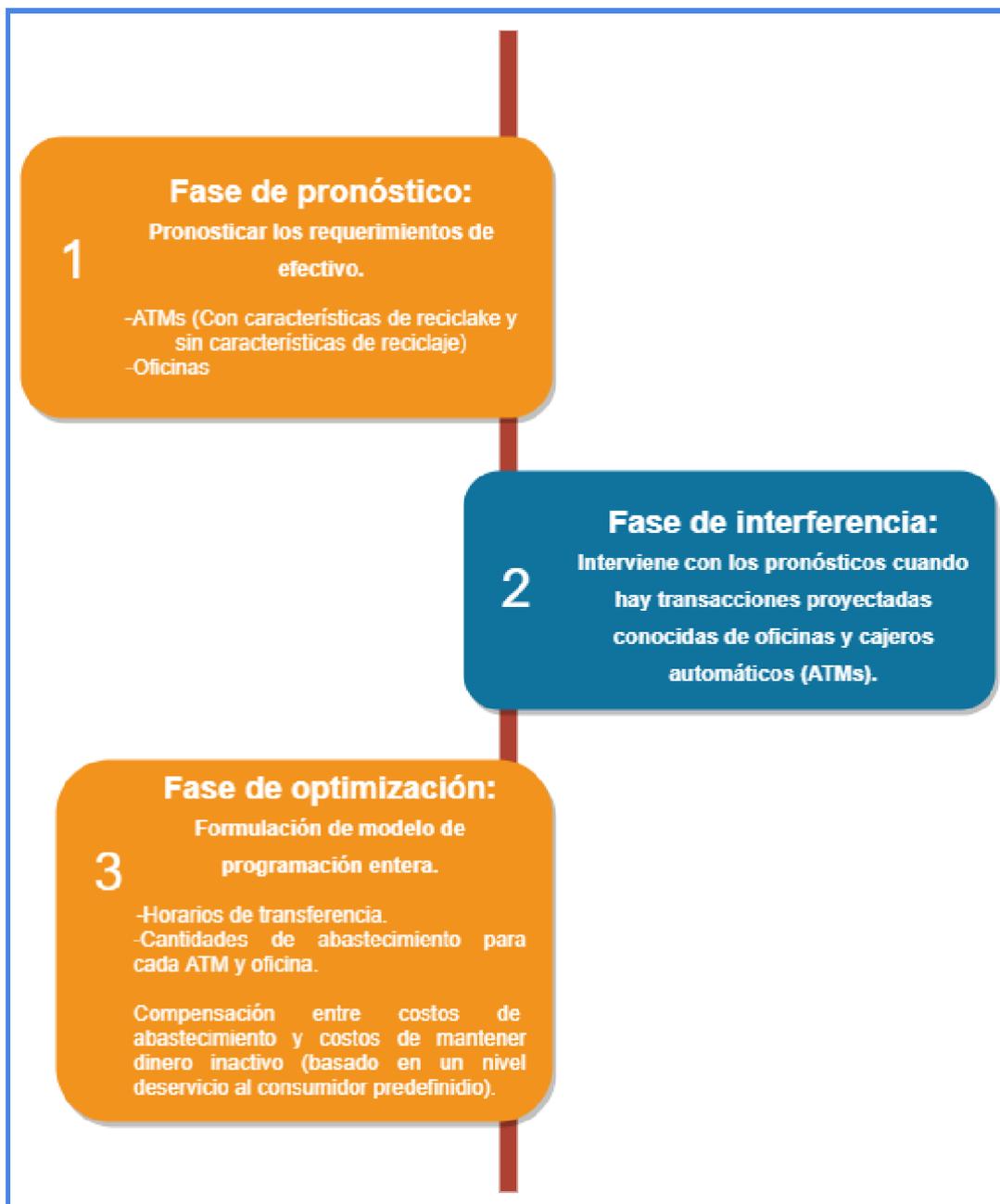


Figura 15. Flujograma del proceso de solución de la problemática.
Fuente: Bilir, C. y Doseyen, A. (2018).

Tal como se visualiza en la Figura 15, el enfoque y metodología a implementar define tres fases. Para la primera, se construye el pronóstico de los requerimientos de efectivo tanto para oficinas como para cajeros automáticos, considerando que existen cuatro tipos de cajeros automáticos en Turquía: cajeros de depósito, cajeros de retiro, cajeros con ambas funcionalidades (pero con caseteras separadas) y los cajeros de reciclaje. Este último tipo de cajeros automáticos posee dos caseteras (una de depósito y otra de retiro), pero a diferencia del de doble funcionalidad, es posible que ambas interactúen, es decir, el efectivo depositado en una de las caseteras puede ser utilizado para el retiro y posterior distribución a clientes, luego de un proceso de validación. Después de finalizada la primera fase, es

posible alterar los pronósticos realizados debido al conocimiento de futuras alteraciones en la demanda debido a cambios en los hábitos del cliente que generen transacciones inusuales. Finalmente, en la última fase, se incluye la optimización para llevar a cabo la obtención del horario de transferencia de los caudales a cada oficina y cajero automático de la compañía.

2.1.3.- Aplicación de herramientas de predicción y optimización

Para la explicación, se dividirá esta sección en dos apartados, ya que Bilir y Doseyen (2018) dividen su solución en fases distintas.

2.1.3.1.- Aplicación de herramientas de predicción

Durante la fase de predicción, se implementó un enfoque de *machine learning* o aprendizaje automático para conocer las transacciones futuras de efectivo de los clientes de oficinas y de cajeros automáticos de Kuveyt Turk. Dicha metodología toma en cuenta las estacionalidades, es decir, considera ciertos patrones que pueden afectar los resultados de demanda, como el día de la semana, mes del año, feriados, cercanía a las vacaciones, periodos de pago de salarios, etc. Dicha estacionalidad se evidencia a partir de índices estacionales, los que son fundamentales para obtener predicciones acertadas del efectivo. También, se aplicaron factores para cada cajero automático, ya que la demanda de un ATM se ve alterada por variables como el nivel de proximidad al centro de la ciudad y nivel de proximidad a los perceptores de salario, las que fueron consideradas en el modelo.

2.1.3.2.- Aplicación de herramientas de optimización

Para la fase de optimización, se formuló un modelo de programación entera con una función objetivo basada en minimizar el costo total de las operaciones. Dicho costo total se compone de dos elementos, para el caso descrito: costo de oportunidad por mantener efectivo inactivo y costos de transferencia de efectivo desde o hacia oficinas o cajeros automáticos. Con respecto al primer componente, este se define a partir de una tasa de interés diaria, la que el banco podría generar si es que el dinero almacenado pudiese ser invertido. Además, se añade un nuevo componente a la función con el objetivo de mejorar la satisfacción del cliente, el que exige que también se minimice la cantidad de *cash-outs*.

Las variables que se definieron en el modelo de optimización fueron: Monto total de dinero retirado y depositado por los clientes en cada sucursal y cajero, cantidad total

a retenerse en cada sucursal y cajero, cantidad total de dinero a transferirse desde y hacia cada cajero automático y cada sucursal. También, existen ciertos parámetros como el nivel de servicio deseado, costo unitario de transferir efectivo y la tasa de interés asociada al costo de mantener efectivo inactivo, que fue previamente explicada.

Finalmente, se definen las restricciones producto de la formulación del modelo de optimización, que son las siguientes: restricciones de capacidad de las caseteras de depósito y retiro, restricciones de saldo de cada casetera de cada cajero automático, restricciones de saldo de la bóveda (almacén) de cada sucursal, restricción de disponibilidad y transporte, restricción para mantener cierto nivel de servicio y restricción de los días en los que se pueden realizar transferencias.

2.1.4.- Resultados obtenidos

La implementación de la solución de predicción y optimización se obtuvo en cuatro fases. Primero, se aplicó una prueba preliminar en 40 cajeros automáticos a finales del año 2015 y a inicios del 2016, se realizó la implementación considerando a todos los cajeros automáticos del banco de estudio. Luego, inició con la implementación en sucursales, pues el modelo integral se aplicó en 37 sucursales durante el mes de febrero del 2016. Finalmente, la fase final permitió implementar el modelo en las 133 sucursales de Estambul para abril del 2016.

Para medir los resultados obtenidos, se pusieron a prueba las siguientes mediciones: nivel de precisión de los pronósticos, influencia en el número de *cash-outs*, medidas de niveles de efectivo inactivo y el ahorro total de costos. A partir de la implementación diaria de los modelos de predicción y optimización para una mejora en la gestión del efectivo del banco Kuveyt Turk, se obtuvo, en primer lugar, que el MAPE promedio para los seis primeros meses de implementación del modelo de predicción fue de 26.44% (Ver detalle mensual en la Tabla 2). Como método de comparación del error porcentual absoluto medio, se obtuvo en la literatura que este indicador se encuentra, en promedio, entre 20% y 45% para la previsión de requerimientos de efectivo (Ekinici, Lu y Duman, 2015). Bilir y Doseyen (2018) hacen hincapié en que este rango suele aumentar cuando los modelos se aplican a situaciones de la vida real, donde se deben tomar en cuenta estacionalidades, por lo que el indicador promedio de precisión se puede considerar como alto.

Tabla 2. Errores porcentuales absolutos medios mensuales para la medición de la precisión de las predicciones de requerimientos de efectivo.

Mes	MAPE (%)
Mes I	26.22
Mes II	25.28
Mes III	27.68
Mes IV	26.88
Mes V	26.97
Mes VI	25.43
Promedio	26.43

Fuente: Bilir, C. y Doseyen, A. (2018).

En segundo lugar, se obtuvo que la implementación generó una ligera disminución de cantidad de *cash-outs* promedio mensuales en 0.07, medida que se registró para 169 cajeros automáticos. En la Tabla 3, se visualizan los resultados de 7 meses, donde se hace una comparativa por tipo de mes entre los incidentes registrados antes y después de la implementación por cajero automático:

Tabla 3. Estadísticas de resumen de las oficinas en Estambul y los cajeros automáticos en Turquía de Kuveyt Turk antes de la optimización.

	Número promedio mensual de incidentes cash-outs (por ATM)							Promedio de 7 meses
	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	
Antes de la implementación (Enero 2015 hasta Julio 2015)	7.1	7.0	8.0	8.2	7.4	8.3	8.5	7.8
Después de la implementación (Enero 2016 hasta Julio 2016)	8.1	7.2	2.7	9.4	9.6	9.1	8.1	7.7

Fuente: Bilir, C. y Doseyen, A. (2018).

Como otra medida comparativa de los resultados, se calcularon los niveles de efectivo inactivo después de la implementación, tanto para sucursales como para cajeros automáticos. Gracias a la aplicación de esta solución, el banco Kuveyt Turk pudo reducir en 22.9% los niveles de efectivo inactivo en las sucursales de Estambul (comparativa realizada entre el mes de enero de 2015 y enero de 2016). Considerando una tasa promedio del 10% para medir los costos de oportunidad, se obtiene que la implementación permitió ahorrar 4.4 millones de liras turcas (TL) anuales, que ahora pueden ser destinados a la tesorería del banco para ser utilizados para fines de inversión. Con respecto a los cajeros automáticos, se

obtuvo una disminución del 22.2% en efectivo inactivo, lo que se traduce en 12.8 millones de liras turcas (TL). Considerando el mismo costo de oportunidad, la implementación en cajeros permitió reducir dicho costo en 1.28 millones de TL anuales.

2.1.5.- Conclusiones

A partir de este caso, se concluye que es necesario considerar el efecto de estacionalidad en modelos de predicción de requerimientos de efectivo en sucursales y cajeros automáticos, con la finalidad de obtener resultados más certeros, que puedan ser aplicados a la banca. Asimismo, un modelo integrado, tanto de predicción y optimización, genera mayores impactos en la gestión del efectivo, tanto a nivel de reducción de niveles de efectivo inactivo y número de *cash-outs* como a nivel de costos operativos. Por último, tal como mencionan Bilir y Doseyen (2018), hasta el año 2018, no existía ningún estudio que incluya un modelo de enrutamiento de vehículos y un modelo de optimización de efectivo al mismo tiempo. La integración aumentaría la eficacia de las cadenas de suministro de efectivo (CSC).

2.2.- Caso 2: An Inventory-Routing Problem with Pickups and Deliveries Arising in the Replenishment of Automated Teller Machines

2.2.1.- Descripción de la problemática

La problemática se sitúa en Países Bajos, país donde los operadores de transporte se enfrentan día a día a la función de distribuir y recolectar efectivo o caudales de cajeros automáticos, tomando en consideración que, en la actualidad, prima la existencia de cajeros automáticos de recirculación (RATM) en el país, máquinas que son capaces de recibir y suministrar billetes, además de comprobar su autenticidad para permitir posteriormente el retiro del efectivo. Ante esta situación, los operadores logísticos o transportistas se ven en la necesidad de diseñar y programar rutas de acuerdo con sus propios criterios, como la experiencia, intuición o criterios cuantitativos, por lo que muchas veces incurren en altos costos de transporte, dado que no toman en cuenta la minimización de las distancias recorridas, que les permita ahorrar en costos asociados a la actividad de recogidas y entregas.

La problemática no aborda únicamente el ruteo para los cajeros automáticos de cierto banco del país, sino que toma en consideración a todos los cajeros automáticos de los Países Bajos, por lo que existen más de 6000 clientes (tanto cajeros automáticos regulares como de recirculación) y 32 centros subregionales de distribución de efectivo, cada centro operando con un vehículo. En este caso, el efectivo se transporta desde y hacia los centros, los cuales

atienden tanto oferta como demanda de efectivo de una subregión específica de Países Bajos. Los centros, ubicados en el mapa de los Países Bajos, se visualizan en la Figura 16.

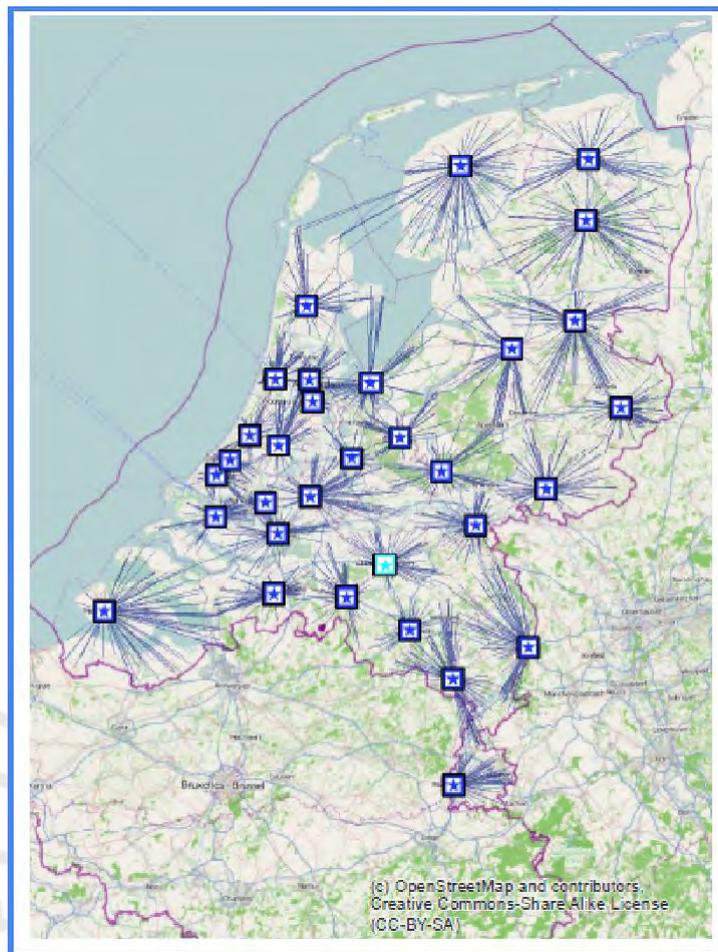


Figura 16. Línea de procesamiento de datos.

