

**UNIVERSIDAD PERUANA UNIÓN**

**ESCUELA DE POSGRADO**

Unidad de Posgrado de Ingeniería



**Toma de decisiones para la venta en unidades de la línea de panes en una industria de alimentos en el Perú basado en redes bipartitas**

Tesis para obtener el Grado Académico de Maestro(a) en Ingeniería de Sistemas con mención en Dirección y Gestión en Tecnología de Información

**Autor:**

Sandrita Aracely Huamán Moreto  
Dany Isidro Villalta Salas

**Asesor:**

PhD. Javier Linkolk López Gonzales

Lima, enero de 2024

## DECLARACIÓN JURADA DE ORIGINALIDAD DE TESIS

Yo Javier Linkolk López Gonzales, docente de la Unidad de Posgrado de Ingeniería, Escuela de Posgrado de la Universidad Peruana Unión.

DECLARO:

Que la presente investigación titulada: **“Toma de decisiones para la venta en unidades de la línea de panes en una industria de alimentos en el Perú basado en redes bipartitas”** los autores, Sandrita Aracely Huamán Moreto y Dany Isidro Villalta Salas tiene un índice de similitud de 12 % verificable en el informe del programa Turnitin, y fue realizada en la Universidad Peruana Unión bajo mi dirección.

En tal sentido asumo la responsabilidad que corresponde ante cualquier falsedad u omisión de los documentos como de la información aportada, firmo la presente declaración en la ciudad de Lima, a los 17 días del mes de enero del año 2024



---

Javier Linkolk López Gonzales

## ACTA DE SUSTENTACIÓN DE TESIS

En Lima, Ñaña, Villa unión a 4 días del mes de enero del año 2024, siendo las 15:00 horas, se reunieron de forma online sincrónica, bajo la dirección del presidente del jurado Dra. Ethel Altez Ortiz, el secretario Dr. Josué Edison Turpo Chaparro y los demás miembros: Mg. Esteban Tocto Cano, Mg Lizeth Geanina Huanca Lopez y el asesor PhD. Javier Linkolk López Gonzales, con el propósito de administrar el acto académico de sustentación de Tesis de Maestría titulada "**Toma de decisiones para la venta en unidades de la línea de panes en una industria de alimentos en el Perú basado en redes bipartitas**", conducente a la obtención del grado de Magíster en Ingeniería de Sistemas con mención en Dirección y Gestión de Tecnologías de Información.

El presidente inició el acto académico de sustentación invitando a los candidatos hacer uso del tiempo determinado para su exposición. Concluida la exposición, el presidente invitó a los demás miembros del jurado a efectuar las preguntas, cuestionamientos y aclaraciones pertinentes, los cuales fueron absueltos por los candidatos. Luego, se produjo un receso para las deliberaciones y la emisión del dictamen del jurado. Posteriormente, el jurado procedió a dejar constancia escrita sobre la evaluación en la presente acta, con el dictamen siguiente:

Candidatos: SANDRITA ARACELY HUAMÁN MORETO y DANY ISIDRO VILLALTA SALAS

CALIFICACIÓN	ESCALAS			Mérito
	Vigesimal	Literal	Cualitativa	
Aprobado	16			Muy bueno

(\*) Ver parte posterior

Finalmente, el presidente del jurado invitó a los candidatos a ponerse de pie, para recibir la evaluación final. Además, el presidente del jurado concluyó el acto académico de sustentación, procediéndose a registrar a registrar las firmas respectivas.



Presidente



Secretario



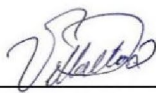
Asesor



Miembro



Miembro



Candidato



Candidato

## **Dedicatoria**

A mi querido hijo Derek Sebastián y a mi esposo José Eber Paz Vílchez, quien día a día me inspira y me motiva.

A mis queridos padres y familia, por sus oraciones a diario y sus consejos.

SAHM

A mi amada esposa Luzyeni Loida Huanca Cotohuanca, la cual me motiva a seguir adelante a alcanzar las metas.

A mis padres y familia, por el apoyo otorgado en todo momento.

DIVS

## **Agradecimiento**

En primer lugar, a Dios por su amor infinito amor y sus bendiciones derramadas en mi vida, a mis queridos padres por su apoyo incondicional, a mi pequeño regalo de Dios quién me motiva cada día a seguir adelante mi Derek Sebastián y a mi esposo por todo el apoyo y compañía de cada día.

A nuestro asesor el PhD. Javier Linkolk López Gonzales, por ser nuestro guía y apoyo en el desarrollo de esta investigación.

Al centro de bienes de la Universidad Peruana Unión, la industria de alimentos Unión, por permitirnos poder realizar este estudio y dejar los aportes para la mejora continua.

SAHM

A Dios en primer lugar ya que sin su amor y cuidado no podríamos obtener nada, el cual a través de sus bendiciones no muestra que podemos salir adelante, a mi esposa

Luzyeni Huanca por todo el apoyo con el cual me motiva a seguir adelante.

A nuestro asesor el PhD. Javier Linkolk López Gonzales por su guía y orientación en el desarrollo de esta investigación.

Al centro de producción de bienes Unión de la Universidad Peruana Unión, por concedernos la oportunidad de realizar la investigación en ella, y dejar así un aporte a la ciencia.

DIVS

## Tabla de contenido

Resumen .....	ix
Capítulo 1 .....	11
1. Planteamiento del Problema .....	11
1.1. Descripción de la problemática.....	11
1.2. Formulación de la pregunta de investigación .....	12
1.2.1. Problema general .....	12
1.3. Objetivos de investigación.....	13
1.3.1. Objetivo general:.....	13
1.4. Presuposición filosófica .....	13
Capítulo II.....	14
2. Fundamentos teóricos de la investigación .....	14
2.1. Antecedentes de la investigación .....	14
2.2. Bases teóricas.....	16
2.2.1. Bipartitas .....	16
2.2.2. CRISP-DM.....	18
Capítulo III .....	22
3. Diseño metodológico .....	22
3.1. Tipo de investigación .....	22
3.2. Diseño de la investigación .....	22
3.3. Población.....	22
Capítulo IV .....	23
4. Propuesta y desarrollo.....	23
4.1. Conocimiento del negocio .....	23
4.2. Conocimiento de los datos .....	25
4.3. Exploración de los datos .....	26
4.4. Modelado de los datos .....	26

4.5.	Evaluación.....	27
4.6.	Implementación del modelo.....	27
	Capítulo V.....	28
5.	Resultados.....	28
	Capítulo V.....	32
6.	Conclusiones y recomendaciones.....	32
6.1.	Conclusiones.....	32
6.2.	Recomendaciones.....	33
	Referencias bibliográficas.....	34

## Índice de figuras

<i>Figura 1.</i> Ejemplo de toma de decisiones. ....	13
<i>Figura 2.</i> Metodología CRISP-DM [26]. ....	19
<i>Figura 3.</i> Mapa de procesos de Unión. ....	24
<i>Figura 4.</i> Matriz de adyacencia (a) y gráfico bipartito (b) para la ciudad de Chimbote en 2023. 28	
<i>Figura 4.</i> Matriz de adyacencia (a) y gráfico bipartito (b) para la ciudad de Juliaca en 2023.....	29
<i>Figura 5.</i> Matriz de adyacencia (a) y gráfico bipartito (b) para la ciudad de Lima en 2023. ....	29
<i>Figura 6.</i> Matriz de adyacencia (a) y gráfico bipartito (b) para la ciudad de Piura en 2023. ....	30
<i>Figura 7.</i> Matriz de adyacencia (a) y gráfico bipartito (b) para la ciudad de Tarapoto en 2023... 30	
<i>Figura 8.</i> Matriz de adyacencia (a) y gráfico bipartito (b) para la ciudad de Trujillo en 2023. ....	31

## **Resumen**

En estos últimos años, la minería de datos a tomado gran importancia en las empresas, porque el uso de la misma ayuda en la toma de decisiones de los directivos. Porque los problemas de toma de decisiones se vuelven cada vez más complicados.

