UNVEILING PATTERNS IN THE NIGHT MARKET: A MACHINE LEARNING AND DEEP LEARNING APPROACH TO CUSTOMER ANALYSIS

THANDAR PHYO

MASTER OF SCIENCE
IN
DIGITAL TRANSFORMATION TECHNOLOGY

SCHOOL OF INFORMATION TECHNOLOGY
MAE FAH LUANG UNIVERSITY
2023

UNVEILING PATTERNS IN THE NIGHT MARKET: A MACHINE LEARNING AND DEEP LEARNING APPROACH TO CUSTOMER ANALYSIS

THANDAR PHYO

THIS THESIS IS A PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN DIGITAL TRANSFORMATION TECHNOLOGY

SCHOOL OF INFORMATION TECHNOLOGY
MAE FAH LUANG UNIVERSITY
2023

UNVEILING PATTERNS IN THE NIGHT MARKET: A MACHINE LEARNING AND DEEP LEARNING APPROACH TO CUSTOMER ANALYSIS

THANDAR PHYO

THIS THESIS HAS BEEN APPROVED
TO BE A PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF SCIENCE
IN
DIGITAL TRANSFORMATION TECHNOLOGY
2023

EXAMINATION COMMITTEE

....CHAIRPERSON

(Asst. Prof. Worasak Rueangsirarak, Ph. D.)

....ADVISOR

(Surapong Uttama, Ph. D.)

....EXTERNAL EXAMINER

(Assoc. Prof. Phasit Charoenkwan, Ph. D.)

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to all individuals who have directly or indirectly contributed to the successful completion of this thesis. I am especially thankful to Mae Fah Luang University, School of Information Technology, for permitting me to develop this thesis and for their invaluable ideas and suggestions.

I extend my heartfelt thanks to Dr. Surapong Uttama for his administrative supervision and for providing invaluable guidelines throughout the preparation of this thesis. I am also deeply grateful to Asst. Prof. Dr. Santichai Wicha for his lectures and support, and to Asst. Prof. Dr. Worasak Rueangsirarak for his guidance and support in navigating various topics in Digital Research Methodology. Additionally, I would like to thank Asst. Prof. Dr. Teeravisit Laohapensaeng, Dean of the School of Information Technology at Mae Fah Luang University, for providing a conducive environment for my research. Moreover, I am grateful to specify that this thesis received a thesis writing grant from Mae Fah Luang University.

Special thanks are due to my parents for their unwavering kindness and assistance. I also extend my respect and gratitude to all my teachers for their helpful recommendations, and to my seniors, friends, and the Mae Fah Luang University staff for their contributions towards the completion of this thesis.

Thandar Phyo

Thesis Title Unveiling Patterns in the Night Market: A Machine Learning and Deep Learning Approach to Customer Analysis

Author Thandar Phyo

Degree Master of Science (Digital Transformation Technology)

Advisor Surapong Uttama, Ph. D.

ABSTRACT

This research conducts an in-depth examination of consumer behavior and business strategy by thoroughly analyzing data from the night market at Mae Fah Luang University. The main goal is to categorize customers into distinct segments to reveal complex purchasing patterns and associations among bought items, thereby illuminating customer preferences and tendencies. Using machine learning techniques, specifically customer segmentation and market basket analysis, the study employs the K-means clustering to divide customers into five segments, with the optimal number determined through the elbow method and clustering validation metrics. Four association rules were derived between products using the Apriori algorithm, with the most robust rule indicating a relationship between street food and snacks and beverages for 50% of the customers. Furthermore, the study addresses the challenge of identifying food items packaged in plastic by implementing object detection techniques, achieving a mean Average Precision (mAP) of 99.4% for unpackaged items and 99.3% for packaged ones, with a combined model mAP of 84.4%. Among various algorithms such as FP-Growth and Eclat, the Apriori algorithm yielded the highest support value of 20% and a confidence level of 50%, resulting in eight association rules with an accuracy of 84% on a new dataset. These insights are crucial for businesses in the food industry, allowing them to tailor marketing strategies and product offerings to better meet the consumer demands, thereby enhancing the understanding of consumer behavior and product associations in the food sector.

Keywords: Night Market, Customer Segmentation, Market Basket Analysis, Machine Learning, Deep Learning, CNN, YOLOv8

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

The night market, also called a night bazaar, is a street market that operates in the evening and offers a more relaxed atmosphere for strolls, shopping, and dining compared to the more business-oriented day markets (Mazlan et al., 2017). The night market's popularity grew due to its convenient location, providing locals with a nearby option for shopping. Its friendly, relaxed atmosphere offered a diverse selection of freshly cooked food and vegetables at reasonable prices. Interactions among residents of different backgrounds fostered community bonds. The overall ambiance and the variety of offerings made the night market a favored leisure spot for locals (Ishak et al., 2012).