Fuente: Van, R. y Coelho, L. (2015).

2.2.2.- Procedimiento

Para el análisis de la problemática, en primer lugar, los autores reconocen la necesidad de especificar el tipo de *Vehicle Routing Problem* al que se están enfrentando. Tal como menciona Van y Coelho (2015), se debe resolver un problema de enrutamiento de inventario con recolecciones y entregas. Entonces, este problema combina dos tipos de clasificaciones: el problema de enrutamiento de inventario (*Inventory Routing Problem*) y el problema de recolecciones y entregas (*Pick-up and Delivery Problem*). Para este último problema, es posible definir varias categorías, donde 1-M-1 hace alusión a una estructura que se enfoca en los movimientos de productos desde el depósito hasta los RATM y luego al depósito de regreso. Otra categoría es M-M, la que refiere a los movimientos de productos entre RATM. En este caso, se trata de un modelo *Pick-up and Delivery* con las dos estructuras, es decir de tipo 1-M-M-1, ya que se deben analizar los movimientos desde los

almacenes hacia RATM, luego visualizar las transferencias de efectivo entre ellos, para finalmente analizar los movimientos de RATM hacia almacenes.

Como siguiente paso, el caso aborda la formulación del modelo matemático que representa la problemática antes definida. Finalmente, para hacer frente al problema, se detalla la solución del modelo a partir de la aplicación de un algoritmo exacto de ramificación y corte, además de la implementación de una heurística de agrupamiento para agilizar la solución posterior con el algoritmo exacto. Estos últimos pasos se detallan en el apartado 2.2.3, donde se realiza una aplicación directa de la disciplina de investigación de operaciones.

2.2.3.- Aplicación de técnicas de optimización

Este apartado se subdivide en tres secciones: Formulación del modelo, aplicación del algoritmo de ramificación y corte, así como la aplicación de heurística de clasificación. Dicha división se debe a que las herramientas de solución se aplican en ese orden.

2.2.3.1- Formulación del modelo

En primer lugar, existen ciertos parámetros como los costos de mantenimiento de inventario en cada RATM (α_i), la capacidad de inventario de cada RATM (C_i), el costo de manejo de inventarios (β) que se incurre en el depósito, la longitud del horizonte de planeamiento (p) con periodos de tiempo discretos (t), la cantidad de vehículos disponibles para alquilar (k), la capacidad de cada vehículo (Q_k), la velocidad promedio (s_k), el costo de alquiler del vehículo por periodo (γ_k), la máxima duración de trabajo del vehículo (S), la penalización por hora extra (δ), el costo asociado al arco (i,j) (c_{ij}), el inventario al inicio del horizonte de planeamiento (I_i^0) y la demanda de cada RATM i para cada periodo t (d_i^t)

Para esta sección, también se incluye la definición de las variables involucradas en el modelo, las que se clasifican por tipo:

- Se incluyen cuatro variables binarias: La variable de utilización de arcos (x_{ij}^{kt}), que equivale a 1 si el arco (i,j) es utilizado en la ruta del vehículo k en el periodo t (que corresponde a un periodo de tiempo dentro del horizonte de planeamiento p); la variable de visitas (y_i^{kt}), que equivale a 1 si y sólo si el RATM i es visitado por el vehículo k en el periodo t ; variable de entregas (

w_i^{kt}), que es equivalente a 1 si y sólo si se realiza una entrega hacia el RATM i por el vehículo k en el periodo t ; y la variable de recogidas (z_i^{kt}), que equivale a 1 si se recoge efectivo del RATM i por parte del vehículo k en el periodo t .

- Además, se incluyen variables de tipo enteras: La variable de nivel de inventarios en cajeros (I_i^t), que refiere a la cantidad de inventarios que posee cada RATM i al final del periodo t ; la variable de cantidades entregadas (q_i^{kt}), que representa a la cantidad de efectivo entregado al RATM i usando el vehículo k en el periodo t ; y la variable de cantidades recogidas (p_i^{kt}), que representa a la cantidad de efectivo recogido del RATM i usando el vehículo k en el periodo t .
- Por último, se toman en consideración las siguientes variables: La cantidad de inventario llevado por el vehículo k en el periodo t fuera y dentro del almacén, que se representan bajo J_k^t y H_k^t , respectivamente. Además, la carga del vehículo k , después de atender al RATM i en el periodo t , que se representa bajo u_i^{kt} . Finalmente, se toma en cuenta a E_k^t , que representa el número de minutos extra que el vehículo k trabaja en el periodo t , es decir, por encima de la máxima duración S .

Luego, se formula la función objetivo del modelo, tal como se visualiza a continuación:

$$\text{minimizar } \sum_{i \in V'} \sum_{t \in T} \alpha_i I_i^t + \sum_{k \in K} \sum_{t \in T} \beta (J_k^t + H_k^t) + \sum_{k \in K} \sum_{t \in T} \gamma_k y_o^{kt} + \sum_{k \in K} \sum_{t \in T} \delta E_k^t \quad (41)$$

Fuente: Van, R. y Coelho, L. (2015).

Tal como Van y Coelho (2015) expresan, se tiene como objetivo minimizar el costo total de mantener inventarios en los cajeros, manejar inventarios en los depósitos y alquilar vehículos. Aquí, se agrega un último término que minimiza los costos de penalización por horas extra.

Finalmente, se definen las restricciones del modelo, que se pueden visualizar a continuación:

$$I_i^t = I_i^{t-1} + \sum_{k \in K} q_i^{kt} - \sum_{k \in K} p_i^{kt} + d_i^t, i \in V', t \in T \quad (42)$$

$$0 \leq I_i^t \leq C_i, i \in V', t \in T \quad (43)$$

$$\sum_{j \in V, i < j} x_{ij}^{kt} + \sum_{j \in V, j < i} x_{ji}^{kt} = \vartheta_i^{kt}, i \in V, k \in K, t \in T \quad (44)$$

$$\sum_{i \in \vartheta} \sum_{j \in \vartheta} x_{ij}^{kt} \leq \sum_{i \in \vartheta} y_i^{kt} - y_m^{kt}, \vartheta \subseteq V, k \in K, t \in T, m \in \vartheta \quad (45)$$

$$w_i^{kt} \leq \varphi_i^{kt}, i \in V', k \in K, t \in T \quad (46)$$

$$z_i^{kt} \leq \varphi_i^{kt}, i \in V', k \in K, t \in T \quad (47)$$

$$q_i^{kt} \leq w_i^{kt} (C_i - I_i^t), i \in V', k \in K, t \in T \quad (48)$$

$$p_i^{kt} \leq z_i^{kt} I_i^t, i \in V', k \in K, t \in T \quad (49)$$

$$w_i^{kt} + z_i^{kt} \leq 1, i \in V', k \in K, t \in T \quad (50)$$

$$s_k \sum_{(i,j) \in \Lambda} c_{ij} x_{ij}^{kt} + r \sum_{i \in V'} y_i^{kt} \leq S + E_k^t, k \in K, t \in T \quad (51)$$

$$u_j^{kt} \leq u_i^{kt} + p_j^{kt} - q_j^{kt} x_{ij}^{kt}, i \in V, j \in V, k \in K, t \in T \quad (52)$$

$$0 \leq u_i^{kt} \leq Q_k, i \in V, k \in K, t \in T \quad (53)$$

$$J_k^t = u_0^{kt} - u_0^{kt}, k \in K, t \in T \quad (54)$$

$$H_k^t = \sum_{i \in V'} x_{i0}^{kt} u_i^{kt}, k \in K, t \in T \quad (55)$$

$$q_i^{kt}, p_i^{kt}, J_k^t, H_k^t \geq 0, i \in V', k \in K, t \in T \quad (56)$$

$$x_{ij}^{kt} \in \{0, 1\}, (i, j) \in \Lambda, k \in K, t \in T \quad (57)$$

$$y_i^{kt} \in \{0, 1\}, i \in V, k \in K, t \in T \quad (58)$$

$$w_i^{kt}, z_i^{kt} \in \{0, 1\}, i \in V', k \in K, t \in T \quad (59)$$

Fuente: Van, R. y Coelho, L. (2015).

Finalmente, Van y Coelho (2015) mencionan que el modelo no es lineal debido a las restricciones (48), (49), (52), (54) y (56), pero que estas se pueden reescribir para garantizar la linealidad.

2.2.3.2- Aplicación del algoritmo de ramificación y corte

Para este caso, se implementó este algoritmo exacto con el objetivo de tener una solución del modelo planteado. La mayoría de las restricciones se pudieron generar, sin embargo, la restricción (45), que garantiza la eliminación de *subtours*, se genera a partir de cortes a medida que se viole dicha restricción. En primer lugar, se elige un nodo aleatorio del árbol y se resuelve un programa lineal, luego se realiza una búsqueda de las restricciones que se infringen y estas se agregan al programa para inicializar la optimización. El proceso itera hasta que se obtiene una solución factible razonable o hasta que no se pueden realizar más cortes al programa.

2.2.3.3- Aplicación de heurística de clasificación

Asimismo, dado que el caso se aplica a los cajeros automáticos de reciclaje de todo el país, el tamaño del modelo es equivalente a más de 200 clientes. Entonces, es necesario brindar soluciones para tamaños realistas, como la de Van y Coelho (2015). Los autores diseñaron una heurística de agrupamiento que brinda mayores facilidades al algoritmo de ramificación y corte, ya que esta técnica se ejecuta previamente. Tal como su nombre lo menciona, la heurística se encarga de clasificar a cada RATM o cliente en uno de los 5 subconjuntos habilitados. *U* es el conjunto de recojos; *W*, el de entregas; *X*, el de posible recojo; *Y*, el de posible entrega y, por último, *Z* es el conjunto de las no visitas. A partir de estas definiciones, cada RATM será ubicado en algún conjunto, de tal manera que, al momento de ejecutar el algoritmo de ramificación y corte, el tiempo de cálculo disminuya. Esto se debe a que ya se conocerá previamente si un RATM necesita un recojo de dinero, una entrega de efectivo, una posible recogida o posible entrega. Entonces, esta heurística proporciona un beneficio, limitando la elección de recojo o entrega y simplificando las posibles soluciones a visualizar en el modelo de ramificación y corte.

2.2.4.- Solución de la problemática y resultados obtenidos

Se llevó a cabo la solución utilizando el lenguaje C++ y una tecnología de IBM, a partir del solucionador CPLEX 12.5.1. A partir de los resultados, se demuestra que el algoritmo planteado es estable, sin embargo, existe una brecha de optimalidad alta aproximada del 51%, con un máximo de 62% y mínimo de 45%, considerando que la brecha de optimalidad refiere a la diferencia porcentual entre el límite superior e inferior del valor objetivo.

Aplicando la técnica de agrupamiento, 7 instancias de las 32 fueron resueltas hasta su solución óptima, mientras que las demás arrojaron una brecha promedio inferior al 0.1%,

entre sus límites superior e inferior de solución. A continuación, se observan los resultados y cambios al aplicar la heurística en la Tabla 4.

Tabla 4. Brecha entre límites para la solución aplicando heurística de clasificación.

Instancia	Límite superior (general)	Límite superior (clustering)	Límite inferior (clustering)	Brecha (%)
1	37290	25699	25643	0.2
2	34077	23891	23850	0.2
3	35899	25595	25577	0.1
4	33963	24989	24950	0.2
5	30506	24755	24755	0.0
6	34215	25421	25395	0.1
7	33985	24979	24964	0.1
8	31838	25068	25054	0.1
9	43821	25191	25182	0.0
10	31956	25071	25071	0.0
11	39340	24648	24641	0.0
12	29866	24314	24302	0.0
13	32565	23664	23645	0.1
14	28124	26499	26499	0.0
15	30685	25706	25706	0.0

16	31621	24256	24166	0.4
17	33620	24832	24817	0.1
18	37026	24503	24458	0.2
19	28201	24709	24686	0.1
20	31526	24752	24738	0.1
21	33029	25344	25314	0.1
22	31670	24733	24686	0.2
23	29386	24752	24738	0.1
24	36755	25013	24961	0.2
25	31768	24157	24135	0.1
26	31971	24744	24744	0.0
27	28715	23772	23764	0.0
28	33453	25344	25344	0.0
29	35922	25853	25853	0.0
30	30079	24530	24526	0.0
31	32504	24082	24031	0.2
32	33662	25340	25314	0.1

Fuente: Van, R. y Coelho, L. (2015).

A partir de la tabla, se puede observar que al utilizar la heurística previamente, se visualiza una gran disminución del límite superior de valores objetivos para todas las instancias, lo

que significa que se genera una elevada mejora debido a la aplicación (Ver Tabla 5). En muchos casos (7), se observa el resultado final mínimo de los costos asociados, mientras que, en otros, la brecha es tal, que existen valores aproximadamente cercanos al óptimo.

Tabla 5. Mejora de resultados al aplicar heurística de clasificación.

Instancia	Límite superior (general)	Límite superior (clustering)	Mejora respecto al caso general (%)
1	37290	22539	39.6
2	34077	23891	29.9
3	35899	25595	28.7
4	33963	24989	26.4
5	30506	24755	18.9
6	34215	25421	25.7
7	33985	24979	26.5
8	31838	25068	21.3
9	43821	25191	42.5
10	31956	25071	21.6
11	39340	24648	37.4
12	29866	24314	18.6
13	32565	23664	27.3

14	28124	26499	5.8
15	30685	25706	16.2
16	31621	18608	41.2
17	33620	21678	35.5
18	37026	22236	40.0
19	28201	24709	12.4
20	31526	19205	39.1
21	33029	18862	42.9
22	31670	24733	21.9
23	29386	19239	34.5
24	36755	19418	47.2
25	31768	19722	39.3
26	31971	19722	38.3
27	28715	23772	17.2
28	33453	20029	40.1
29	35922	21382	40.5
30	30079	24530	18.5
31	32504	24082	25.9
32	33662	21011	37.6

Fuente: Van, R. y Coelho, L. (2015).

Gracias a esta aplicación, se pudo mejorar, en promedio, un 49.9% de la brecha presente entre los límites inferiores y superiores de la solución del modelo, lo que genera que, en conjunto, soluciones estables y efectivas de diseño de rutas de abastecimiento y recogida de efectivo en Países Bajos.

2.2.5.- Conclusiones

Se concluye que es de vital importancia la aplicación previa de una heurística de clasificación si el problema de enrutamiento de vehículos cuenta con una extensa cantidad de clientes. En este caso, se pudo observar una mejora promedio del 49.9% entre la brecha de los límites y se pudo obtener la optimalidad en varias instancias, gracias a la clasificación previa de los nodos en conjuntos.

2.3.- Caso 3: Cash demand forecasting in ATMs by clustering and neural networks

2.3.1.- Descripción de la problemática

La problemática se sitúa en una institución calificada entre los 700 bancos más importantes del mundo ubicada en Inglaterra. Se analizó que la precisión en los pronósticos de consumo en los cajeros automáticos tenía una implicancia directa en los costos, ya que, se incurrían en costos adicionales por abastecimientos de emergencia si es que los pronósticos eran demasiado bajos y costos elevados en posesión de inventarios si es que los pronósticos estaban sobreestimados. Estos costos representaban alrededor de 35% o 60% de los costos totales de funcionamiento de un cajero automático y los errores en pronósticos estaban generando pérdidas de clientes por insatisfacción debido al mal servicio, el cual estaba directamente asociado a la indisponibilidad de efectivo.