Por lo tanto, en este artículo se propone un algoritmo probabilístico de toma de decisiones de recomendación, basado en una proyección de red bipartita. En primer lugar, combinando la información de toma de decisiones de los almacenes de venta de los productos y la venta de los panes, se establece un gráfico bipartito que conecta al almacén y al producto. Para ello se utilizó la metodología CRISP-DM. Los resultados muestran que nuestro método propuesto supera a un método de predicción simple, logrando la oportuna toma de decisiones.

Palabras clave: Bipartites; Toma de decisión; CRISP-DM.

## **Abstract**

In recent years, data mining has taken on great importance in companies, because its use helps managers make decisions. Because decision-making problems are becoming more and more complicated.

Therefore, this paper proposes a probabilistic recommendation decision-making algorithm based on a bipartite network projection. First, by combining the decision-making information from the product sales warehouses and the sales panels, a bipartite graph is established that connects the warehouse and the product. For this, the CRISP-DM methodology was used. The results show that our proposed method outperforms a simple prediction method, achieving timely decision making.

Keywords: Bipartites; Decision making; CRISP-DM

## Capítulo 1

### Planteamiento del Problema

#### 1.1. Descripción de la problemática

En la actualidad vivimos en un mundo muy competitivo, donde las empresas se están enfocando en las ventas, la capacidad de realizar una previsión precisa se ha convertido en un factor fundamental para el logro del éxito en cualquier tipo de negocio. Los pronósticos inexactos para la venta de productos pueden generar condiciones de falta de stock o costos inflados debido al exceso de las mismas. Cuando un producto específico está agotado, las empresas pierden las ganancias de la venta. Si la falta de stock ocurre muy seguido, ocasiona insatisfacción en los consumidores que puede llevar a cambiarse a la competencia. [1,2] Y los problemas de toma de decisiones se vuelven cada vez más complicados. [3–5]

Debido a esto, el Centro de Producción de Bienes Unión de la Universidad Peruana Unión, dedicada a la producción de alimentos saludables, no es ajeno a todas las necesidades de las demás empresas, también requiere surgir y crecer en el mercado a través de pronósticos de ventas, porque con el incremento de pedidos y ventas, la producción se realiza sin apoyo de un método de pronóstico adecuado, careciendo de herramientas tecnológicas que le faciliten el trabajo y la identificación oportunas de las mismas. Si bien es cierto, el centro de producción cuenta herramientas a través de juicio de expertos que le permite realizar pronósticos con los datos anteriores a las ventas, esto no contempla el comportamiento ni la variabilidad de la demanda basada en los registros históricos, lo que puede cambiar de un periodo a otro. El presente artículo se enfoca en el desarrollo de un modelo de redes neuronales para la toma de decisiones en el pronóstico de ventas de unidades de pan en una industria de alimentos en Perú, aplicando la metodología CRISPDM (Cross-Industry Standard Process for Data Mining). La industria alimentaria en Perú, con un

énfasis particular en la línea de panes, representa un sector económico significativo que enfrenta constantemente el desafío de adaptarse a las fluctuantes demandas del mercado y las preferencias cambiantes de los consumidores. En este contexto, la capacidad de predecir con precisión las ventas futuras es crucial para la gestión eficiente de la cadena de suministro, la optimización de los niveles de inventario y la planificación estratégica.

Este estudio no solo contribuye a la literatura académica en el campo de la minería de datos y las ciencias de la computación, sino que también proporciona insights valiosos para los profesionales de la industria alimentaria, ofreciendo una herramienta poderosa para la toma de decisiones basada en datos. Con un enfoque en el mercado peruano, el artículo también pone de relieve la importancia de adaptar las soluciones de minería de datos a contextos y desafíos locales específicos, un aspecto crucial para el éxito en la aplicación de tecnologías avanzadas en diferentes regiones geográficas y sectores económicos. Este estudio está estructurado de la siguiente manera. El capítulo 1, describe el planteamiento del problema, el capítulo 2 describe los fundamentos teóricos de la investigación, el capítulo 3 muestra el diseño de la metodología, en el 4 capítulo muestra los resultados y finalmente en el capítulo 5 muestra las conclusiones y recomendaciones.

## **1.2. Formulación de la pregunta de investigación**

### **1.2.1. Problema general**

¿De qué manera el algoritmo probabilístico basado en una proyección de red bipartita ayuda a la toma de decisiones para la venta en unidades de la línea de panes en una industria de alimentos en el Perú?

### 1.3. Objetivos de investigación

#### 1.3.1. Objetivo general:

Determinar el modelo de marketing de servicios apoyado en ecuaciones estructurales es predictor de la satisfacción del usuario e imagen institucional en la Universidad Peruana Unión.

### 1.4. Presuposición filosófica

En este estudio presenta una presuposición filosófica basada en las Sagradas escrituras de la Biblia, en la siguiente imagen podremos encontrar algunos ejemplos de personajes que tuvieron que tomar decisiones buenas o malas.

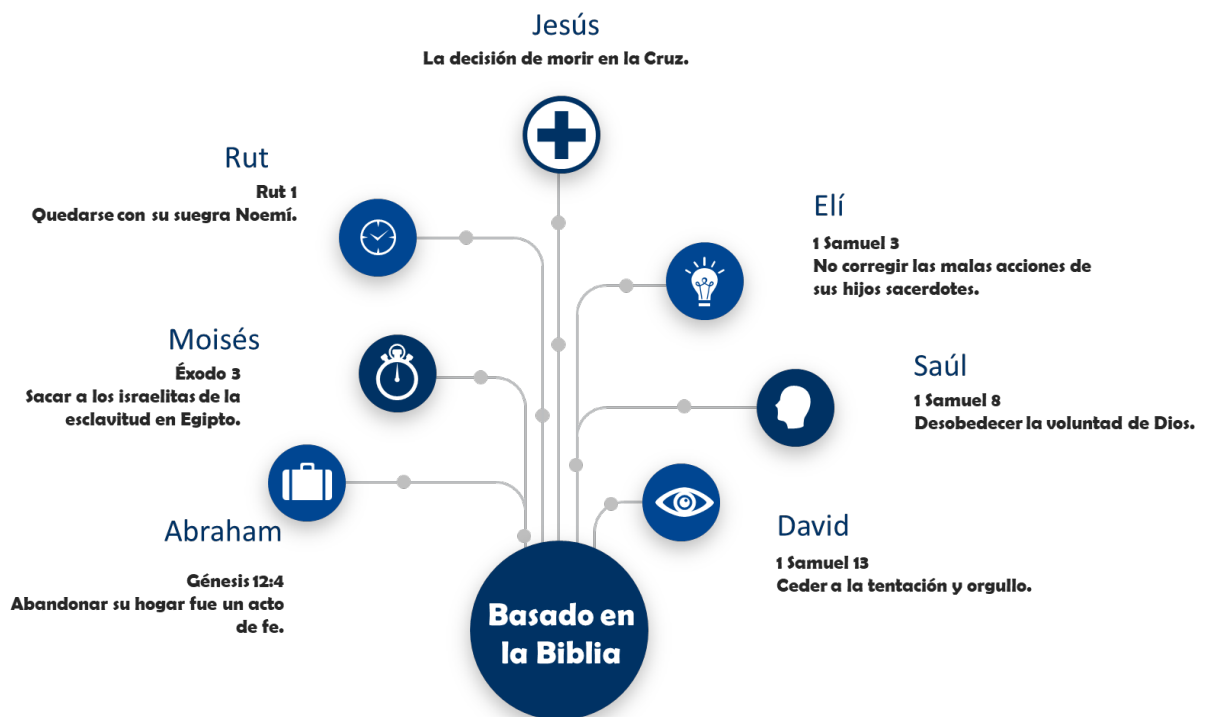


Figura 1. Ejemplo de toma de decisiones.