People prefer night markets due to the cooler weather, vibrant atmosphere with colorful lights, extensive variety of street food, entertainment options like live music, and the adventurous nighttime setting, making it a memorable experience. With 1256 night markets spread across Asia, including 34 in Myanmar, 194 in Taiwan, 200 in India, 214 in Malaysia, and 237 in Thailand, these nocturnal bazaars attract millions of visitors annually ("List of night markets in Asia," 2024). For instance, the Seoul Bamdokkaebi Night Market in South Korea alone welcomed approximately 221 million visitors from 2015 to 2017, highlighting the immense popularity of these cultural hubs ("Seoul night market visitor number South Korea 2015-2017", 2022). People prefer night markets over day markets since they provide a distinctive and alluring blend of cooler weather, atmosphere, food diversity, and entertainment.

Figure 1.1 A Vibrant Night Market Scene at Mae Fah Luang University

Figure 1.2 Night Market Food Culture

This research explores the impact of night markets on local economies, especially focusing on the night market held at Mae Fah Luang University every

Wednesday and Friday. Mae Fah Luang University is situated in Chiang Rai, Thailand. The night market at Mae Fah Luang University spans the main campus square, showcasing a variety of stalls and vendors offering a range of products and services (see Figure 1.1 and Figure 1.2). The market features a mix of local artisans and food vendors, creating a lively and diverse atmosphere. The layout includes designated areas for food stalls, craft vendors, and entertainment stages, with pathways for visitors to navigate the market easily.

The night market offers a wide range of products, especially Thai cuisine, including meals, snacks, fruits, drinks, beverages, and desserts. The food offerings at this market are diverse and cater to different tastes and dietary preferences. In addition to oven-baked chicken and rice and egg fried rice, the market offers various noodles, providing options for noodle lovers. For those seeking healthier options, there are Vietnamese spring rolls, as well as boiled sweet potatoes and boiled pumpkins. These dishes are complemented by a selection of nuts, including sunflower seeds, almond nuts, and pumpkin seeds, adding a crunchy and nutritious element to the market's food offerings.

The market also offers a diverse range of products, including a variety of fruits such as sliced apples, watermelons, strawberries, melons, and pineapples, along with beverages like coconut water, smoothies, orange juice, and cold-pressed juices such as pomegranate, mulberry, and pineapple. Additionally, the dessert category features banana waffles, chocolate waffles, brownies, and various types of Korean fish-shaped pastries. The market also offers clothing, accessories, fresh vegetables, and a variety of stalls, making it a one-stop destination for both food and non-food shopping.

This diverse array of offerings, coupled with its bustling atmosphere and high footfall, highlights the market’s significance as a key economic and social hub within the university community. On average, the night market attracts around 500 customers every 20 minutes, with this number increasing to as many as 1,000 customers during peak hours from 5:30 pm to 7:00 pm. With 5,000 to 10,000 customers visiting every week, the market's popularity and variety of offerings underscore its importance in stimulating economic growth and fostering community spirit.

1.2 Problem Identification

The inaugural stage of this comprehensive study delves into the critical process of pinpointing a significant research issue that lies at the heart of the investigation. This issue centers on the intriguing and complex realm of shopping behaviors, particularly within the distinct demographic segments of residents, as they navigate the myriad offerings within their respective stores.

The first step in this research process involves identifying the research issues, which are not understanding and categorizing customers into meaningful segments and lacking an understanding of consumers' shopping behaviors at the night market. The second step is analyzing data from night markets, especially for market basket analysis, which presents various challenges. In a previous study on market basket analysis, (Phyo & Uttama, 2023), images of products were collected from customers at this night market. However, these images sometimes showed food items and products in plastic bags or boxes (see Figure 1.3), making it difficult to identify the items. In other cases, the food items were not in plastic bags, which facilitated easier identification (see Figure 1.4).

Figure 1.3 An Example of Night Market Food Wrapped in Plastic Bags

Figure 1.4 An Example of Night Market Food without the Plastic Bag

The focal point of this phase is to illuminate the pervasive lack of comprehensive comprehension regarding the intricacies and dynamics that underlie the decision-making processes, preferences, and tendencies exhibited by these distinct groups of consumers. By delving into the intricacies of how residents engage with and navigate the shopping environment, the research aims to unravel the enigmatic web of factors, motivations, and influences that shape their choices, paving the way for a deeper understanding of the underlying mechanisms driving their interactions with the stores.