Además, se identificó que existían ciertos patrones de estacionalidad que no estaban siendo analizados, dado que, se comprobó que los jueves y viernes había mucha mayor afluencia de clientes en los cajeros automáticos. Sin embargo, los métodos de pronósticos no reflejaban este comportamiento, lo que representaba un gran problema para la empresa.

2.3.2.- Procedimiento

Para el análisis de la problemática, los autores identifican la necesidad de agrupar los cajeros automáticos con patrones similares debido a que les resulta difícil realizar un análisis para cada cajero y consideran que podrían existir cajeros con comportamientos similares. El

procedimiento empieza con la construcción de un modelo de series de tiempo multiplicativo para analizar los montos retirados diarios de cada cajero automático, de manera que reorganizados según el día de semana, puedan ser analizados desde la perspectiva de una serie de tiempo. Posteriormente, se modela la tendencia de serie de tiempo utilizando regresión lineal y se asume que no hay efectos cíclicos ni irregulares en la demanda para poder hallar el factor estacional.

El siguiente paso consiste en discretizar los índices obtenidos de tal manera que se conviertan en un atributo clusterizable o agrupable. En otras palabras, se tradujo la secuencia continua de valores del índice estacional en valores discretos al agruparlos en intervalos por día de la semana y asignarles una etiqueta de clase. Posterior a este paso, se calculó el puntaje de distancia propuesto por Levenshtein utilizando el método SAM o de secuencia alineada, el cual expresa mediante un indicador qué tan similar es un centro de un cajero automático con respecto a su etiqueta de clase asignada. Siguiendo con el proceso y utilizando la distancia SAM calculada, se aplica el algoritmo de *clustering* de Taylor-Butina, de tal manera que quedan agrupados los cajeros según los valores de estacionalidad en cada día similares.

Finalmente, para cada grupo de cajeros automáticos se le aplica cuatro tipos de redes neuronales para realizar el pronóstico de la demanda o consumo diario en los cajeros automáticos. Cabe resaltar que cada tipo de red neuronal aplicada será detallada en la aplicación de herramientas. Se muestra a continuación el diagrama de bloques del procedimiento descrito.

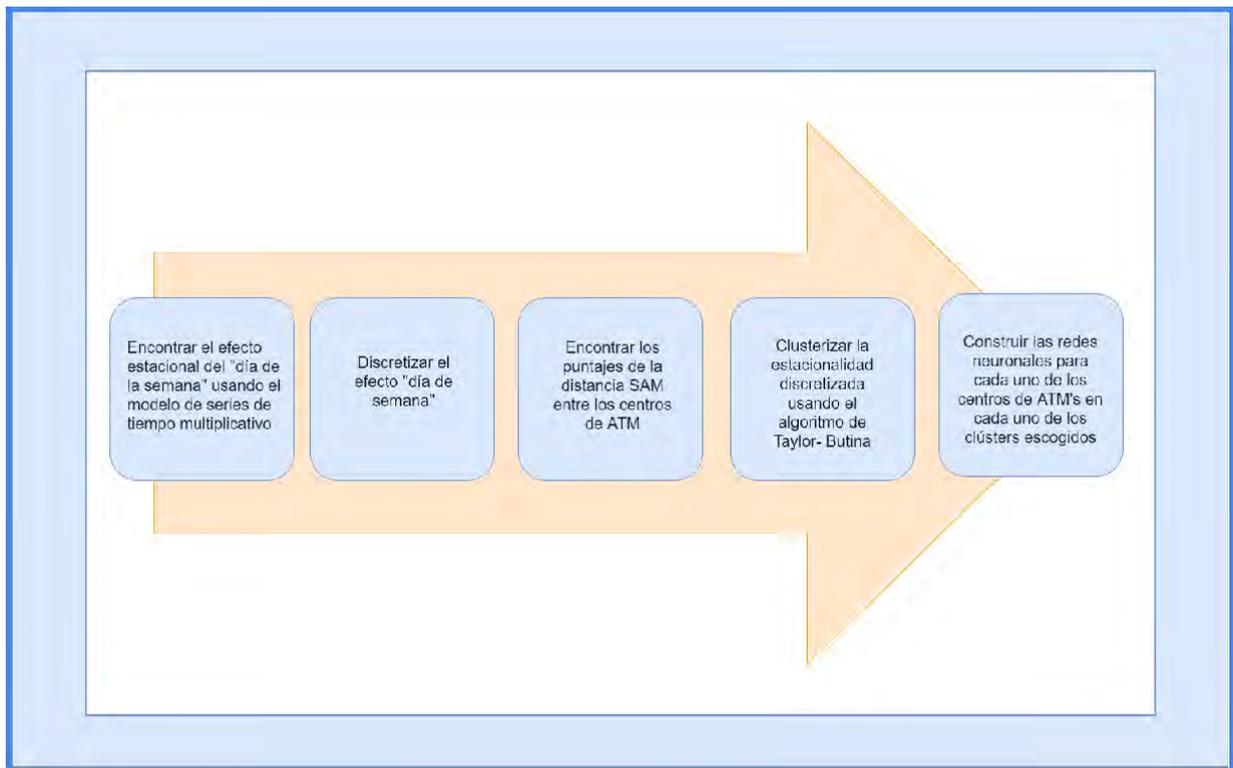


Figura 17. Diagrama de bloques del procedimiento.

Fuente: Venkatesh, K. y Vadlamani R. (2014).

2.3.3.- Aplicación de herramientas de predicción

Este apartado se subdividirá en tres secciones:

2.3.3.1.- Aplicación del modelo de series de tiempo multiplicativo con discretización de la estacionalidad

Antes de aplicar el modelo, se separaron los montos de retiros en los cajeros automáticos según el día de la semana y se aplicó la siguiente ecuación:

$$Y = T * S * C * I \quad (60)$$

Donde:

Y= Valor de la serie de tiempo para respectivo día

T= Representación de la tendencia

S= Representación de la estacionalidad

C= Representación del movimiento cíclico

I= Representación del componente irregular

Se asume que no existe representación del movimiento cíclico ni del componente irregular, por lo que se le asigna el valor de 1. Se pretende hallar la estacionalidad de

los retiros diarios en cada uno de los cajeros automáticos, por lo que, se despejó el factor estacional de la ecuación resultando:

$$S = Y/T \quad (61)$$

Para discretizar la estacionalidad en el periodo de días de semana, el índice de estacionalidad fue dividido en cuatro intervalos correspondientes a los cuartiles, asignándole etiquetas de clase del uno al cuatro, en las cuales el número cuatro representa al valor más alto de estacionalidad. Se muestra a continuación los intervalos utilizados para la discretización.

Tabla 6. Cuartiles usados para la discretización del índice estacional.

Día / Cuartil	1	2	3	4
Lunes	<0.69	0.69-0.77	0.77-0.81	>0.81
Martes	<0.75	0.75-0.81	0.81-0.85	>0.85
Miércoles	<0.95	0.95-1.03	1.03-1.08	>1.08
Jueves	<1.34	1.34-1.46	1.46-1.55	>1.55
Viernes	<1.21	1.21-1.29	1.29-1.38	>1.38
Sábado	<0.65	0.65-0.73	0.73-0.8	>0.8
Domingo	<0.77	0.77-0.83	0.83-0.87	>0.87

Fuente: Venkatesh, K. y Vadlamani R. (2014).

2.3.3.2.- Aplicación de modelo de clustering utilizando el método de secuencia alineada o distancia SAM

Se aplica el método de secuencia alineada o SAM por sus siglas en inglés para calcular la similitud de cada cajero automático en la dimensión discretizada hallada anteriormente. Este método maneja secuencias de longitud variable e incorpora información secuencial en una medida de distancia, similar a la euclidiana, representada por una puntuación o puntaje. Mientras más alto sea el puntaje, significa que existe un mayor esfuerzo por tratar de igualar la secuencia de comparación y por ende no son similares.

Las operaciones elementales en el cálculo de la distancia SAM son las inserciones, las eliminaciones y las sustituciones, que sumados dan el efecto presentado en la ecuación (62):

$$SAM_{dist} = \min \left[\sum_{r=1}^R n * ps \right] \left(\begin{matrix} * c \\ i_e \\ d_e \end{matrix} \right) \left(\begin{matrix} R \\ d_e \end{matrix} \right) \quad (62)$$

Fuente: Venkatesh, K. y Vadlamani R. (2014).

El cálculo de este indicador servirá como input directamente para aplicar el algoritmo de Taylor-Butina para *clustering*, la cual agrupa puntos de datos con listas de vecinos cercanos utilizando el indicador y denomina *singletons* a los puntos sin comportamiento similar a la de algún vecino. El modelo asigna a todos los singletons al clúster más cercano, pero los mantiene con esta etiqueta, decidiendo internamente la cantidad de *clusters*, lo que lo hace un modelo eficaz.

2.3.3.3.- Aplicación de redes neuronales en el pronóstico de la demanda

Como paso posterior a la agrupación de cajeros con características similares, se aplicará un modelo de redes neuronales para predecir la demanda, el cual se compone de cuatro arquitecturas o capas: MLFF, WNN, GRNN Y GMDH.

Estas capas son entrenadas en el periodo de una semana, con la finalidad de tomar en cuenta los consumos de hasta siete días anteriores al analizado. Para comparar el rendimiento de la red, se utilizará el error porcentual absoluto medio simétrico o SMAPE. A continuación, se detalla cada capa utilizada en el pronóstico:

- A. **Wavelet neural network (WNN)**: Es una red neuronal de avance con una capa oculta que se construye a partir de una familia de ondículas a partir de una función $W(x)$ conocida como *ondícula madre* que se encuentra confinada en un intervalo de *ondículas hijas*.
- B. **Método grupal de manejo de datos (GMDH)**: Esta red neuronal es un enfoque autoorganizado basado en la clasificación de modelos y evaluación por criterios predefinidos. Esta red neuronal ofrece una solución jerárquica, probando muchos modelos simples y conservando los mejores para componer una solución optimizada.
- C. **Red neuronal de alimentación directa multicapa (MLFF)**: Esta es una de las estructural de redes neuronales más común, que es simple y efectiva. Está estructurada en redes de nodos organizados en tres capas: capas de entrada, capas ocultas y capas de salida, en la que las capas de entrada y salida sirven como nodos para almacenar las entradas y salidas del modelo, respectivamente. Los pesos en los nodos son aleatorios, y la retroalimentación con un algoritmo estándar de retro propagación, hace que se entrene el algoritmo.

D. **Red neuronal de regresión generalizada (GRNN):** Se considera como una red de función de base radial (RBF) normalizada en la que existe cierta unidad oculta centrada en cada fase de entrenamiento. Estas unidades se denominan núcleos y representan funciones de probabilidad. Esta red neuronal es, de hecho, un modelo de regresión no paramétrica implementada en forma de red neuronal de retroalimentación.

2.3.4.- Solución de la problemática y resultados obtenidos

Se ha dado solución al problema siguiendo con el procedimiento descrito. En el caso específico de las redes neuronales, se utilizó una tasa de aprendizaje de 0.5 a 0.7 y una tasa de impulso de 0.02 a 0.06, por resultar las mejores luego de realizar diversas pruebas y simulaciones. Luego de medir el SMAPE en cada una de las capas aplicadas, se verificó que se tuvo mejores resultados en la capa GRNN con un SMAPE promedio de 18.44%. Además, el algoritmo de Taylor-Butina determinó que cuatro *clusters* era la cantidad óptima para la agrupación de cajeros, por lo que se aplicó las cuatro capas de redes neuronales para cada uno de los *clusters*. Se muestra a continuación el SMAPE obtenido en cada uno de los *clusters* descritos (Ver Tabla 7):

Tabla 7. Resultados de la aplicación de redes neuronales en los clusters.

Cluster	MLFF	WNN	GRNN	GMDH
Cluster 1(45)	19.81	19.63	17.67	19.56
Cluster 2 (21)	21.67	20.79	18.83	20.24
Cluster 3 (22)	22.6	21.93	19.38	21.63
Cluster 4 (23)	20.33	20.07	17.89	19.86
Promedio	21.10	20.60	18.44	20.32

Fuente: Venkatesh, K. y Vadlamani R. (2014).

Además, se logró comprobar que los segmentos siguen efectos estacionales parecidos con picos los jueves, lo cual se muestra en la Figura 18.

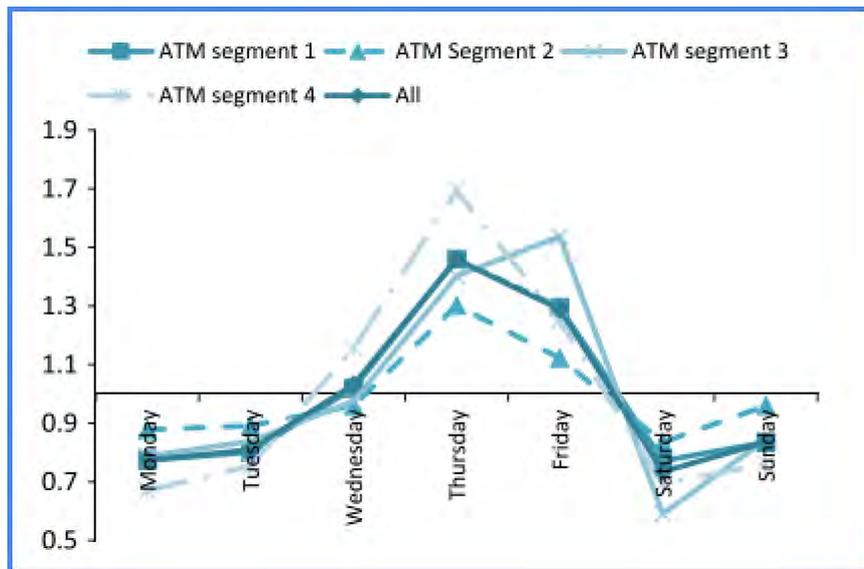


Figura 18. Índices de estacionalidad por cluster.
Fuente: Venkatesh, K. y Vadlamani R. (2014).

2.3.5.- Conclusiones

Se concluye que el SMAPE obtenido en la capa GRNN tuvo un rendimiento menor al de la fuente consultada por el autor, con un valor de 18,44%. Además, el hecho de clusterizar los cajeros automáticos similares reduce en gran medida la tarea computacional en la etapa de pronóstico, mejorando así las predicciones.

Finalmente, se concluye que la agrupación de cajeros tiene una gran implicancia en el planeamiento gerencial de abastecimientos, ya que se puede agrupar abastecimientos paralelos en cajeros automáticos con comportamiento similar, lo que generaría muchos ahorros en transporte al tener rutas previamente establecidas.

2.4.- Caso 4: A new game-theoretical multi-objective evolutionary approach for cash-in-transit vehicle routing problem with time windows (A Real life Case)

2.4.1.- Descripción de la problemática

La problemática se sitúa en Teherán, capital de Irán, en la que se analizó un banco central y treinta y ocho sucursales bancarias. Se determinó que las rutas para abastecer los cajeros automáticos no cumplían con los estándares de seguridad requeridos, puesto que se presentaban robos frecuentes en ciertas zonas de la capital. Además, existían ciertos vehículos que llegaban a los puntos de abastecimiento fuera del horario de atención. Estos hallazgos hicieron notoria la falta de herramientas apropiadas para el diseño de rutas en el proceso de abastecimiento de cajeros automáticos. Se muestra la localización de los puntos en el mapa (Ver Figura 19).