Fuente: Elaboración propia

## Capítulo II

### Fundamentos teóricos de la investigación

#### 2.1. Antecedentes de la investigación

Muchas empresas, fabricantes, distribuidores y minoristas suelen enfrentarse a dos problemas críticos. El primero es decidir la cantidad de producción o pedido de productos, el segundo es hacer la asignación de un recurso limitado. Por lo tanto, una compensación entre pedir demasiado o poco es una decisión crucial para un minorista. [6] Hay algunos factores importantes que tienen un impacto en las ventas futuras. Estos factores pueden identificarse analizando los patrones de las ventas totales de una tienda minorista o de un producto específico. Importante conocer, que cada producto tiene un nivel diferente de dificultad de pronóstico. Algunos productos tienen un consumo estable a lo largo del año. [7,8] Otros artículos, como moda y muebles, contienen tendencias y estacionalidad en su patrón de ventas que causan complejidad en el proceso de pronóstico. [2]

Cao en su investigación sobre alojamiento y con la continua expansión de la escala de alojamiento compartido, le resulta bastante complicado para los usuarios buscar servicios o información adecuados. Por ello, proponen un algoritmo probabilístico de toma de decisiones de recomendación difusa y vacilante basado en una proyección de red bipartita [3]

Wienclaw menciona que, una tendencia es la dirección persistente en la cual algo se mueve en el corto, mediano o largo plazo. Identificar esta tendencia permite a los administradores mejorar la planeación al conocer las necesidades futuras. [9] Por ello, la toma de decisiones es un factor clave de éxito y una buena decisión genera valor para la empresa. Entre ellas las previsiones de ventas, que es un factor clave de éxito en la gestión de las empresas. [8,10]

Los algoritmos de engrosamiento se han utilizado con éxito como una estrategia poderosa para abordar problemas de aprendizaje automático con uso intensivo de datos

definidos en redes bipartitas, como la agrupación, la reducción de dimensionalidad y la visualización.[11] Recientemente, los estudios de predicción de enlaces en redes complejas y a gran escala se han convertido en particular en el foco de interés para investigadores de diversos campos científicos. Muchas redes complejas creadas a partir de datos del mundo real contienen una estructura bipartita por naturaleza. Las redes bipartitas son un tipo de redes complejas que representan las interacciones entre los diferentes grupos de nodos. [12] método acreditable para comprimir redes bipartitas, sino que también destaca una posible forma de resolver mejor un desafío de larga data en la ciencia de la información moderna: cómo hacer una recomendación personal. [13] Este trabajo es un estudio del algoritmo de recomendación personal que emplea la proyección de una red bipartita ponderada de consumo-producto [13]

Se han realizado diversas investigaciones respecto a los métodos utilizados para las predicciones. Podemos encontrar en esta investigación, donde propusieron la proyección de red bipartita para la predicción de la asociación de miARN-enfermedad (BNPMDA) basada en asociaciones conocidas de miARN-enfermedad, similitud funcional de miARN, similitud semántica de enfermedad y similitud del núcleo del perfil de interacción gaussiana y concluyeron con un excelente desempeño de BNPMDA se atribuye principalmente a los siguientes factores importantes. En primer lugar, el algoritmo básico de recomendación bipartita supera a algunos métodos clásicos anteriores para recomendar artículos a los usuarios en el ámbito empresarial. Esta superioridad garantizó la eficacia básica de nuestro método propuesto. En segundo lugar, el algoritmo de recomendación bipartita mejorado podría reflejar las calificaciones de sesgo de las enfermedades y los miARN en el proceso de inicialización de recursos y el proceso de transferencia de recursos. Este tipo de algoritmo de calificación integrada podría aprovechar al máximo la información de similitud tanto de los miARN como de las enfermedades para mejorar aún

más la precisión de la predicción [14]. Propusieron un método de predicción de enlaces para la red bipartita compleja. Llegando a la conclusión. Los resultados de los experimentos demuestran que nuestro método propuesto puede resolver eficazmente el problema de predicción de enlaces en redes bipartitas complejas [15].

## **2.2. Bases teóricas**

### **2.2.1. Bipartitas**

La teoría de las redes bipartitas ha sido ampliamente estudiada por diversos autores en múltiples disciplinas. Una de las bases teóricas fundamentales proviene de los trabajos pioneros de Euler y König en teoría de grafos, donde establecieron los fundamentos matemáticos para el estudio de las redes y grafos bipartitos [16,17]. Posteriormente, en la década de 1930, Hall formuló el Teorema del emparejamiento bipartito, que establece condiciones necesarias y suficientes para la existencia de un emparejamiento perfecto en una red bipartita, lo cual es crucial en numerosas aplicaciones prácticas [18].

Otro enfoque teórico importante proviene de la teoría de conjuntos, que proporciona una base conceptual para entender las relaciones entre los nodos de una red bipartita. Autores como Frank Harary han contribuido significativamente a la aplicación de la teoría de conjuntos en el estudio de las redes bipartitas, explorando aspectos como la representación de las redes bipartitas mediante matrices de adyacencia y las implicaciones de las intersecciones de conjuntos en la estructura y el análisis de estas redes [19].

Una red bipartita, una forma especializada de red compleja, se caracteriza por la presencia de dos tipos distintos de nodos y bordes que establecen conexiones exclusivamente entre nodos de diferentes tipos. Este paradigma estructural resulta particularmente ventajoso cuando se modelan relaciones entre entidades con características dispares, pero donde las conexiones entre ellas tienen una importancia considerable. En el contexto examinado en

este artículo, la aplicación de redes bipartitas se ha empleado estratégicamente para iluminar la intrincada dinámica que rige las relaciones entre varios almacenes y la venta de pan.

Las redes bipartitas son un tipo de red donde dos conjuntos de nodos (productos y usuarios) se conectan si existe una interacción entre ellos (por ejemplo, una compra). Estas redes se han utilizado para modelar la venta de productos en línea con gran éxito [20].

Chen y otros propone un método basado en aprendizaje profundo para predecir la demanda de productos en redes bipartitas. El método logra resultados de última generación en comparación con métodos tradicionales [21].

Wang y otros, propone un sistema de recomendación que toma en cuenta la heterogeneidad de los usuarios para recomendar productos con mayor precisión [22].

Liu y otros, utiliza técnicas de aprendizaje automático para detectar transacciones fraudulentas en redes bipartitas. El método logra una alta precisión en la detección de fraude [23].

Y Zhao con otros investigadores, utiliza grafos de Markov para modelar la competencia entre productos en redes bipartitas. El modelo permite a los vendedores comprender mejor la dinámica de la competencia y tomar mejores decisiones estratégicas [24].

La razón detrás del aprovechamiento de las redes bipartitas en este estudio radica en la necesidad de capturar y analizar de manera integral la interacción matizada entre las instalaciones de almacenamiento y los aspectos comerciales de la distribución de pan. Las ciudades, como nodos de un tipo, representan las ubicaciones de estos almacenes, mientras que los nodos del otro tipo denotan las transacciones de venta de pan. Los bordes, que sirven como conductos de conexión, vinculan exclusivamente los almacenes con las ventas de pan, proporcionando una representación visual clara de la compleja red de relaciones dentro de la red de logística y distribución de la cadena de suministro de pan.

Este enfoque metodológico no sólo facilita una comprensión más profunda de las intrincadas relaciones entre las instalaciones de almacenamiento y la dinámica de las ventas de pan, sino que también establece un marco para la aplicación más amplia del análisis de redes bipartitas en dominios donde entidades de distintos tipos participan en interacciones cruciales. Al explorar el caso específico de los almacenes y las ventas de pan, este artículo contribuye a promover el discurso sobre la ciencia de las redes y los sistemas complejos, mostrando la versatilidad de las redes bipartitas para desentrañar estructuras relacionales complejas dentro del contexto de la logística y las actividades económicas asociadas con un producto específico. en este caso, pan.

### **2.2.2. CRISP-DM**

Este trabajo se desarrolló utilizando la metodología CRISP-DM (Cross Industry Process Model for Data Mining) que sigue un enfoque orientado a objetivos, que continúa siendo ampliamente aceptado en proyectos de minería de datos a través de algoritmos de minería de aprendizaje automático de datos. [27, 28] Esta metodología proporciona un enfoque de ciclo de vida en proyectos de inteligencia artificial aplicada [29] y se considera como ideal para el descubrimiento de conocimiento. Consta de 6 fases secuenciales y recurrentes: (i) Conocimiento del negocio, (ii) Conocimiento de los datos, (iii) Exploración de datos, (iv) Modelado de datos, (v) Evaluación y (vi) Implementación de modelo de datos. Estas etapas se describen a continuación.

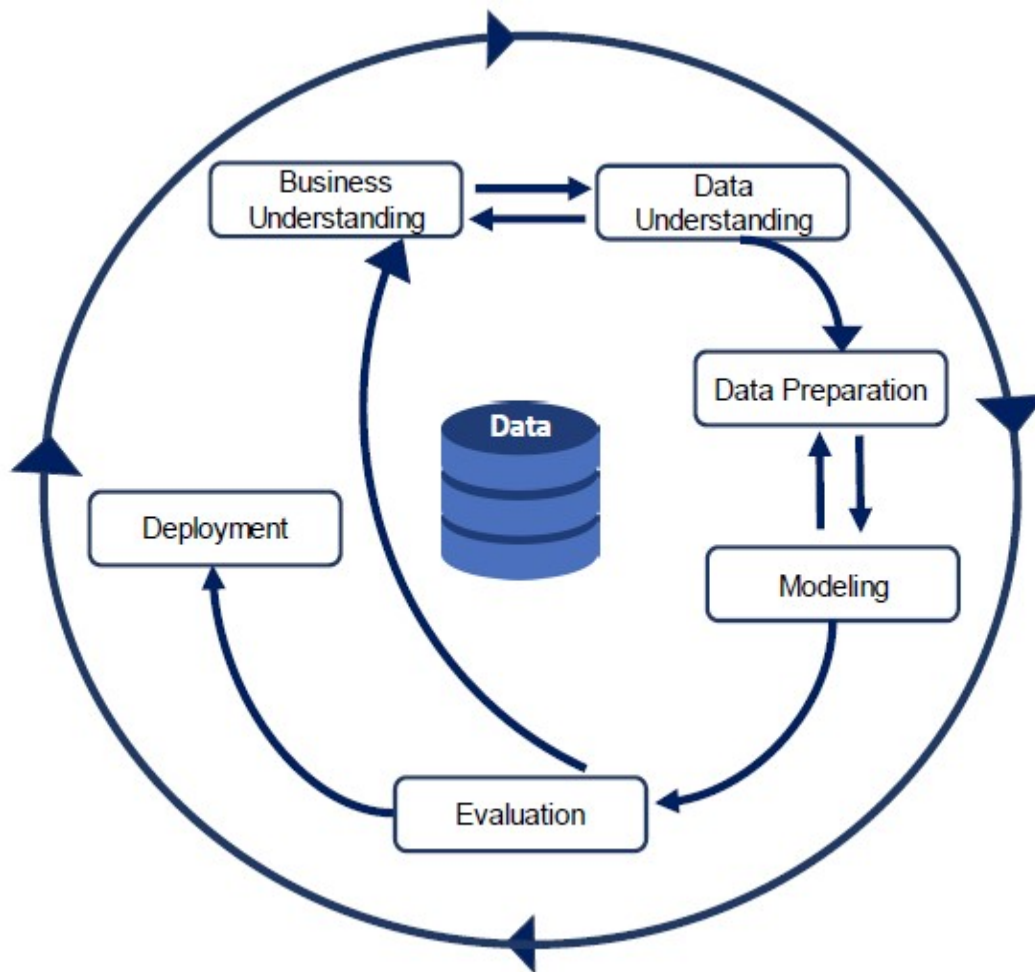


Figura 2. Metodología CRISP-DM [26].

Fuente: Elaboración propia

### Conocimiento del negocio

La evaluación de la situación del negocio es fundamental para obtener una visión general de los recursos disponibles y necesarios. En esta etapa, resulta de gran importancia establecer claramente el objetivo de la minería de datos. Para ello, se debe describir el tipo de minería de datos a emplear, por ejemplo, la clasificación y definir los criterios de éxito, como la precisión. Es esencial desarrollar un plan de proyecto obligatorio para continuar con el proceso de manera efectiva [36,37]. En esta fase se obtiene conocimientos sobre las actividades de Unión, su estructura organizacional, ambiente interno, relaciones e interacciones. Estos detalles se realizarán para comprender mejor el negocio y poder plasmar bien sus objetivos.

### Conocimiento de los datos

En esta etapa, resultan imprescindibles la obtención de datos de diversas fuentes, su exploración y descripción, así como la verificación de su calidad. En términos más específicos, la guía del usuario detalla el proceso de descripción de datos, que involucra análisis estadístico y la identificación de atributos y sus relaciones [36,37]. Se requiere el listado de los recursos del proyecto, esto incluye la carga de datos. Además de la recolección inicial de los datos para poder establecer un primer contacto con el problema y familiarizarse.

### Exploración de datos

La etapa de selección de datos implica establecer criterios precisos para incluir o excluir información. Si los datos presentan deficiencias, es factible abordar este problema mediante su limpieza. Adicionalmente, en función del modelo previamente definido en la fase inicial, es necesario generar atributos derivados. Es relevante destacar que existen múltiples métodos disponibles para llevar a cabo estos pasos, los cuales varían según el modelo utilizado [36,37].

### Modelado de datos

Durante la etapa de modelado de datos, se procede a seleccionar la técnica de modelado más adecuada y se construye tanto el caso de prueba como el modelo en sí. Todas las técnicas de minería de datos pueden ser empleadas en este proceso. Sin embargo, la elección específica dependerá tanto del problema comercial en cuestión como de los datos disponibles. Lo más relevante es justificar claramente la elección realizada. Para la construcción del modelo, resulta esencial establecer parámetros específicos. Finalmente, para evaluar el modelo correctamente, es apropiado compararlo frente a criterios de evaluación y seleccionar aquellos que muestren mejores resultados [36,37].

## Evaluación

Durante la etapa de evaluación, se contrastan los resultados con los objetivos comerciales definidos previamente. Es crucial realizar una interpretación adecuada de los resultados y establecer las acciones adicionales que se deben implementar. Asimismo, es importante llevar a cabo una revisión exhaustiva de todo el proceso en su conjunto [36,37]

## implementación de modelo datos

La guía del usuario proporciona una descripción general de la fase de implementación, que puede adoptar la forma de un informe final o un componente de software. En esta fase, se planifica la implementación, así como la supervisión y el mantenimiento del proyecto, tal como se detalla en la guía [36,37].

## **Capítulo III**

### **Diseño metodológico**

#### **3.1. Tipo de investigación**

Valderrama [26] menciona que la investigación tecnológica consiste en trabajos sistemáticos basados en conocimientos existentes, obtenidos mediante investigación y/o experiencia práctica. Se dirigen a la fabricación de nuevos materiales, productos o dispositivos; a establecer nuevos procesos, sistemas y servicios, o la mejora sustancial de los mismos.