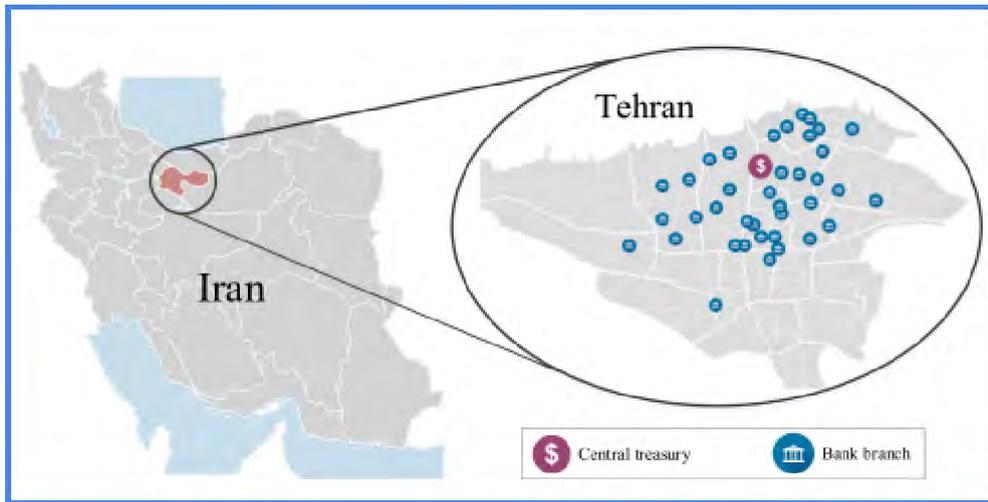


Figura 19. Localización geográfica de Teherán y los bancos localizados en el sector.

Fuente: Ghannadpour, S. F., Zandiyeh, F. (2020)

Una primera solución del problema resultaba en configurar las rutas de distancia más corta; sin embargo, esto implicaba un riesgo muy alto, al no tomar en cuenta las zonas de riesgo alto y no asegurar abastecer los cajeros dentro de los horarios de atención. Esto decanta en la necesidad de elaborar rutas que minimicen costos de transporte y a su vez tengan la mayor seguridad posible, teniendo en cuenta el número de comisarías cercanas y distritos de alta criminalidad además de asegurar que cumplan con los requisitos exigidos por cada una de las sucursales bancarias. Se muestra la ruta con la distancia más corta en un área acotada a sólo ocho sucursales bancarias. (Ver Figura 20).

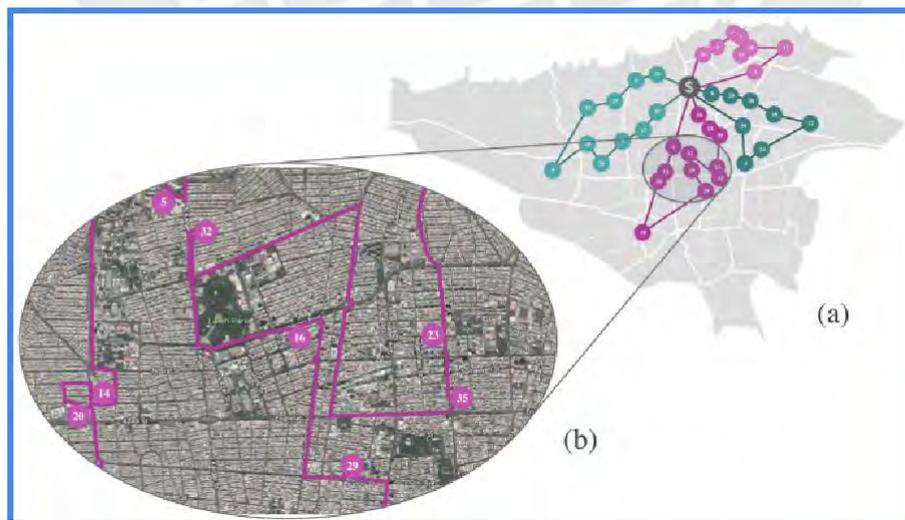


Figura 20. Solución acotada de distancia más corta.

Fuente: Ghannadpour, S. F., Zandiyeh, F. (2020)

2.4.2.- Procedimiento

Se optó por modelar los factores afectados en el riesgo para una posterior predicción y análisis del peligro potencial que significa tener accidentes que involucran cantidades elevadas de efectivo. El enfoque seleccionado fue el de la teoría de juegos, ya que este método examina las estrategias de la competencia y adopta soluciones pertinentes. Además, se minimiza el riesgo del efectivo en tránsito mediante un enfoque de toma de decisiones multicriterio y un algoritmo genético. El método es bautizado como “Game-MOGLNS”, lo que en español significa “Método genético multiobjetivo con teoría de juegos y búsqueda de grandes vecindarios”.

El procedimiento consiste en tres fases principales, las cuales son el relevamiento de información, algoritmo genético y reducción de repetibilidad de enlaces. En la primera fase, se organiza la información disponible en el periodo de análisis, incluyendo distancias, tiempos de viaje, ubicaciones de cámaras de seguridad y consumos en los cajeros. Además, se recolecta datos sobre las rutas establecidas en periodos pasados y las estrategias que se usaban para disminuir el riesgo por robos en las empresas transportadoras de caudales. Con esta información, se estima la probabilidad de robo por zona utilizando teoría de juegos y la probabilidad de éxito del robo, utilizando la ponderación aditiva simple (SAW), desde el punto de vista del ladrón.

En la segunda etapa, se realiza la optimización de las rutas mediante un algoritmo genético híbrido multiobjetivo, el cual incluye nueve sub etapas marcadas en el procedimiento. La primera sub etapa consiste en generar el cromosoma principal usando la heurística de *push forward insertion*. Los cromosomas del algoritmo representan a cadenas de clientes, en la que cada gen del cromosoma es representado por un cliente. Se asignan las rutas desde el centro del cromosoma o punto de partida, de tal manera que no se violen las restricciones y que la ruta creada esté permitida, la cual se muestra en la siguiente figura (Ver figura 21).

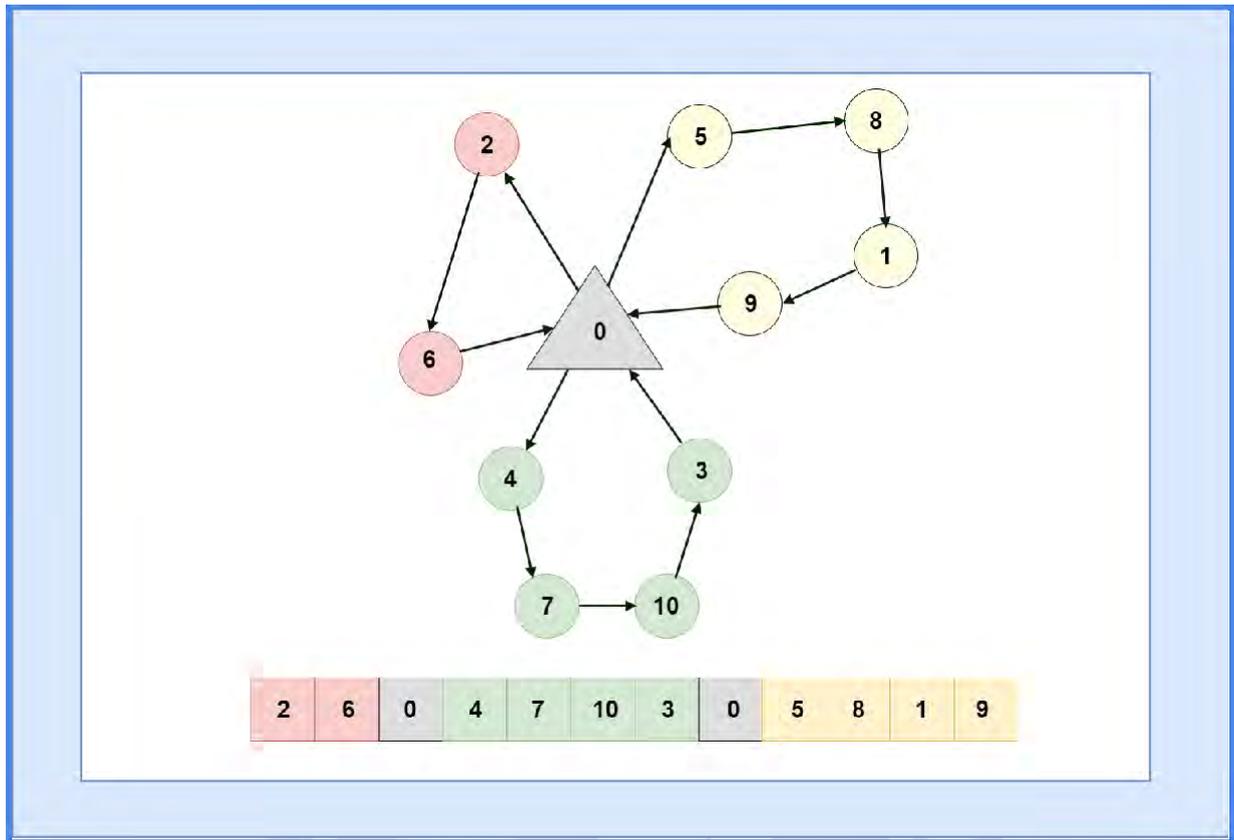


Figura 21. Representación típica de un cromosoma.
 Fuente: Ghannadpour, S. F., Zandiyeh, F. (2020)

Las siguientes subetapas consisten en optimizar las rutas iniciales mediante métodos de selección, combinación transversal de descendientes y mutación. Posteriormente a estas subetapas, se generan nuevas poblaciones mediante *montañismo*, las cuales son mejoradas mediante la *escalada* y *elitismo*, heurísticas propias del algoritmo genético.

Finalmente, en la tercera etapa, se busca reducir la cantidad de rutas repetidas por las que pasa un vehículo, puesto que el autor lo considera riesgoso, dado que los ladrones pueden identificarla y la probabilidad de un ataque en el próximo viaje aumentaría. Se muestra, en la Figura 22, el diagrama de flujo del procedimiento explicado.

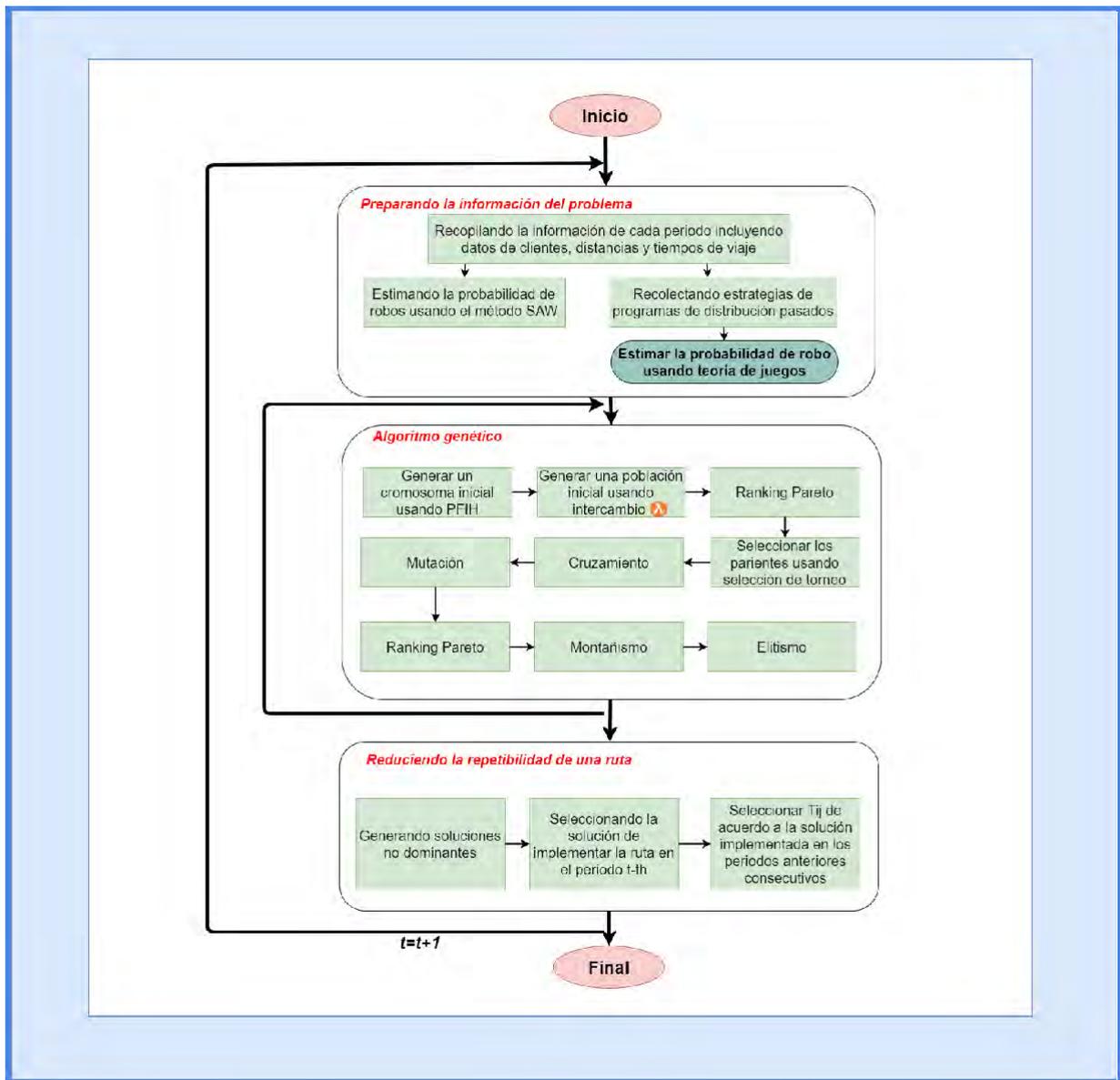


Figura 22. Procedimiento para configurar la ruta óptima.
Fuente: Ghannadpour, S. F., Zandiyeh, F. (2020)

2.4.3.- Aplicación de herramientas

2.4.3.1.- Teoría de juegos

Se utiliza esta herramienta para realizar un análisis de riesgos. Se tomarán dos roles en la aplicación de la herramienta, *el defensor*, el cual alude a las empresas transportadoras de caudales y *el atacante*, refiriéndose al ladrón que intenta apropiarse del efectivo transportado. Se asume que, en primer lugar, el defensor tomará una estrategia defensiva para proteger su recurso y el atacante, conociendo esa posición, decide un lugar estratégico para atacar. Se modelan las probabilidades de éxito en cada estrategia y se dimensiona el riesgo de tomar cada una de las rutas posibles. Se repite el juego tantas veces hasta encontrar la ruta que minimice los

riesgos, utilizando los árboles de decisión. Se presenta, en la Figura 23, la representación gráfica de la estimación de estas probabilidades.