Cordero [ 27] también indica que es reproducción de los saberes en las prácticas, para emplearlos en beneficio del conjunto de los participantes en estos procedimientos y en el contexto social. Porque da soluciones para la toma de decisiones de los administrativos del centro de bienes de producción.

#### **3.2. Diseño de la investigación**

Hernández, Fernández y Batista [28] mencionan que las investigaciones no experimentales se caracterizan por los estudios que se efectúan sin la manipulación voluntaria de las variables. En otros términos, solo se observan los acontecimientos fenomenológicos en su real contexto en el marco del análisis.

#### **3.3. Población**

Los registros de las ventas de enero del 2018 hasta abril del 2023. El tamaño de la población es mayor a 40 mil registros de los años mencionados.

## **Capítulo IV**

### **Propuesta y desarrollo**

En este capítulo se presenta los procedimientos del desarrollo de la red bipartitas para la toma de decisiones de la venta en unidades de la línea de panes del centro de bienes de Universidad Peruana Unión.

#### **4.1. Conocimiento del negocio**

En esta fase se obtiene conocimientos sobre las actividades del centro de bienes de la Universidad Peruana Unión, su estructura organizacional, ambiente interno, relaciones e interacciones. Estos detalles se realizarán para comprender mejor el negocio y poder plasmar bien sus objetivos.

Unión es un Centro de Producción de la Universidad Peruana Unión, dedicado esencialmente al rubro de la panificación desde 1929. Las líneas de producción son panes, cereales, galletas, pasteles, jugos, panetones y otros productos que se elaboran con materia prima de la más alta calidad que pasan por un análisis riguroso por especialistas del área de Aseguramiento de Calidad que aseguran sabor y nutrición en todos nuestros productos.

Actualmente Unión, está estructurado mediante el mapa de procesos con procesos estratégicos, operativos y de soporte, como se puede apreciar en la siguiente imagen:

**MAPA DE PROCESOS**

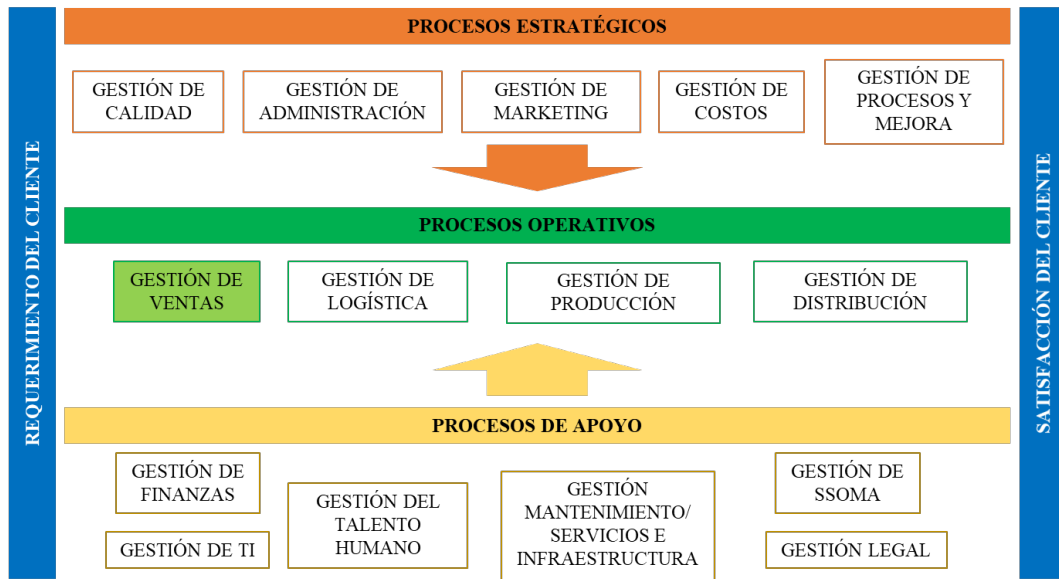


Figura 3. Mapa de procesos de Unión.

Fuente: Elaboración propia

En los procesos operativos, encontramos a la gestión de ventas, gestión de producción, gestión de logística y la gestión de distribución. Estos procesos operativos son esenciales y giran en torno a la producción y venta de productos. En los procesos de la gestión de ventas, encontramos a tres áreas o canales que lo conforman: el canal clásico que se dedica a la venta directa mediante agentes de ventas de bodegas, distribuidores; el canal moderno a la venta a los supermercados y el canal Ecommerce a la venta en Outlets y corporativo. Las ventas de los productos se realizan a diario y son registradas en una base de datos que tiene el centro de bienes Unión. Estos registros se almacenan por la descripción del producto, cantidad, categoría de producto, tipo de almacén que realiza la venta (canal clásico, moderno, ecommerce), entre otros.

La categoría seleccionada para esta investigación es de la línea de panes con los siguientes nombres: Pan Americano Sandwich, Pan Blanco Familiar Superbom, Pan Fuente, Pan

Integral Familiar Superbom, Pan Integral Mediano Unión, Pan Integral Sándwich, Pan Guerrero, Pan Fibra Integral, Pan Granos y Semillas, Pan Multisemillas con Avena, Pan Multisemillas con Miel de Abeja.

Con el transcurrir de los años, Unión tiene una marca posicionada en el mercado reconocido como una línea de productos saludables, que forma parte de la red mundial de productos alimenticios de la Iglesia Adventista del Séptimo Día. Hoy es una de las empresas líderes en nutrición y prevención de salud con 3 marcas: Unión, Vita rey y Superbom. Este estudio busca realizar un modelo de red bipartitas de la línea de producción de panes, por ser la línea más vendida y ser de ayuda en la toma de decisiones para la administración.

#### **4.2. Conocimiento de los datos**

Familiarizarse con los datos teniendo presente los objetivos del negocio

Recopilación inicial de datos: Los registros de las ventas de los productos se almacenan en un ERP que tiene Unión, estos registros se guardan diariamente. Para realizar el modelo de red bipartitas se seleccionó los registros del mes de enero del 2018 hasta el 30 de abril del 2023.

Descripción de los datos: Los registros son las ventas diarias de los productos de las diversas líneas que ofrece Unión.

Exploración de los datos: Según los datos almacenados, se pudo evidenciar que todos los campos solicitados no son requisito importante para registrar, a excepción de algunos campos y por ello se seleccionó el producto y el almacén de venta.

Verificación de calidad de datos: Se realizó la comprobación que todos los campos estén completos.

Además, se analizan los datos recopilados para identificar patrones, tendencias y posibles problemas de calidad de datos.

### **4.3. Exploración de los datos**

En la actualidad los registros de ventas del centro de bienes se almacén en el ERP, por ello se seleccionó los campos requeridos para realizar la red bipartita.

Selección de los datos: Los campos requeridos son: descripción del producto vendido, cantidad de producto vendido, fecha de producto vendido, almacén que vende (lugar)

Limpieza de datos: Se revisó que todos los campos estén completos y no haya ningún campo nulo o vacío, además se eliminó los datos duplicados, inconsistentes o faltantes para garantizar la calidad de los datos.

Construcción de datos: Se armó una sola data de los años seleccionados y de los campos requeridos para poder integrarlos. Se transformó los datos en el formato adecuado para la construcción de la red bipartita, creando una matriz que represente las ventas de cada tipo de pan en cada almacén.

### **4.4. Modelado de los datos**

Para el modelado de la red bipartitas se seleccionó este modelo, ya que permite analizar las ventas de los productos en relación con los almacenamientos de distribución de pan. Las ciudades, como nodos de un tipo, representan las ubicaciones de estos almacenes, mientras que los nodos del otro tipo denotan las transacciones de venta de pan. Los bordes, que sirven como conductos de conexión, vinculan exclusivamente los almacenes con las ventas de pan, proporcionando una representación visual clara de la compleja red de relaciones dentro de la red de logística y distribución de la cadena de suministro de pan.