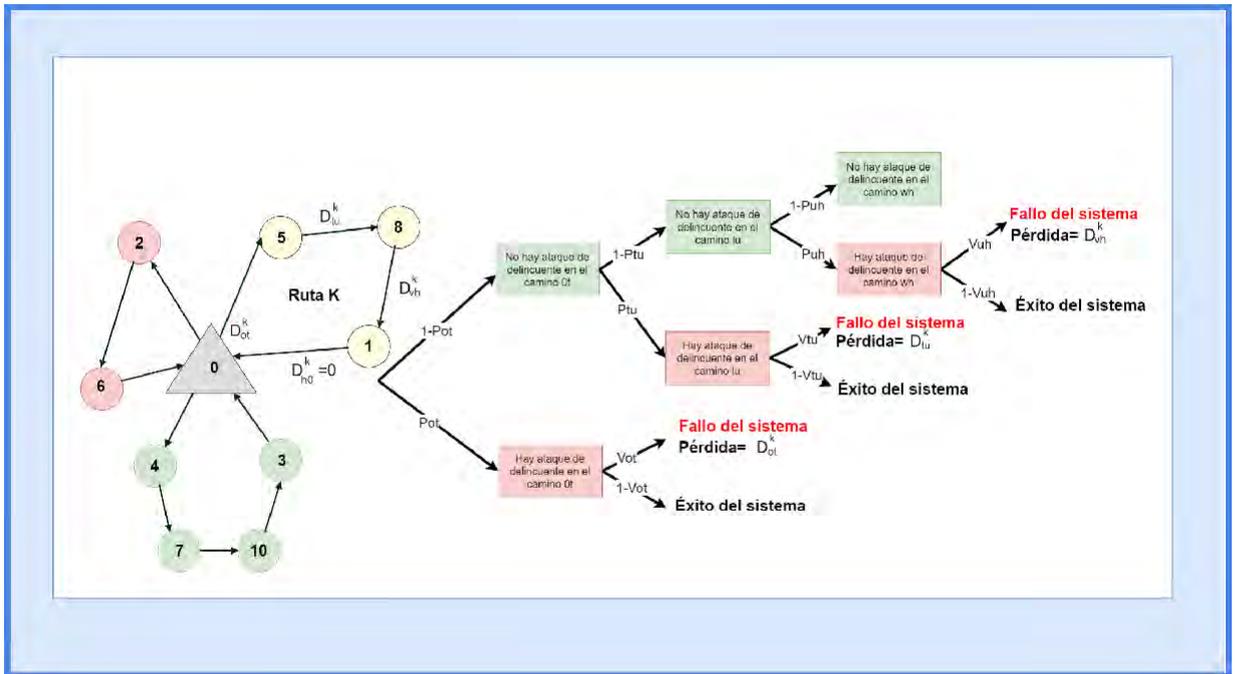


Figura 23. Procedimiento para calcular el riesgo por robos.
Fuente: Ghannadpour, S. F., Zandiyeh, F. (2020)

2.4.3.2.- Vehicle Routing Problem with Time Window (VRPTW)

Se aplica el modelo de ruteo de vehículos para optimizar las rutas escogidas considerando los horarios de atención, además de incluir como uno de los objetivos, la minimización del riesgo. El centro de distribución sería el banco de reserva o bóveda, denominado nodo central, del que partirán todos los vehículos hacia sus centros de reparto. Al final del servicio, los vehículos transportadores de caudal volverán a este centro para iniciar con viajes posteriores o terminar el servicio. Cabe resaltar que cada vehículo cuenta con restricciones propias, como lo son la capacidad de transporte, los viajes máximos diarios y el horario laboral del personal que los maneja. Cabe resaltar que este algoritmo sólo es aplicado en una etapa preliminar, ya que, por la complejidad computacional, se decide aplicar un algoritmo genético.

2.4.3.3.- Algoritmo genético

Para la aplicación inicial del modelo, se utiliza la heurística de impulso empujando hacia adelante (PFIH), que consiste en formar rutas con costo diferencial mínimo y agruparlas con puntos con similares costos sin violar las restricciones del vehículo,

cumpliendo los horarios. La ruta se cerrará en cuanto se incumplan alguna de las restricciones.

Posteriormente, se optimizan las rutas en la fase de selección, en las que se aplican algoritmos especializados que ordenan a la población (nodos de demanda) de manera aleatoria. Utilizando el análisis de Pareto, se escogen las rutas aleatorias con menores costos agrupadas por el algoritmo y son seleccionadas para pasar a la siguiente etapa, estas agrupaciones son llamadas cromosomas. Cada uno de estos cromosomas *padres* son agrupados formando descendencias mediante la aplicación del algoritmo de *mejor costo - mejor ruta* o BCBRC por sus siglas en inglés. Se escoge la mejor descendencia utilizando el mismo criterio que en la etapa pasada, el cual es el menos costoso. Cabe resaltar que el concepto de costo en este algoritmo abarca los costos de transporte y de oportunidad de incurrir en pérdidas por robos o abastecimientos fuera de horario.

Los descendientes que pasaron a esta fase son replicados, por lo que esta fase es nombrada como *Mutación*. Estas réplicas buscan evitar que los óptimos elegidos solo sean óptimos en localidades cercanas. Además, se verifica que las rutas no contradigan los resultados obtenidos en la teoría de juegos. Para estas fases se utilizan tres tipos de algoritmos los cuales son descritos a continuación:

A. Mutación IDR (Intercambio de clientes dentro de rutas diferentes):

Se elige uno de los puntos de abasto o clientes al azar y se busca el punto más cercano, si este pertenece a otra ruta se utiliza uno de los tres algoritmos siguientes.

- Fusión de un solo punto: uno de ambos puntos es expulsado de su ruta y es incluido en la ruta del otro punto en esta mutación.
- Intercambio de rutas: se intercambian las rutas entre ambos puntos
- Fusión de segmentos: en este algoritmo, uno de los puntos y sus adyacentes son agregados a la ruta del otro punto seleccionado.

B. Contrarrestar:

Se seleccionan dos puntos al azar de una ruta y se invierten.

C. Intercambio de rutas:

Se eligen dos puntos al azar y las rutas se intercambian desde los puntos seleccionados.

Se presenta, a continuación, un ejemplo del algoritmo de fusión de un solo punto, parte del modelo de mutación IDR (Ver Figura 24):

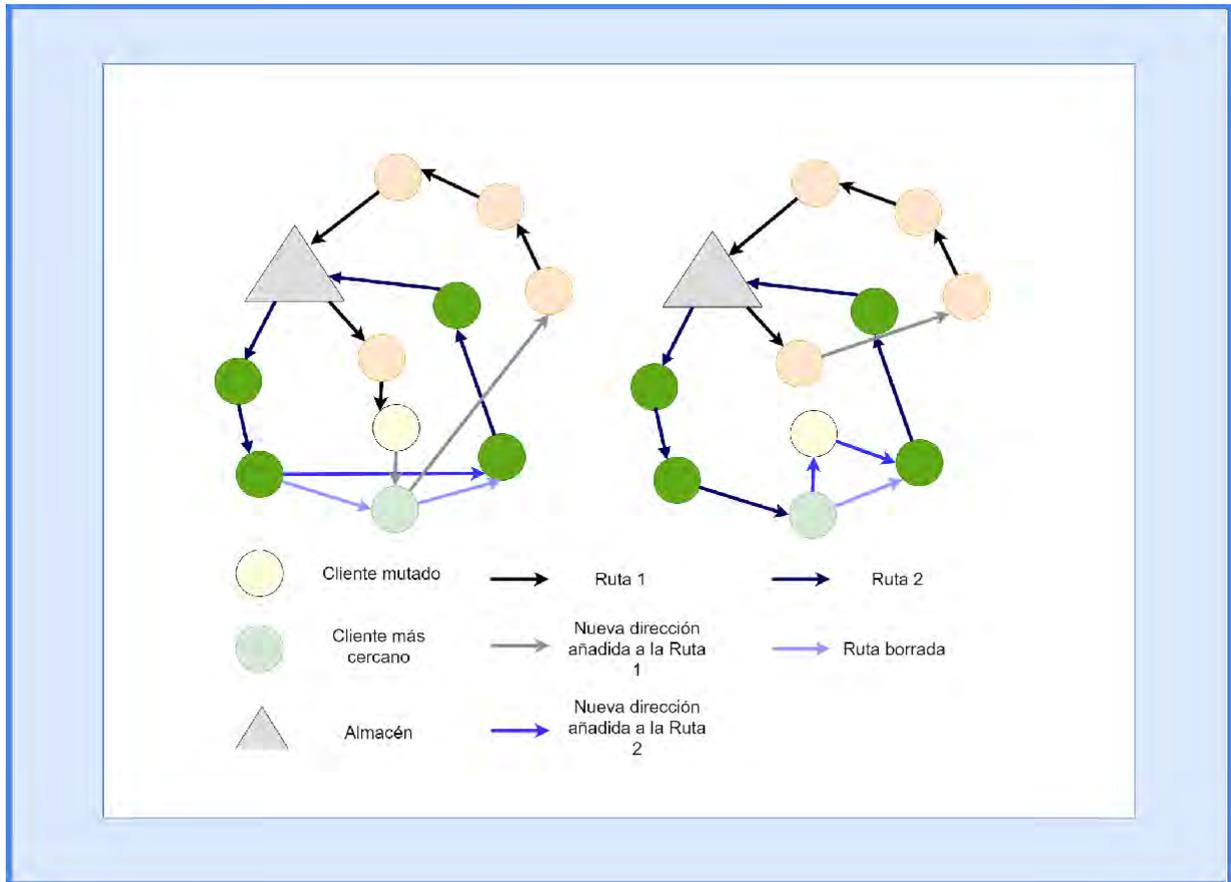


Figura 24. Algoritmo de fusión de un solo punto.
Fuente: Ghannadpour, S. F., Zandiyeh, F. (2020).

Luego de haber generado una nueva población, como resultado de los algoritmos anteriores, se procede a mejorar la calidad de la población utilizando los mecanismos de *escalada* y *elitismo*. El procedimiento de mejora consiste en la aplicación de siete operadores de búsqueda locales (VND) y otro algoritmo heurístico.

Finalmente, se reduce la repetibilidad de las rutas seleccionadas. El autor considera que repetir rutas es un indicador de riesgos y podría aumentar la probabilidad de robos en la ruta, siguiendo el resultado obtenido por la implementación de la teoría de juegos. Se reinicia el algoritmo cada vez que detecta rutas repetidas hasta generar una solución que cumpla con esta restricción.

2.4.4.- Solución de la problemática y resultados obtenidos

Se validó el algoritmo en dos fases. En la fase inicial, se probó el modelo en una muestra pequeña y se midió el efecto de los factores en cada fase, verificando la reducción óptima del riesgo. En la segunda fase, se aplica en todo el alcance. Se obtuvo que la distancia recorrida total fue de 319 km y el riesgo promedio alcanzado fue de 0%. Esta reducción en el riesgo significó un aumento de 5.36% en la distancia total recorrida.

Con respecto a la fase inicial de aplicación del algoritmo VRPTW, se muestra el resultado, en el que las rutas son diferenciadas con colores diferentes según el riesgo. Además, cabe resaltar que se ha considerado factores externos como la cantidad de símbolos de seguridad, presencia policial, rutas de escape y lugares de alta peligrosidad para estimar el riesgo (Ver Figura 25).

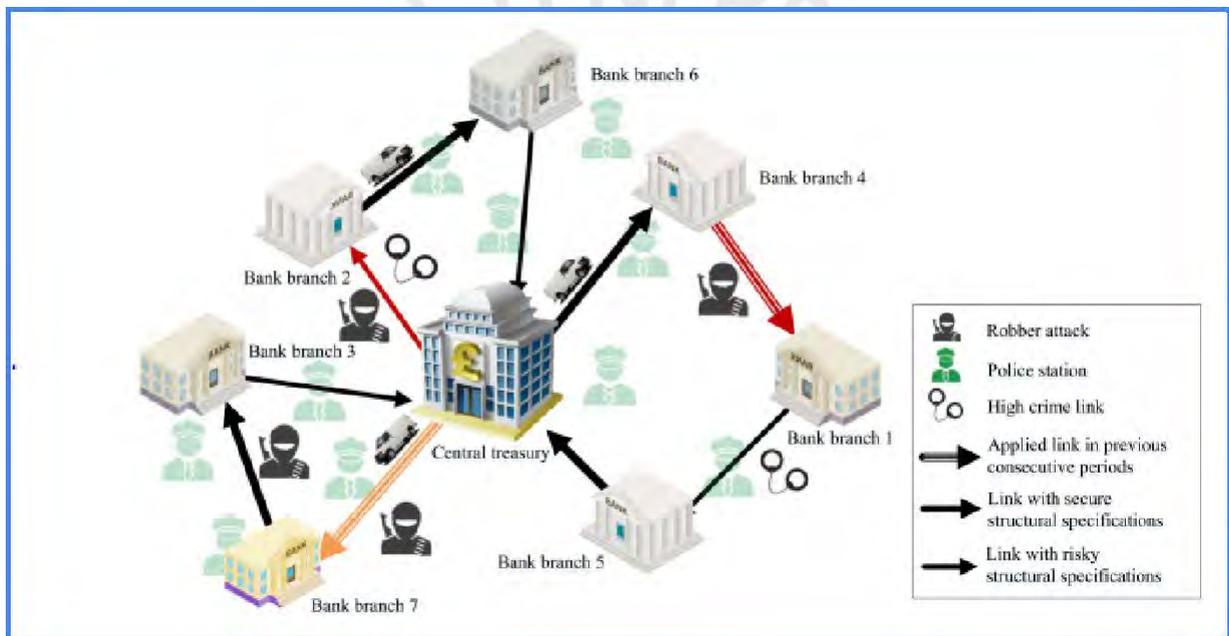


Figura 25. Resultado de aplicación de algoritmo VRPTW

Fuente: Ghannadpour, S. F., Zandiyeh, F. (2020)

Para medir el resultado del algoritmo genético híbrido aplicado, se hizo uso de la prueba de rangos con signo de Wilcoxon, en la que se obtiene optimalidad. Asimismo, se mide el desempeño del algoritmo genético utilizando el método RSM, el cual integra el uso de un software gráfico.

2.4.5.- Conclusiones

Se concluye que la adición de factores como el riesgo por robos complican en gran magnitud la solución de los modelos. Además, se evidenció que la aplicación de la teoría de juegos

representa una herramienta poco explorada pero muy útil, ya que, el autor afirmó ser el primero en utilizarla en aplicaciones de ruteo de vehículos.

Asimismo, se observó que la inclusión del factor de riesgo sólo aumentó la distancia recorrida en aproximadamente 5%, lo que evidencia que vale la pena añadir factores relevantes al análisis, puesto que no impacta en gran magnitud a los costos, desde la perspectiva de la implementación de un algoritmo genético

2.5.- Caso 5: Improvement of Demand Forecasting Models with Special Days

2.5.1.- Descripción de la problemática

En la ciudad de Londres, los primeros cajeros automáticos fueron instalados hace más de 50 años. A raíz del uso constante de estas máquinas por parte de los ciudadanos, los bancos en Londres se han visto en la necesidad de predecir la demanda de efectivo tanto en cajeros, como en otras sucursales, con la finalidad de optimizar los requerimientos de efectivo en los cajeros automáticos, evitando exceso de caudal inactivo y que los cajeros caigan en un estado de *cash-out*, lo que impactaría negativamente en la experiencia al cliente.

Si bien se han expresado muchas soluciones en la literatura mundial de este tipo de problemática, las nuevas soluciones se centran únicamente en métodos de inteligencia computacional avanzada, sin embargo, estos exigen una amplia cantidad de recursos de almacenamiento y procesamiento de muestras de entrenamiento. Por ello, se evidencia la necesidad de contar con nuevas alternativas de pronósticos de demanda que permitan obtener niveles de precisión similares, utilizando menos recursos.