#### **4.5. Evaluación**

Se analizó la estructura de la red bipartita y se evaluó la capacidad para representar eficazmente las relaciones entre los tipos de pan y los almacenes de venta. Se pueden utilizar medidas como la densidad de la red y la centralidad de los nodos para evaluar su calidad.

#### **4.6. Implementación del modelo**

Se integró la red bipartita en los sistemas existentes de gestión de ventas y la gestión de producción para utilizarla en la optimización de la distribución y la estrategia de marketing de los productos de pan.

Se monitorea el rendimiento de la red bipartita en producción y se realizan ajustes según sea necesario para garantizar su eficacia a largo plazo en la maximización de las ventas de unidades de pan en los almacenes.

## Capítulo V

### Resultados

Como hemos visto a través de esta investigación, las redes bipartitas son muy robustas para poder mostrar las interacciones como en caso de los almacenes de distribución y los panes que se comercializan a través de ellos, siendo esta una de las características principales del uso de las redes bipartitas.

Los resultados presentados en esta sección muestran la matriz de adyacencia y el gráfico bipartito para la venta de pan entre ciudades del Perú.

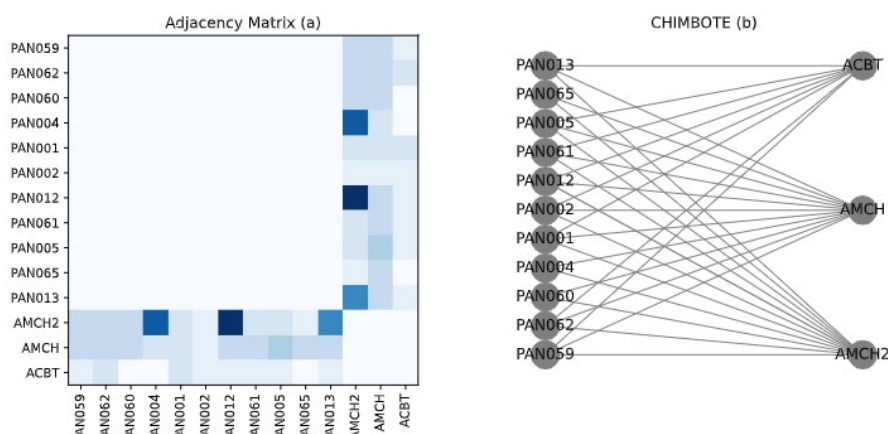


Figura 4. Matriz de adyacencia (a) y gráfico bipartito (b) para la ciudad de Chimbote en 2023.

Fuente: Elaboración propia

Las ventas en unidades las cuales corresponden a Chimbote tienen como más representativos el Código PAN004, que representa al Pan Fuente, y el Código PAN012 que representa al Pan Integral Mediano Unión, ambos en el almacén AMCH2.

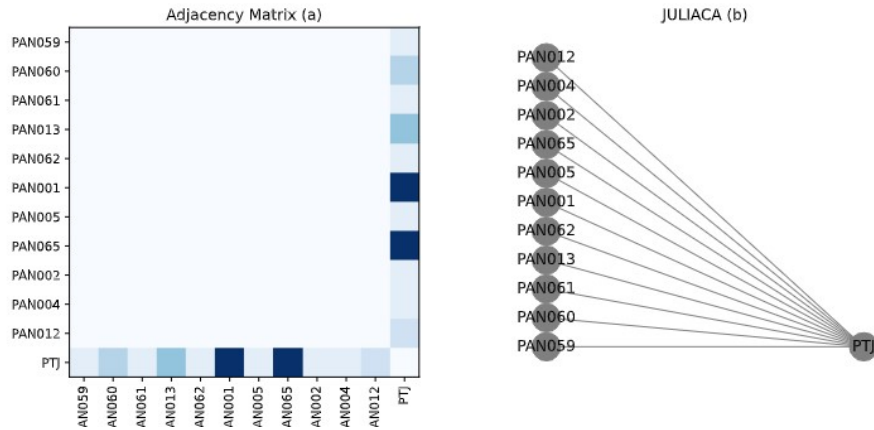


Figura 5. Matriz de adyacencia (a) y gráfico bipartito (b) para la ciudad de Juliaca en 2023. Fuente: Elaboración propia

El almacén de Juliaca indicado con el Código PTJ, se presentan los códigos más resaltantes los panes PAN001 y el PAN065 correspondientes al Pan Americano Sandwich y el Pan Molde Blanco Guerrero.

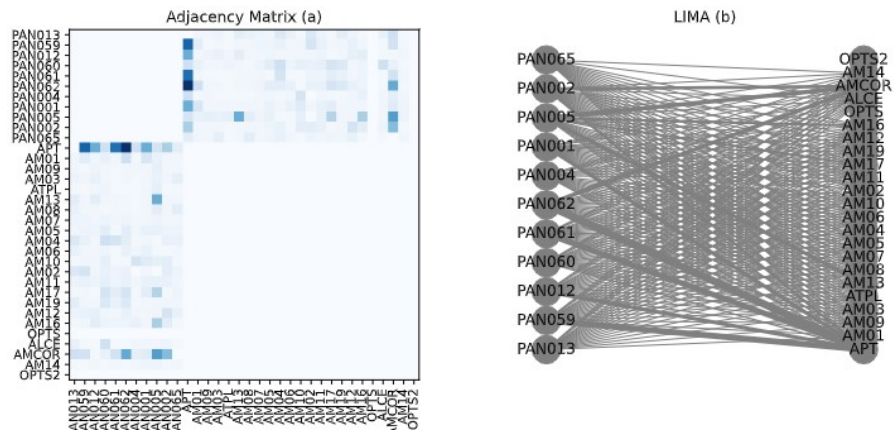


Figura 6. Matriz de adyacencia (a) y gráfico bipartito (b) para la ciudad de Lima en 2023. Fuente: Elaboración propia

En el almacén de Lima (APT) se observa que los panes que son más recurrentes son PAN062, PAN061, PAN059, PAN012 Y PAN001 los cuales hacen referencia a Pan Multisemillas con Avena, Pan Multisemillas con Miel de Abeja, Pan Fibra Integral, Pan Integral Mediano Unión y Pan Americano Sandwich respectivamente

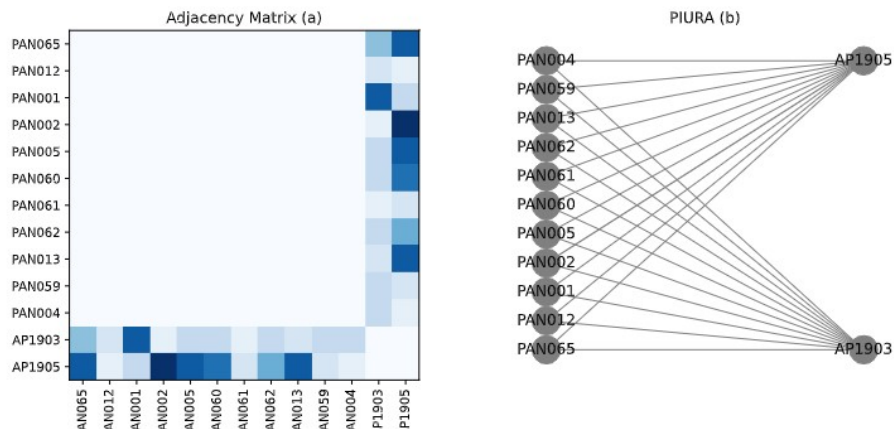


Figura 7. Matriz de adyacencia (a) y gráfico bipartito (b) para la ciudad de Piura en 2023.  
Fuente: Elaboración propia

En el almacén de Piura, se observa que el pan más recurrente es PAN001, PAN002, PAN005, PAN060, lo que representa al Pan Americano Sandwich, Pan Integral Familiar Superbom, Pan Blanco Familiar Superbom, Pan Granos y Semillas.

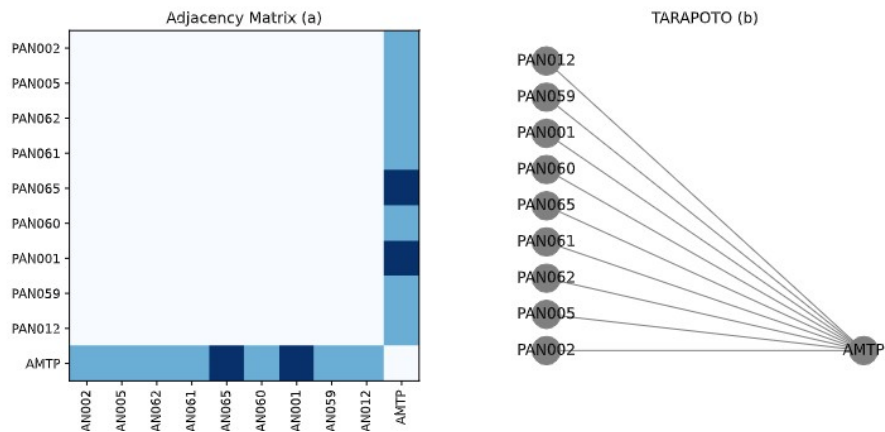


Figura 8. Matriz de adyacencia (a) y gráfico bipartito (b) para la ciudad de Tarapoto en 2023.  
Fuente: Elaboración propia

En el almacén de Tarapoto, se observa que el pan más recurrente es PAN001, PAN065, lo que representa al Pan Americano Sandwich, Pan Molde Blanco Guerrero.

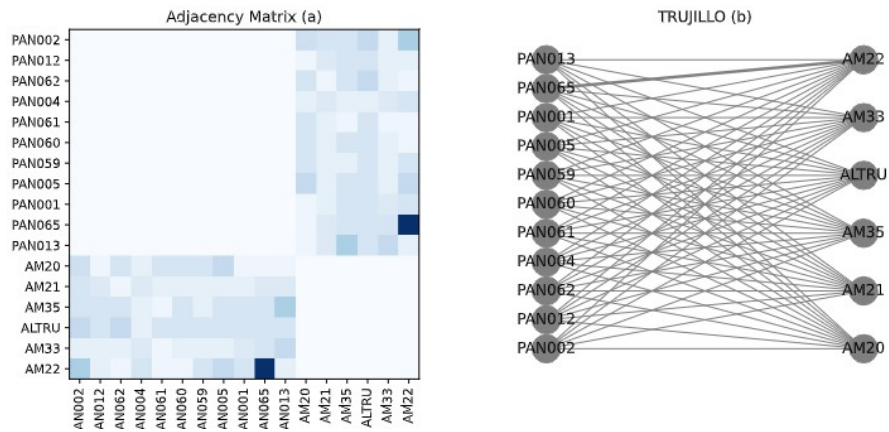


Figura 9. Matriz de adyacencia (a) y gráfico bipartito (b) para la ciudad de Trujillo en 2023.  
Fuente: Elaboración propia

Y en el almacén de Trujillo, se observa que los panes más vendidos son PAN065, PAN013, PAN002, lo que representa al Pan Molde Blanco Guerrero, Pan Integral Sandwich y el Pan Integral Familiar Superbom.

## Capítulo V

### Conclusiones y recomendaciones

#### 6.1. Conclusiones

En este estudio, desarrollamos un método para la toma de decisiones, basado en redes bipartitas que será de ayuda para pronosticar las ventas de productos en la línea de panes.

La construcción de la red bipartita nos permitió identificar patrones claros de venta de unidades de pan en los diferentes almacenes. Esto nos ayuda a comprender mejor qué tipos de pan tienen una demanda más alta en ciertos lugares.

Al analizar la red bipartita, pudimos identificar oportunidades para optimizar la distribución de los tipos de pan en los almacenes. Por ejemplo, podemos ajustar los niveles de inventario para satisfacer la demanda de manera más eficiente.

La visualización de la red bipartita nos proporcionó información valiosa sobre las relaciones entre los tipos de pan y los almacenes de venta. Esto nos ayuda a desarrollar estrategias de marketing más efectivas para promover ciertos tipos de pan en ubicaciones específicas.

La evaluación de la red bipartita también nos permitió ver las áreas donde podríamos mejorar nuestra estrategia de ventas. Por ejemplo, identificamos almacenes que podrían beneficiarse de una mayor diversificación de productos de pan para satisfacer mejor las preferencias de los clientes locales.

La implementación de la red bipartita no solo nos proporciona información valiosa en el presente, sino que también establece una base sólida para la toma de decisiones futuras en términos de gestión de inventario, planificación de la cadena de suministro y estrategia de marketing.

La implementación de la red bipartita bajo la metodología CRISP-DM ha demostrado ser una herramienta poderosa para comprender y optimizar la venta de unidades de pan en nuestros diversos almacenes de venta. Este enfoque nos ha permitido tomar decisiones más informadas y estratégicas para impulsar el éxito del negocio en el mercado de productos de pan.

## **6.2. Recomendaciones**

De acuerdo con los resultados obtenidos, estos pueden servir de base para realizar futuros trabajos sobre pronósticos de ventas en unidades en la línea de panes en industrias de alimentos en Perú, que permitirán ampliar los campos de aplicación del método usado y establecer mejores pronósticos.

Al seguir esta metodología CRISP-DM adaptada al contexto de la venta de unidades de pan en diversos almacenes, se puede desarrollar e implementar una red bipartita efectiva para optimizar las estrategias de ventas y marketing.

Refinar la estrategia de distribución, utilizando los patrones identificados en la red bipartita. Personalizar la oferta de productos por ubicación, lo que implicar la introducción de nuevos tipos de pan o la promoción de productos existentes que tienen una alta demanda en ciertos almacenes.

Utilizando la información sobre las relaciones entre los tipos de pan y los almacenes de venta, se pueden desarrollar campañas de marketing más efectivas y dirigidas.

Al seguir estas recomendaciones, se puede mejorar la eficiencia operativa y aumentar las ventas en los almacenes de venta de unidades de pan.

## Referencias bibliográficas

- [1] Aliaga, G. (2019). Calidad de servicio, satisfacción y Lealtad de los usuarios del servicio de consulta externa de las clínicas en el Perú (Universidad San Ignacio de Loyola). Retrieved from <http://repositorio.usil.edu.pe/handle/USIL/9386>
- [2] Islam, S.; Amin, S.H. Prediction of probable backorder scenarios in the supply chain using Distributed Random Forest and Gradient Boosting Machine learning techniques. *Journal of Big Data* 2020, 7, 1–22.
- [3] Ensafi, Y.; Amin, S.H.; Zhang, G.; Shah, B. Time-series forecasting of seasonal items sales using machine learning – A comparative analysis. *International Journal of Information Management Data Insights* 2022, 2, 100058. <https://doi.org/https://doi.org/10.1016/j.ijime.2022.100058>.
- [4] Cao, Q.; Liu, X.; Wang, Z.; Zhang, S.; Wu, J. Recommendation decision-making algorithm for sharing accommodation using probabilistic hesitant fuzzy sets and bipartite network projection. *Complex & Intelligent Systems* 2020, 6, 431–445. <https://doi.org/10.1007/s40747-020-00142-7>.
- [5] Harary, F. *Graph theory*. Addison-Wesley Series in Mathematics. Reading, Mass. etc.: Addison-Wesley Publishing Company. ix, 274 p. (1969)., 1969.
- [6] Kevork, S.; Kauermann, G. Bipartite exponential random graph models with nodal random effects. *Social Networks* 2022, 70, 90–99. <https://doi.org/https://doi.org/10.1016/j.socnet.2021.11.002>.
- [7] Punia, S.; Singh, S.P.; Madaan, J.K. From predictive to prescriptive analytics: A data-driven multi-item newsvendor model. *Decision Support Systems* 2020, 136, 113340. <https://doi.org/https://doi.org/10.1016/j.dss.2020.113340>.
- [8] Chopra, S.; Meindl, P.; Kalra, D.V. *Supply chain management: strategy, planning, and operation* (Vol. 232), 2013.
- [9] Acharya, A.; Singh, S.K.; Pereira, V.; Singh, P. Big data, knowledge co-creation and decision making in fashion industry. *International Journal of Information Management* 2018, 42, 90–101. <https://doi.org/https://doi.org/10.1016/j.ijinfomgt.2018.06.008>.
- [10] Wienclaw, R. *Forecasting Methods for Management*. Research Starters-Business 2015, pp. 1–6.