Por otro lado, la demanda de efectivo tiene ciertas particularidades, específicamente debido a que se ve afectada por los denominados “días especiales”, como feriados o días festivos, por lo que se vuelve aún más necesario integrar estas particularidades en un pronóstico, para conocer la demanda real y, como consecuencia, tomar correctas decisiones de planificación con respecto al abastecimiento de cajeros automáticos (ATM).

2.5.2.- Procedimiento

Para proponer una solución a la problemática planteada, se optó por realizar, en primer lugar, un preprocesamiento de los datos relacionados a la demanda de efectivo en 11 cajeros automáticos de Londres. Se realizó un análisis de datos, donde, en este caso, se halló la media, desviación estándar y el número de valores faltantes. Luego, se realizó una conversión del *dataset* en un formato de serie de tiempo, donde se colocó el valor de “0” a

los registros faltantes. Como siguiente paso, se realizó el manejo de data faltante, donde se reemplazó al registro equivalente a "0" por los valores medios de cada cajero automático.

Asimismo, se añadieron nuevas columnas que expresan el planteamiento de los días especiales notables en Reino Unido, que en este caso fueron 19, tales como el Día de San Valentín y el Día de San Patricio. Cada día especial era representado con el valor de "1" en la columna correspondiente, mientras que los que no, tomaban el valor de "0" en el conjunto de datos. Además, se consideraron a los días sábado y domingo como días especiales, así como a los días de una semana previa a un día festivo, debido a que muchas personas podrían preferir realizar las compras de manera anticipada. Era imprescindible la inclusión de esta información debido a que es un factor relevante para obtener resultados más acertados de previsión de la demanda de efectivo en cajeros.

El conjunto de datos acumulado se dividió no sólo por día, sino que también se especificó el mes, año, día de semana y trimestre. Ello se debe a que se podían evidenciar ciertas estacionalidades tanto a lo largo del año como dentro de una semana, donde los fines de semana y los meses de verano eran propensos a tener un comportamiento superior de demanda, con respecto a los días de semana y los meses de invierno, respectivamente.

Una vez preprocesada la data, se procedió a ejecutar ciertos métodos o técnicas de pronóstico, con la finalidad de conocer la que mejor funcionaba para estimar las necesidades de efectivo en cajeros automáticos. Para realizar la comparación adecuada del desempeño de cada método, se utilizó el indicador de error denominado *Symmetric mean absolute percentage error* (SMAPE).

2.5.3.- Aplicación de herramientas de pronóstico

Si bien en la literatura previa investigada por Catal, Fenerci, Ozdemir y Gulmez (2015), se llegó a la conclusión de que el uso de la red neuronal de regresión general (GRNN) proporcionaba el mejor rendimiento con un ratio SMAPE equivalente a 18.44% (Venkatesh, Ravi, Prinzie y Van Den Poel, 2014), en este caso, se tomó la decisión de aplicar algoritmos convencionales para la solución de la problemática, dado que estos permitían obtener resultados en menores tiempos de cálculo, utilizando menores recursos para almacenamiento y procesamiento de datos, además de facilitar la explicación de funcionamiento del modelo a los clientes.

Con la finalidad de potenciar el modelo, se utilizaron los días especiales de Reino Unido y, finalmente, se trabajaron modelos como redes neuronales artificiales, suavizado exponencial

y el modelo autorregresivo integrado de media móvil (ARIMA), utilizando la plataforma de software SAS. Por un lado, se aplicó la herramienta SAS Forecast Server para los enfoques de ARIMA y suavizado exponencial, donde el último utilizó las fórmulas del apartado 1.3.1.B.2.c del presente documento. Por otro lado, se utilizó la herramienta SAS Enterprise Miner para aplicar el enfoque de redes neuronales, además de realizar el preprocesamiento del conjunto de datos por medio de la misma herramienta.

Es necesario precisar que también se consideró la utilización de un algoritmo de *clustering* jerárquico antes de utilizar directamente la red neuronal, donde se analizó el modelo lineal generalizado (GLM), así como la red neuronal artificial perceptrón multicapa (MLP) y la red de función de base radial (RBF). Con respecto al agrupamiento, se obtuvo cuatro clases al final del análisis.

2.3.4.- Solución de la problemática y resultados obtenidos

Para realizar la comparación de los resultados, se utilizó el ratio de error de pronóstico denominado SMAPE, el que permite calcular el nivel de desempeño de un modelo en cuestión de precisión a partir de la siguiente fórmula:

$$SMAPE = \frac{1}{n} \sum_{t=1}^n \frac{|X_t - F_t|}{(X_t + F_t)/2} * 100 \quad (63)$$

Donde:

X_t= Valor real

F_t= Valor de pronóstico

n= Tamaño del periodo de prueba

Utilizando dicho indicador, se obtuvo el mejor rendimiento utilizando el suavizado exponencial, incluyendo el conjunto de columnas de días especiales. Para los datos de entrenamiento, se obtuvo un valor equivalente al 19,21%, mientras que para un horizonte de predicción de 56 días, se obtuvo un valor del 21,57%. Con el uso de redes neuronales, se obtuvo el mejor ratio con el método GLM, siendo este de 40,64%. Al aplicar *clustering* antes de la red neuronal, se obtuvo un mejor rendimiento (39.73%), sin embargo, no fue suficiente como para alcanzar el desempeño del método de suavizado exponencial. A continuación, se presentan los resultados de aplicación del método GLM para el clúster número 1 en la Figura 26, donde se contrastan los resultados de media y pronóstico obtenidos tanto con los datos de entrenamiento como con los datos de prueba.

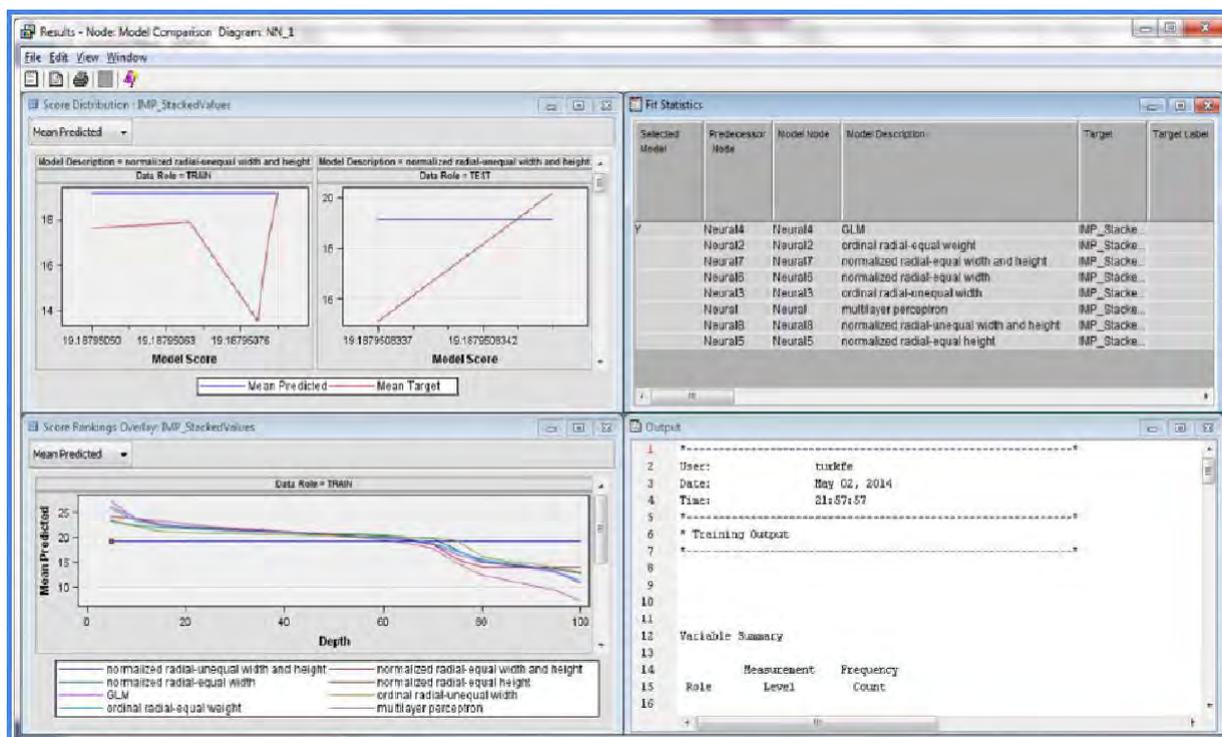


Figura 26. Resultados de aplicación del método GLM para el clúster 1.
Fuente: Catal, C., Fenerci, A., Ozdemir, B. y Gulmez, O. (2015).

2.5.5.- Conclusiones

Se concluye que métodos tradicionales de *forecasting* pueden resultar igual de efectivos que métodos computacionales avanzados, con respecto al pronóstico de la demanda de efectivo de cajeros automáticos (ATM). En este caso, se concluye que el método de suavizado exponencial resulta ser el de mejor desempeño para conocer los retiros de efectivo a futuro. Asimismo, se concluye que es de vital importancia la incorporación de los días especiales del país de análisis en el conjunto de datos que se posee, debido a que impacta favorablemente en el nivel de precisión de cualquier pronóstico utilizado.

CAPÍTULO 3. DIAGNÓSTICO DEL SECTOR BANCARIO

En este capítulo, se abordará el diagnóstico del sector bancario, a partir de tres acápites: Análisis de la situación actual del sector bancario, análisis de indicadores financieros de la Banca Comercial y análisis de *stakeholders* en la Banca Comercial. Este análisis es de vital importancia para conocer el estado actual, comportamiento, cambios y grupos de interés de estas instituciones dentro de la cadena de suministro del efectivo.

3.1.- Análisis de la situación actual del sector bancario

En la actualidad, el sector bancario peruano se encuentra integrado por un conjunto de instituciones, tales como el Banco Central de la Reserva (BCR), el cual se encarga de regular las operaciones del sector financiero nacional y gestionar las reservas del país; El Banco de la Nación, que es el responsable de las operaciones bancarias del sector público y la Banca Comercial y de Ahorros, que está integrada por entidades financieras que se comportan como intermediarios que receptionan dinero del público y a su vez conceden créditos en diversas modalidades (Ayala, s/f).

La situación económica actual y la pandemia producida por el coronavirus no ha sido ajena al sector financiero. Según estimaciones publicadas en septiembre por la empresa consultora Apoyo Consultoría, se proyecta que el PBI peruano tenga su mayor caída en los últimos 100 años, disminuyendo en 13%; sin embargo, las empresas del sector bancario no se verían afectadas en tal magnitud. Un ejemplo de aquello se refleja en las estimaciones realizadas por Scotiabank, en las que se proyectan crecimientos de 0.3% al finalizar el año, aunque su proyección disminuyó en 2% debido al COVID-19.

Por otro lado, la pandemia, aunque indirectamente, ha tenido efecto sobre las tasas de interés cobradas a ciertos sectores de la población. Aunque es de conocimiento que las tasas de interés son determinadas por el libre mercado, la coyuntura actual ha obligado al estado, por medio de la Superintendencia de Banca y Seguros y los programas de Fortalecimiento Patrimonial de Empresas (FOPE) y el fondo Reactiva Perú, a realizar esfuerzos para bajar los costos y ofrecer mejores condiciones a todos los tipos de clientes, desde grandes empresas hasta personas naturales. Este se ve evidenciado en el reporte del mes de octubre de 2020 publicado por el Banco Central de Reserva del Perú, en el que se evidencia que la tasa de interés activa de mercado (TAMN) solo aumentó 0.1% con respecto a Septiembre y se mantiene estable en el horizonte de 2020 con valores que se acercan al 12% efectivo anual.

La Banca Comercial en el Perú se compone principalmente de dieciséis entidades financieras, de las cuales se destacan, el Banco de Crédito, Banco BBVA, Interbank y Scotiabank. Cabe señalar que el 27 de abril de 2020 se sumó el Banco de China, siendo autorizada por la SBS y catalogada como una empresa bancaria de operaciones múltiples (RPP, 2020). El desempeño actual de cada una de estas entidades puede ser justificado desde distintos puntos de vista, entre los que destacan las colocaciones, las captaciones, el apalancamiento y entre otros.

3.2.- Análisis de indicadores financieros de la Banca Comercial

Para obtener un diagnóstico adecuado sobre la gestión bancaria y los cambios suscitados este año, se deben analizar los siguientes indicadores financieros tanto del sector como de las empresas que lo componen: ratios de solvencia, de rentabilidad y de liquidez. Estos son, junto con el nivel de morosidad, los principales centros de control en la gestión por objetivos (Lizarzaburu y Del Brio, 2016). A continuación, se realizará un análisis del estado del sector en base a sus indicadores financieros, realizando una comparativa entre los valores registrados en Septiembre del 2019 y Septiembre del 2020, con la finalidad de conocer los cambios del sector con respecto al periodo anterior.

Con respecto a la solvencia, un indicador representativo para conocer la capacidad de pago de los bancos es el ratio de capital global, que hace referencia al patrimonio como una parte del total de activos y contingentes (SBS, 2010). La Superintendencia (2011) establece que el requerimiento exigido debe ser del 10%. Tal como se observa en la Figura 27, el mínimo ratio del sector en el 2020 le corresponde al banco Pichincha, aunque supere el límite inferior mínimo estipulado por la SBS, con un 13.26%. En general, casi el 74% de las empresas del sector bancario mejoraron su ratio de solvencia para el año 2020, generando un incremento del 2% en el ratio promedio del sector, que pasó de 17.16% al 17.52%. En este comparativo no se considera a la nueva entrada al sector bancario, puesto que el Banco de China es considerado como una de las instituciones financieras más influyentes del mundo, según *The Banker*, por lo que su ratio supera excesivamente los indicadores comunes del Perú, con un 514 %.

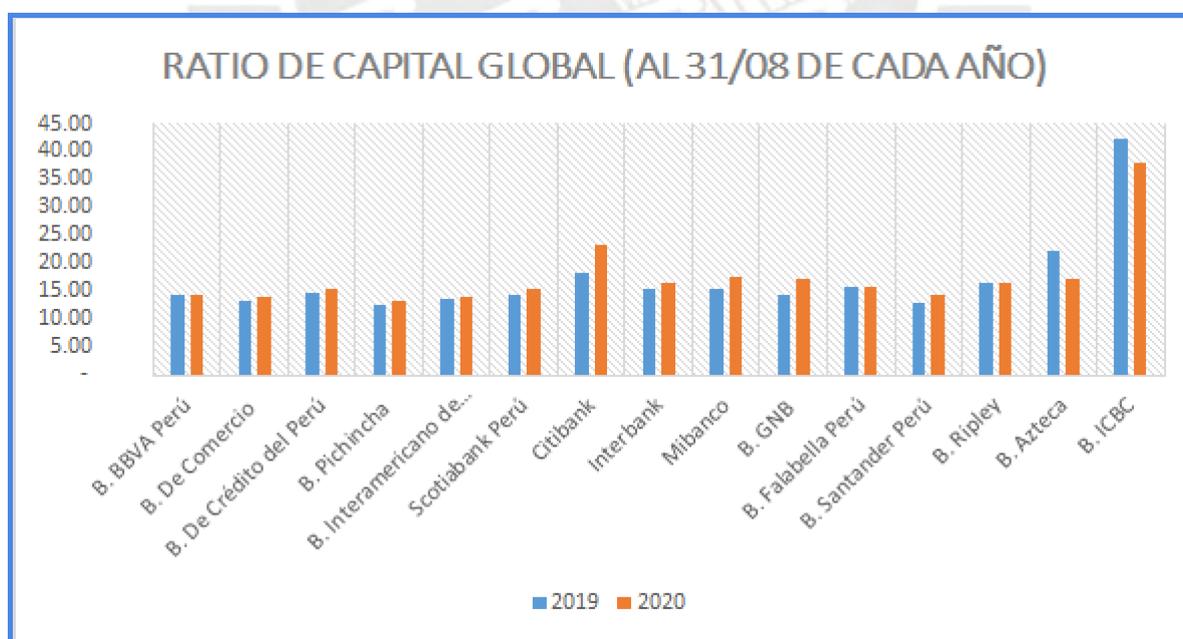


Figura 27. Indicador de solvencia: Ratio de capital global 2019-2020.