- [11] Thomassey, S. Sales forecasts in clothing industry: The key success factor of the supply chain management. *International Journal of Production Economics* 2010, 128, 470–483. *Supply Chain Forecasting Systems*, <https://doi.org/https://doi.org/10.1016/j.ijpe.2010.07.018>.
- [12] Valejo, A.D.B.; de Oliveira dos Santos, W.; Naldi, M.C.; Zhao, L. A review and comparative analysis of coarsening algorithms on bipartite networks. *The European Physical Journal Special Topics* 2021, 230, 2801–2811. Version January 25, 2024 submitted to *Journal Not Specified* 10 of 10
- [13] Aslan, S.; Kaya, M. Topic recommendation for authors as a link prediction problem. *Future Generation Computer Systems* 2018, 89, 249–264.
- [14] Zhou, T.; Ren, J.; Medo, M.; Zhang, Y.C. Bipartite network projection and personal recommendation. *Physical review E* 2007, 76, 046115.
- [15] Chen, X.; Xie, D.; Wang, L.; Zhao, Q.; You, Z.H.; Liu, H. BNPMDA: Bipartite Network Projection for MiRNA–Disease Association prediction. *Bioinformatics* 2018, 34, 3178–3186, <https://academic.oup.com/bioinformatics/articlepdf/34/18/3178/48920873/bioinformatic> <https://doi.org/10.1093/bioinformatics/bty333>.
- [16] Euler, L. (1736). *Solutio problematis ad geometriam situs pertinentis*. *Commentarii academiae scientiarum Petropolitanae*, 8, 128-140
- [17] König, D. (1936). *Theorie der endlichen und unendlichen Graphen*. Chelsea Publishing Company.
- [18] Hall, P. (1935). On representatives of subsets. *Journal of the London Mathematical Society*, 10(1), 26-30.
- [19] Harary, F. (1962). Theorem 1.5. En "Graph Theory" (p. 10). Addison-Wesley.
- [20] Adomavicius, G., & Tuzhilin, A. (2015). *Recommender systems for e-commerce*. Springer Science & Business Media.
- [21] Chen, X., Zhang, Y., & Wang, D. (2023). Product demand prediction in bipartite networks with deep learning. *IEEE Transactions on Knowledge and Data Engineering*, 35(4), 1014-1027. <https://ieeexplore.ieee.org/document/9312014>
- [22] Wang, J., He, X., & Li, J. (2024). Product recommendation in bipartite networks with attention to user heterogeneity. *IEEE Transactions on Knowledge and Data Engineering*. <https://ieeexplore.ieee.org/document/9728482>

- [23] Liu, H., Wang, J., & Li, X. (2022). Fraud detection in bipartite networks with machine learning. *IEEE Access*, 10, 33421-33434.
- [24] Zhao, X., Li, Y., & Chen, X. (2023). Competition analysis in bipartite networks with Markov graphs. *IEEE Transactions on Network Science and Engineering*. <https://ieeexplore.ieee.org/document/9352832>
- [25] Aslan, S.; Kaya, M. Predicting Links in Complex Bipartite Networks Based on Strengthening Projections. In *Proceedings of the 2018 International Conference on Artificial Intelligence and Data Processing (IDAP)*. IEEE, 2018, pp. 1–6.
- [26] Valderrama, S. *Pasos para elaborar proyectos de investigación científica*, 2015.
- [27] Cordero, Z.R.V. *La investigación aplicada: una forma de conocer las realidades con evidencia científica*. *Revista educación* 2009, 33, 155–165.
- [28] Hernández Sampieri, R.; Fernández Collado, C.; Baptista Lucio, P.; et al. *Metodología de la investigación*; Vol. 4, McGraw-Hill Interamericana México, 2018.
- [29] Wilson, R.L.; Sharda, R. Bankruptcy prediction using neural networks. *Decision Support Systems* 1994, 11, 545–557. [https://doi.org/https://doi.org/10.1016/01679236\(94\)90024-8](https://doi.org/https://doi.org/10.1016/01679236(94)90024-8).
- [30] Zhang, G.; Eddy Patuwo, B.; Y. Hu, M. Forecasting with artificial neural networks:: The state of the art. *International Journal of Forecasting* 1998, 14, 35–62. [https://doi.org/https://doi.org/10.1016/S0169-2070\(97\)00044-7](https://doi.org/https://doi.org/10.1016/S0169-2070(97)00044-7).
- [31] Gupta, A.; Shrinath, P. Link Prediction based on bipartite graph for recommendation system using optimized SVD++. *Procedia Computer Science* 2023, 218, 1353–1365. *International Conference on Machine Learning and Data Engineering*, <https://doi.org/https://doi.org/10.1016/j.procs.2023.01.114>.
- [32] Liew, C.; Labadin, J.; Kok, W.; Eze, M. A methodology framework for bipartite network modeling. *Applied Network Science* 2023, 8, 6. <https://doi.org/10.1007/s41109-023-00533-y>.
- [33] Ayele, W.Y. Adapting CRISP-DM for Idea Mining: A Data Mining Process for Generating Ideas Using a Textual Dataset. *International Journal of Advanced Computer Science and Applications* 2020, 11. <https://doi.org/10.14569/IJACSA.2020.0110603>.
- [34] Huber, S.; Wiemer, H.; Schneider, D.; Ihlenfeldt, S. DMME: Data mining methodology for engineering applications – a holistic extension to the CRISP-DM

- model. *Procedia CIRP* 2019, 79, 403–408. 12th CIRP Conference on Intelligent Computation in Manufacturing Engineering, 18-20 July 2018, Gulf of Naples, Italy, <https://doi.org/https://doi.org/10.1016/j.procir.2019.02.106>.
- [35] Chapman, P.; Clinton, J.; Kerber, R.; Khabaza, T.; Reinartz, T.P.; Shearer, C.; Wirth, R. *CRISP-DM 1.0: Step-by-step data mining guide*. 2000.
- [36] Wirth, R.; Hipp, J. *CRISP-DM: Towards a standard process model for data mining*. In *Proceedings of the Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*. Manchester, 2000, Vol. 1, pp. 29–39.
- [37] Chapman, P.; Clinton, J.; Kerber, R.; Khabaza, T.; Reinartz, T.; Shearer, C.; Wirth, R.; et al. *CRISP-DM 1.0: Step-by-step data mining guide*. SPSS inc 2000, 9, 1–73.
- [38] Schröer, C.; Kruse, F.; Gómez, J.M. A Systematic Literature Review on Applying CRISP-DM Process Model. *Procedia Computer Information Systems / ProjMAN 2020 International Conference on Project MANagement / HCist 2020 - International Conference on Health and Social Care Information Systems and Technologies 2020, CENTERIS/ProjMAN/HCist 2020*, <https://doi.org/https://doi.org/10.1016/j.procs.2021.01.199>.