Fuente: Superintendencia de Banca y Seguros (2019-2020).

Por otro lado, con respecto a los indicadores de rentabilidad, se visualizó una caída de la rentabilidad económica en cerca del 94% de las entidades bancarias en el Perú. De acuerdo con la Superintendencia de Banca y Seguros (2010), este ratio indica que tan bien la empresa ha hecho uso de sus activos para generar ganancias. Tal como se visualiza en la Figura 28, muchas entidades bancarias se han visto totalmente afectadas debido al COVID-19, lo que se plasma en sus disminuciones en niveles de rentabilidad, incluso algunas, como el Banco Falabella Perú, el Banco Ripley y el Banco Azteca, poseen rentabilidades económicas negativas al mes de Septiembre de 2020. En resumen, la rentabilidad sobre activos del sector ha tenido una disminución del 77% con respecto al año previo, impactando en un cambio del 1.82% al 0.42% del ratio promedio del sector.

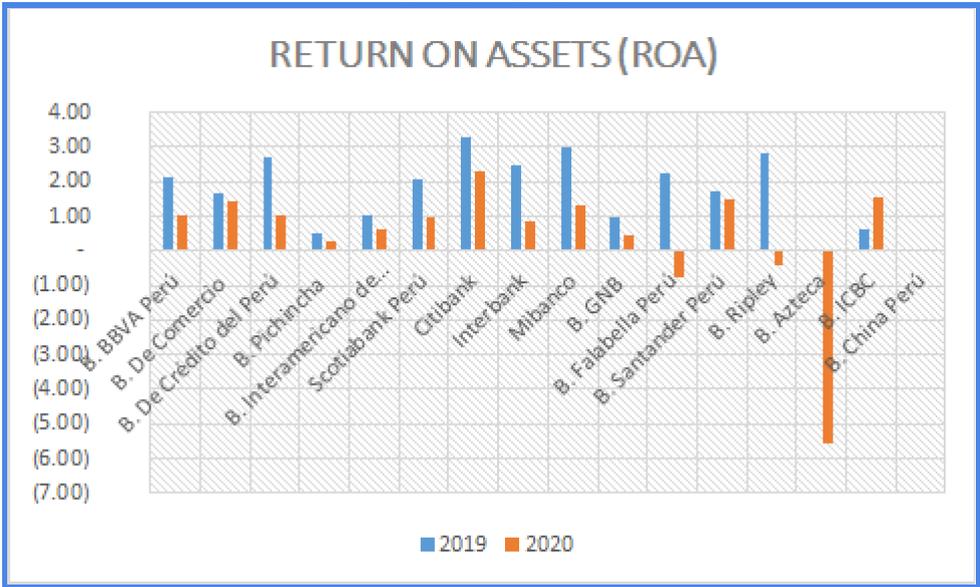


Figura 28. Indicador de rentabilidad: Rentabilidad económica de las entidades bancarias en el Perú 2019-2020.

Fuente: Superintendencia de Banca y Seguros (2019-2020).

Asimismo, otro indicador útil para medir el desempeño del sector bancario es la rentabilidad de capital, el que refleja los beneficios obtenidos de los accionistas a raíz de su patrimonio y que es mayormente utilizado para tomar decisiones de inversión a futuro (Superintendencia de Banca y Seguros, 2010). En el caso del sector bancario, aproximadamente el 94% del total de las instituciones han presenciado una disminución de su rentabilidad de capital con respecto al año 2019, observándose un comportamiento similar al de la rentabilidad económica (Ver Figura 29). Sin embargo, la mayor cantidad de estas aún mantiene resultados positivos, por lo que los beneficios sobre capital promedio del sector hasta Septiembre del 2020 equivalen al 4.42%.

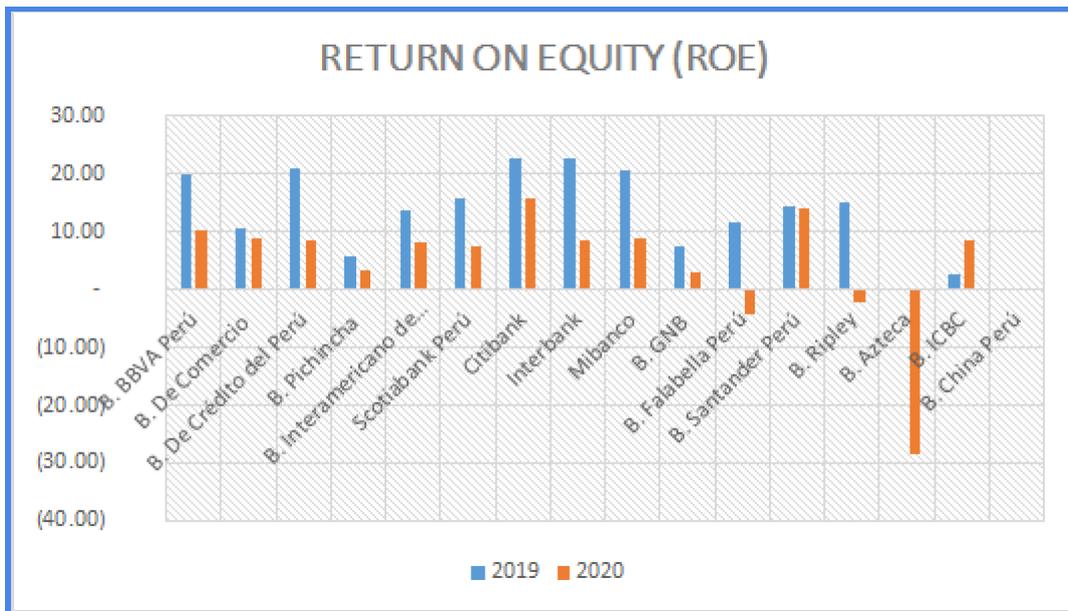


Figura 29. Indicador de rentabilidad: Rentabilidad de capital de las entidades bancarias en el Perú 2019-2020.

Fuente: Superintendencia de Banca y Seguros (2019-2020).

Con respecto a la liquidez, el indicador de liquidez en moneda nacional (MN) permite conocer el promedio mensual de los saldos diarios de activos líquidos con respecto al promedio mensual de saldos de pasivos de corto plazo (Superintendencia de Banca y Seguros, 2010). De acuerdo con la SBS, las instituciones financieras deben mantener un indicador mínimo del 8% en moneda nacional, por lo que se valida con la Figura 30, que incluso el indicador mínimo visualizado entre los años 2019 y 2020 es del 18%, por parte del Banco Pichincha durante el año 2019. En general, de acuerdo a la comparativa, el ratio promedio de liquidez del sector ha incrementado en un 45%, llegando hasta un 53% para Septiembre del 2020, por lo que se prevee que continúe incrementando la liquidez del sector, lo que permite impulsar el crecimiento de las colocaciones.

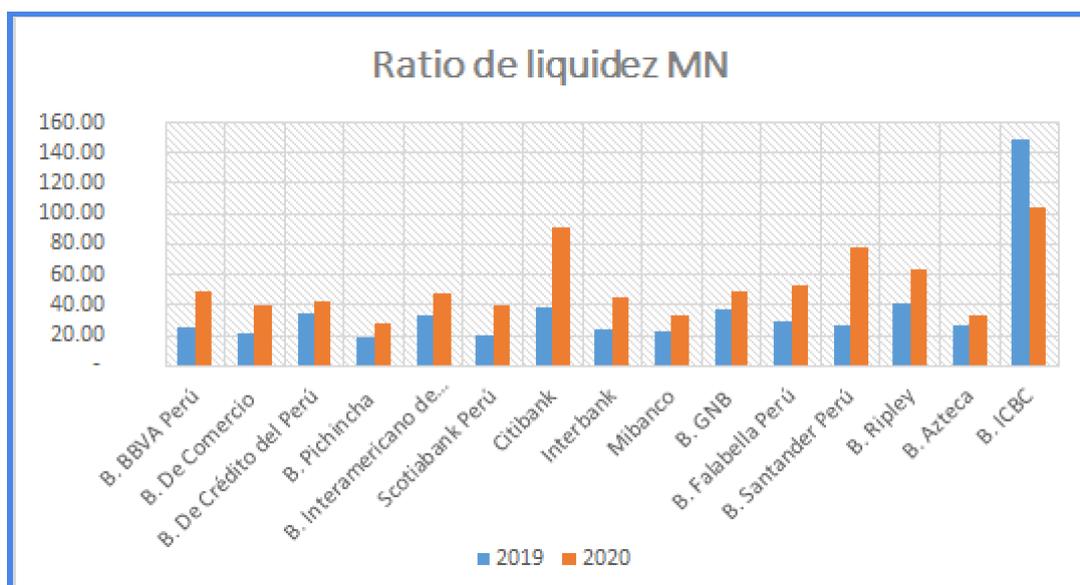


Figura 30. Indicador de liquidez de las entidades bancarias en el Perú 2019-2020.
Fuente: Superintendencia de Banca y Seguros (2019-2020).

Por otro lado, con la finalidad de evidenciar el nivel de morosidad del sector, se utiliza un indicador de calidad de activos que expresa los créditos atrasados con respecto a los créditos directos, donde los atrasados hacen referencia a los créditos que se encuentran vencidos o en cobranza judicial (Superintendencia de Banca y Seguros, 2010). Con la Figura 31 se demuestra efectivamente que los clientes del sector bancario han disminuido su capacidad de pago de créditos con respecto al año 2019, por lo que el sector tuvo que afrontar un incremento del 24% en este indicador, debido a un paso del 3.56% al 4.4% entre los periodos de Septiembre de 2019 y Septiembre de 2020. En particular, el banco Azteca presenta altos porcentajes de cartera atrasada, con respecto a las instituciones del sector, debiéndose en mayor medida a las altas tasas de interés a las que Azteca otorga créditos, así como a la mayor dificultad de pago de deudas por parte de los clientes, debido a la crisis económica a causa del COVID-19.

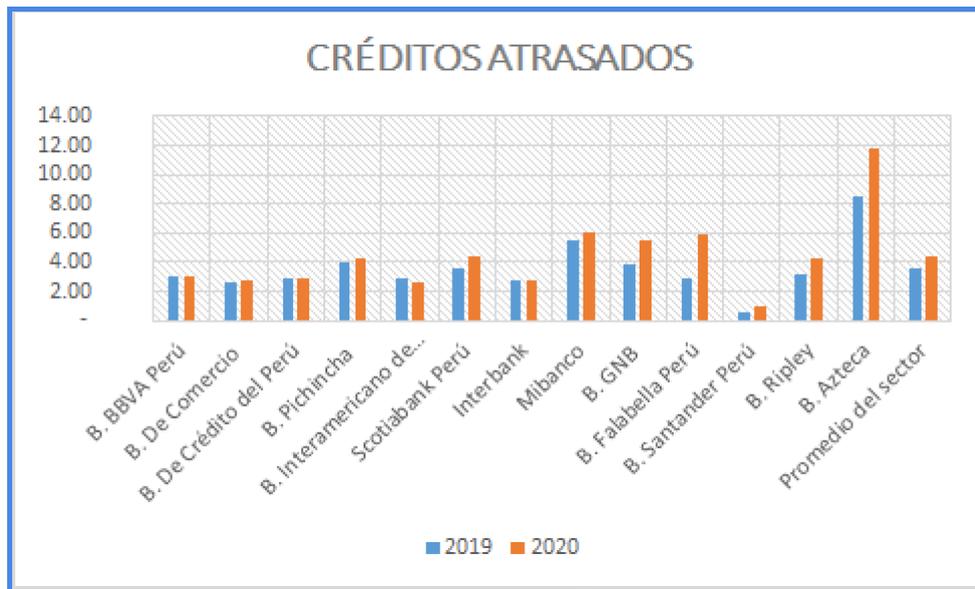


Figura 31. Indicador de créditos atrasados de las entidades bancarias en el Perú 2019-2020.
Fuente: Superintendencia de Banca y Seguros (2019-2020).

Por último, para conocer cómo ha evolucionado el índice de morosidad promedio durante el año 2020, se puede visualizar la Figura 32. En ella, se observa que el mes de Abril fue un mes crítico para el sector bancario debido a que se obtuvo el mayor índice de morosidad registrado este año, el cual fue equivalente al 3.38%. Ello se debe a que se trata del mes siguiente al confinamiento sanitario, que imposibilitó que muchas empresas continúen operando o que muchas personas trabajen.



Figura 32. Índice de morosidad promedio durante el año 2020.
Fuente: Asociación de Bancos del Perú (2020).

Realizando un mayor detalle de ello, se observó que además de obtener los índices más altos durante el mes de abril, el segmento de pequeñas y medianas empresas (PyMES) ha sido el de mayores porcentajes de morosidad, alcanzando valores de un poco más del 9% durante el presente año (Ver Figura 33).

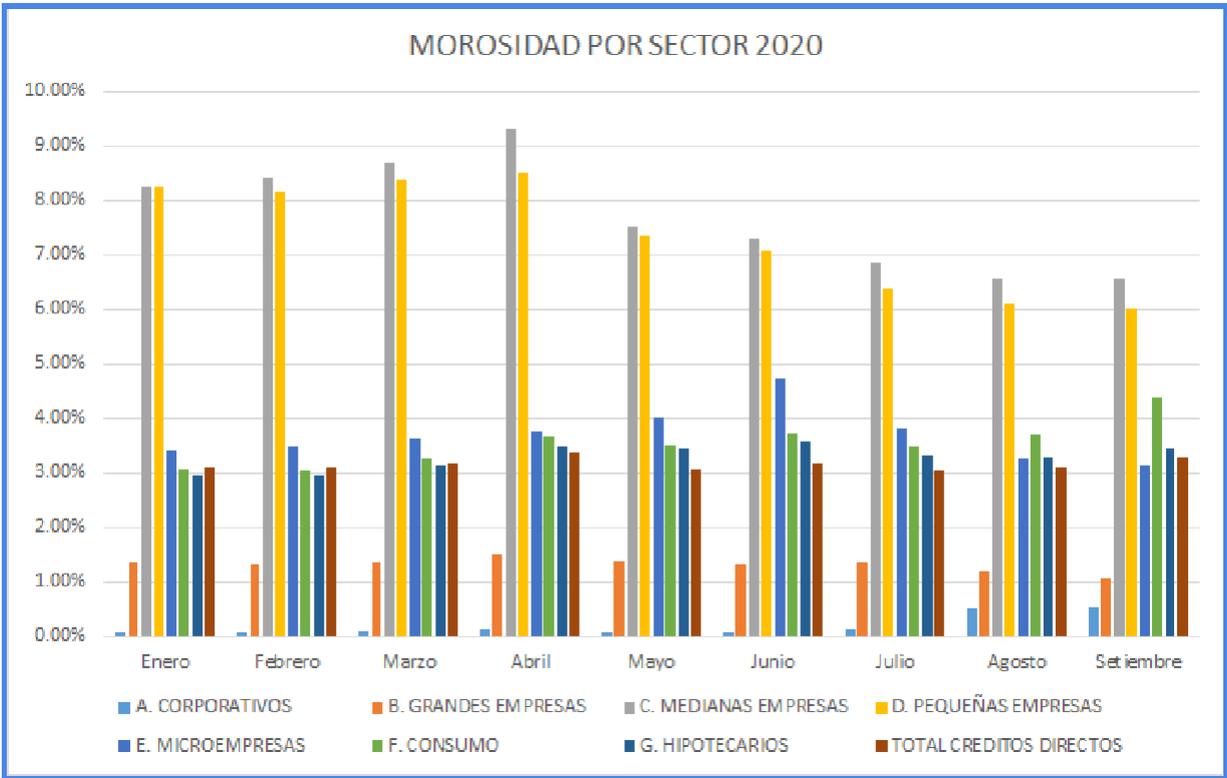


Figura 33. Índice de morosidad promedio por segmento, durante el año 2020.
 Fuente: Asociación de Bancos del Perú (2020).

3.3.- Análisis de stakeholders en la Banca Comercial

Con la finalidad de realizar una gestión efectiva de caudales de dinero, los bancos peruanos adquieren servicios de proveedores, tales como logística de valores, gestión de cajeros automáticos (ATM), así como distribución de tarjetas de crédito y débito (PROSEGUR, 2020). Estas empresas privadas brindan servicios de transporte de valores, incluyendo recojos, envíos y custodia de billetes, monedas y cheques. Asimismo, incluyen los cobros por recuento de billetes, materiales utilizados para la normalización de las cargas y otros servicios. En el Perú, solo existen dos empresas transportadoras de caudal, las cuales son Prosegur y Hermes, las cuales abastecen en su totalidad a toda la oferta nacional.

Por otro lado, se encuentran los entes reguladores del sector, tales como la Superintendencia de Banca y Seguros (SBS) y la Asociación de Bancos Peruanos (ASBANC). Estas empresas se encargan de promover el fortalecimiento del sistema proporcionando servicios de información, en igual proporción, a cada uno de los bancos que

operan en el país. Las empresas asociadas han firmado un acuerdo ético que las obliga a reportar el desarrollo de actividades, las cuales son discutidas en juntas mensuales con todos los involucrados.

CAPÍTULO 4. CONCLUSIONES

- Con respecto a la distribución de efectivo, se concluye que existe un sinnúmero de técnicas y algoritmos con las que se puede efectuar un programa de rutas que permitan abastecer tanto sucursales, como oficinas y cajeros automáticos de un sistema bancario. Aunque la investigación casuística de algoritmos tome lugar en Irán y Países Bajos, en la realidad peruana, también se pueden aplicar, debido a que existen similitudes con respecto a las realidades de otros países con respecto a sistemas de transporte, riesgo y niveles de operación, además de tomar en consideración que en una ciudad, la cantidad de puntos de servicio a abastecer superan la centena de unidades, por lo que se requiere una combinación efectiva de métodos, que permitan obtener tiempos mínimos de acción ante posibles desabastecimientos, además de distancias cortas que minimicen los costos de transporte en los que incurren las instituciones bancarias.
- Es importante unificar diversas herramientas que permitan optimizar la cadena de suministro, en particular, la del efectivo, dado que por separado, podrían no cubrir las expectativas o no optimizar la solución al máximo. Por ejemplo, en el caso 2.4, se utilizaron, bajo la nomenclatura de algoritmo genético, nueve distintos algoritmos secuenciales que optimizaron la solución en cada una de sus etapas. Además, en ese mismo caso, se incluyeron herramientas de la investigación de operaciones y la teoría de juegos, lo que optimizó la ruta seleccionada en gran magnitud.
- Con respecto a las técnicas de pronósticos, se concluye que no necesariamente los métodos computacionales avanzados actualmente usados para pronosticar la demanda son superiores en nivel de precisión con respecto a las técnicas tradicionales de *forecasting*. Estos últimos pueden llegar a tener niveles de desempeño similares, incluso utilizando una menor cantidad de recursos de almacenamiento y procesamiento de datos.
- Se concluye que la demanda de efectivo tiene un comportamiento particular, en comparación con otros productos. En este caso, los días festivos y los periodos para aplicar estacionalidad serán variables importantes a incorporar en un modelo de

pronóstico de requerimiento de efectivo en cajeros automáticos, debido a que, a partir de evidencia casuística, su inclusión permitió mejorar el desempeño y precisión del pronóstico en un 4.5%, tomando como medida de precisión al error SMAPE. Además, porque a partir de los hallazgos mostrados en el caso 2.3, se evidenció que el no considerarlos generaba caídas inesperadas en los cajeros automáticos y dimensionamientos excesivos en algunos días de semana.

- Complicar un modelo matemático al añadir variables o restricciones no representa una disminución de la optimalidad en gran medida. En el desarrollo del caso 2.4, el añadir el factor seguridad a las rutas de abastecimiento tan solo aumentó la distancia total recorrida en 5.36% y disminuyó el riesgo promedio por robos a 0%. Se concluye que en estudios posteriores, se debe encontrar una combinación óptima entre costos y nivel de complicación del modelo, dado que, si bien es cierto que mientras más factores se incluyen, el modelo tiene resultados más precisos, los costos podrían elevarse al punto de superar la situación actual, lo que llevaría a una paradoja.
- Es posible que una herramienta ajena al alcance del estudio permita desarrollar las problemáticas de la cadena de suministro del efectivo. Por ejemplo, se utilizó la teoría de juegos para estimar la probabilidad de éxito de un robo, modelando las estrategias seguidas por el delincuente y la posterior respuesta del sistema. De esta manera se integran herramientas diversas en un sector que parecía estar limitado a las herramientas del data mining y de predicción clásicas.
- Se concluye que al aplicar el método predictivo de redes neuronales, la red general de regresión generalizada (GRNN) resulta ser la de mejor desempeño para un tipo de data como la demanda del efectivo. En el caso 2.3, la capa GRNN resultó la más efectiva entre las estudiadas para la predicción de cajeros automáticos utilizando clústeres. Esta capa obtuvo un SMAPE promedio de 18.44%, 2% por encima de sus competidores en promedio.
- Existe una alta importancia con respecto a la digitalización de los procesos y actividades, por lo que, en el panorama actual, las herramientas investigadas en el presente trabajo, tendrán mayor acogida debido a la necesidad del sector bancario de innovar. En la actualidad, si bien algunas herramientas son aplicadas por los principales proveedores peruanos de servicios de abastecimiento de efectivo, las entidades tomarán mayor control con respecto a sus propias operaciones de

predicción y distribución, con la finalidad de garantizar elevados niveles de servicio al cliente, minimizando los costos operativos de las entidades bancarias.

- Durante el 2020, el sector bancario, en general, presenta una disminución en las utilidades, aunque se hayan realizado esfuerzos para impulsarlo a través de iniciativas como el fondo Reactiva Perú. Los ratios de rentabilidad expresaron que el sector experimentó caídas considerables en los tres primeros trimestres del año 2020. El retorno sobre activos (ROA) disminuyó un 77%, mientras que el retorno sobre el patrimonio (ROE) lo hizo en un 68%. Ello evidencia que, en primer lugar, las entidades bancarias han requerido mayores activos para generar igual o menor valor, con la finalidad de subsistir durante la pandemia producida por el brote de coronavirus, donde estos han sido entes cruciales para brindar estabilidad financiera a la sociedad, al ser su principal grupo de interés. En segundo lugar, entre las principales consecuencias se encuentran distintas iniciativas por parte de los bancos como la reducción de tasas activas a las que se ofrecen los créditos bancarios, con la finalidad de atraer clientes potenciales, así como la disminución de costos operativos para lograr una estabilidad en términos de rentabilidad.

CAPÍTULO 5. BIBLIOGRAFÍA

Aggarwal, C., (2015). *Data Mining*. New York, USA: Springer.

Álvarez, M. y Parra, J., (2013). *Teoría de grafos*. Universidad del Bío-Bío, Chillán.

Recuperado de

http://repobib.ubiobio.cl/jspui/bitstream/123456789/1953/3/Alvarez_Nunez_Marcelino.pdf

Asociación de Bancos del Perú (2020). Estadísticas del Sistema Financiero. Recuperado de <https://www.asbanc.com.pe/Paginas/Estadistica/Estadisticas.aspx?posTabActivo=1>

Asociación de Bancos del Perú (2020). La Asociación. Recuperado de <https://www.asbanc.com.pe/Paginas/Conocenos/Conocenos.aspx#:~:text=Asbanc%20es%20una%20instituci%C3%B3n%20gremial,fortalecimiento%20del%20sistema%20financiero%20privado.>

Ayala, S. (s/f). El sistema financiero peruano: *Gestiopolis*. Recuperado de <https://www.gestiopolis.com/sistema-financiero-peruano/>

Ballou, R., (2004). *Logística: Administración de la cadena de suministro*, México D.F., México: Prentice Hall.

Banco Central de Reserva del Perú (2020). Nota Semanal N° 40. Recuperado de <https://www.bcrp.gob.pe/docs/Publicaciones/Nota-Semanal/2020/ns-40-2020.pdf>

Bilir, C. y Doseyen, A., (2018). Optimization Of ATM And Branch Cash Operations Using An Integrated Cash Requirement Forecasting And Cash Optimization Model. *BMIJ*, 6 (1), 237-255.

- Bramer, M., (2020). *Principles of Data Mining*, London, England: Springer.
- Caicedo, A., Wagner, G. y Mendez, R., (2010). *Introducción a la teoría de grafos*, Armenia, Colombia: Ediciones Elizcom.
- Carreño, A., (2018). *Cadena de suministro y logística*, Lima, Perú: Fondo Editorial Pontificia Universidad Católica del Perú.
- Catal, C., Fenerci, A., Ozdemir, B., y Gulmez, O., (2015). Improvement of Demand Forecasting Models with Special Days. *Procedia Computer Science*, 59, 262–267.
- Chase, R., Jabos, F. y N. Aquilano., (2009). *Administración de Operaciones*, México D.F., México: Mc Graw Hill.
- Chopra, S. y Meindl, P., (2016). *Supply Chain Management. Strategy, Planning, and Operation*, United States of America: Pearson Education.
- Daza, J., Montoya, J. y Narducci, F., (2009). Resolución del problema de enrutamiento de vehículos con limitaciones de capacidad utilizando un procedimiento metaheurístico de dos fases. *Revista EIA*, 1 (12), 23-38. Recuperado de <http://www.scielo.org.co/pdf/eia/n12/n12a03.pdf>
- Ekinci, Y., Lu, J.-C., y Duman, E., (2015). Optimization of ATM cash replenishment with group-demand forecasts. *Expert Systems with Applications*, 42 (7), 3480 - 3490.
- Gestión (2020, 25 de Marzo). *Scotiabank reduce de 2.4% a 0.3% su proyección de crecimiento del PBI en 2020 debido al coronavirus*. Lima. Recuperado de <https://gestion.pe/economia/coronavirus-peru-scotiabank-reduce-de-24-a-03-su-proyeccion-de-crecimiento-del-pbi-en-2020-debido-al-covid-19-nndc-noticia/>
- Ghannadpour, S. F., & Zandiyeh, F. (2020). A new game-theoretical multi-objective evolutionary approach for cash-in-transit vehicle routing problem with time windows (A Real life Case). *Applied Soft Computing*, 106378.
- Golden, B., Raghavan, S. y Wasil, E., (2008). *The Vehicle Routing Problem: Latest Advances and New Challenges*, New York, United States of America: Springer.
- Heizer, J. y B. Render., (2009). *Principios de Administración de Operaciones*, México: Prentice Hall.
- Hermes (2020). Atención Integral de Cajeros Automáticos. Recuperado de <https://www.hermes.com.pe/logistica-de-valores/sector-financiero-y-comercial/atencion-integral-de-cajeros-automaticos>
- Hillier, F. y Lieberman, G., (2010). *Introducción a la investigación de operaciones*, México D.F., México: Mc Graw Hill.
- Jungnickel, D., (2008). *Graphs, Networks and Algorithms*, Augsburg, Germany: Springer.
- Kong, M., (2010). *Investigación de operaciones. Programación lineal, Problemas de transporte y Análisis de redes*, Lima, Perú: Fondo Editorial Pontificia Universidad Católica del Perú.
- Krajewski, L., Ritzman, L. y M. Malhotra., (2016). *Administración de Operaciones*, México: Prentice Hall.

Lizarzaburu, E. y del Brío, J., (2016). Evolución del sistema financiero peruano y su reputación bajo el índice Merco. Período: 2010-2014. *Suma de Negocios*, 7(16), 94–112.

Nan, Z., and W. Long. 2009. "The Mode of Cash Logistical Management Based on Cash Supply Chain in Bank Industry of China." *Financial Theory & Practice* no. 4.

Oliveira, A., (2004). *Heurísticas para problemas de ruteo de vehículos*. Recuperado de <https://www.fing.edu.uy/inco/pedeciba/bibliote/reptec/TR0408.pdf>

Prosegur (2020). Sistema de Gestión Integral de Cajeros Automáticos. Recuperado de <https://www.prosegur.com.pe/empresas-instituciones/cash/gestion-integral-de-cajeros-automaticos>

Rajamani, D., Neil, H. y C. Sriskandarajah (2005). A Framework to Analyze Cash Supply Chains. *Production and Operations Management*, 15(4).

RPP (2020, 27 de Abril). *Bank of China: El tercer banco más importante del mundo llega a Perú*. Lima. Recuperado de <https://rpp.pe/economia/economia/bank-of-china-el-tercer-banco-mas-importante-del-mundo-llega-a-peru-noticia-1261491?ref=rpp>

Sapag, N., Sapag, R. y J. Sapag., (2014). *Preparación y Evaluación de Proyectos*, México: Mc Graw Hill.

Superintendencia de Banca, Seguros y AFP (2020). Acerca de la SBS. Recuperado de <https://www.sbs.gob.pe/acercadelasbs#:~:text=La%20Superintendencia%20de%20Banca%2C%20Seguros,activos%20y%20financiamiento%20del%20terrorismo.>

Superintendencia de Banca, Seguros y AFP (2010). Glosario: Registro de indicadores y términos del sistema financiero.

Superintendencia de Banca, Seguros y AFP (2019). Indicadores de empresa bancarias. Recuperado de https://www.sbs.gob.pe/app/stats_net/stats/EstadisticaBoletinEstadistico.aspx?p=1#

Superintendencia de Banca, Seguros y AFP (2020). Indicadores de empresa bancarias. Recuperado de https://www.sbs.gob.pe/app/stats_net/stats/EstadisticaBoletinEstadistico.aspx?p=1#

Taha, H., (2012). *Investigación de operaciones*, México D.F., México: Prentice Hall.

Toth, P. y Vigo, D., (2017). *Vehicle Routing: Problems, Methods, and Applications*, United States of America: Society for Industrial and Applied Mathematics.

Van, R., (2014). Optimizing logistics processes in Cash Supply Chains (Tesis doctoral). Vrije Universiteit, Amsterdam.

Van, R., Coelho, L., Laporte, G. y Vis, I., (2016). An Inventory-Routing Problem with Pickups and Deliveries Arising in the Replenishment of Automated Teller Machines. *Transportation Science*, 50(3), 1077–1091.

Venkatesh, K., Ravi, V., Prinzie, A., Van Den Poel, D., (2014). Cash demand forecasting in ATMs by clustering and neural networks. *European Journal of Operational Research*, 232, 383-392.

Zaki, M. y W. Meira., (2020). *Data Mining and Machine Learning*, New York, U.S.A.: Cambridge.