1.3 Research Objective

The objective of this research is to use technology to achieve the following:

1.3.1 To collect data from the Mae Fah Luang University's night market and implement data preprocessing approaches

1.3.2 To utilize machine learning techniques to cluster the data into respective groups, thereby identifying prevalent purchasing patterns among students and understanding their preferred product combinations

1.3.3 To conduct object detection of food items using deep learning techniques for product object detection from the night market

1.3.4 To feed the expected results to market basket analysis approaches to find associations between food items

1.4 The Importance of Research

The research outlined here plays a pivotal role in bridging a critical gap in understanding shopping behaviors at night markets by focusing on the often-overlooked perspective of consumers. By investigating both consumer and resident viewpoints within the night market context, this research aims to shed light on the intricate dynamics influencing their choices and interactions within stores. Customer segmentation, market basket analysis, and deep learning techniques for food object detection, employed as key methodologies, offer a comprehensive lens to comprehend the nuanced shopping behaviors and preferences. By diving into these complexities, the study seeks to unravel the multitude of factors, motivations, and influences guiding consumers and residents in their interactions within the night market. Using machine learning and deep learning techniques to analyze numerical data representing various facets of customer behaviors and patterns is a significant step forward. It allows for the identification of prevalent purchasing behaviors, the clustering of customers into distinct groups, and an understanding of preferred product combinations. This quantitative approach provides empirical evidence and tangible insights into consumer behaviors, contributing significantly to our understanding of how individuals engage within the night market setting. The importance of this research lies in its potential to serve as a foundational piece, paving the way for more comprehensive and insightful analyses. It offers a deeper understanding of customer behavior within the night market and establishes a framework for future studies. Ultimately, this research facilitates a more nuanced comprehension of consumer interactions with stores in this unique context, enabling better-informed strategies for businesses and policymakers alike.

1.5 Research Questions

This study aims to answer these research questions as follows:

1.5.1 How can a comprehensive methodology be formulated to categorize customers into meaningful segments within the night market context, aiming to refine targeted marketing strategies and personalize consumer experiences?

1.5.2 What strategies or methodologies could be employed to deepen the understanding of the consumer's shopping behaviors specific to night market environments, facilitating more effective tailoring of offerings and experiences to accommodate diverse consumer preferences and needs?

1.6 Scope of Study

This thesis collects and analyzes two distinct datasets to understand consumer behavior and product preferences at Mae Fah Luang University's night market. The first dataset contains consumer information gathered from the night market and includes nine features: Age, Gender, Occupation, Initial Budget, Purchased Products, Spending Amount, Number of Visited Times, Arrival Hour, and Time Spent at the night market. This dataset comprises over 100 records of consumer information. Due to the limited amount of data, various machine learning algorithms will be applied to enhance the performance and accuracy of the models. This part of the research focuses on the impact of the night market, machine learning, customer segmentation, and market basket analysis.

The second dataset consists of 1,000 annotated images, each accompanied by its respective annotation file. Half of the dataset, totaling 500 images, contains multi-dish compositions, while the remaining 500 are single-dish images. This combined approach, utilizing both the consumer transaction data and annotated food images, aims to provide a comprehensive understanding of consumer behaviors and product combinations at the night market by using deep learning techniques.

CHAPTER 2

LITERATURE REVIEW

2.1 Theoretical Background

This portion examines the existing body of literature concerning concepts and theories associated with customer analysis topics including customer segmentation, market basket analysis, machine learning, and deep learning techniques.

2.1.1 Customer Segmentation

Customer segmentation is a critical process that involves classifying markets and pinpointing groups of customers with similar needs and expectations. This methodology enables companies to develop tailored strategies for various customer segments, adapt to evolving trends, and project future growth trajectories. Significantly, it bolsters a company’s market position and empowers customers to influence strategic decisions, thereby enhancing the company's market value (Singh et al., 2023). In this research, the k-means clustering algorithm has been employed for customer segmentation.

2.1.2 K-Means Clustering

This algorithm is grounded in the concept of partitioning and utilizes the elbow method to ascertain the optimal number of clusters, denoted as ‘K’. It employs the Sum of Square Errors (SSE) metric to evaluate the proximity of centroids to the given data points, potentially testing various K values. The objective is to identify the most suitable K value where the SSE is minimized. Additionally, the Euclidean distances between data points are calculated to assign them to the nearest centroid, thus forming clusters. This process iteratively computes the cluster centroids using the mean of the cluster’s data points, continuing until the centroids' positions stabilize and no further changes occur (Mehta et al., 2021). The performance of the K-means algorithm can be evaluated by validation metrics such as Silhouette, Davies Bouldin, and Calinski-Harabasz.

2.1.3 Data Mining

Data mining is the process of examining vast datasets to uncover patterns and extract new insights. It is extensively used in Artificial Intelligence (AI) and Data Science (Sharma & Babbar, 2023). Data mining techniques are designed for different types of data, each with unique characteristics. The four main types include structured, unstructured, semi-structured, and multi-dimensional data. Several prominent data mining methods include clustering, classification, association rule mining, and regression analysis. Noteworthy applications of these methods span across business intelligence, healthcare, finance and banking, and manufacturing (El Ghezzaz, 2023).

2.1.4 Market Basket Analysis

This research focuses on market basket analysis, a specific application of data mining that examines patterns in consumer purchases to uncover relationships between products. It is a method used by businesses to discover which items are often bought together. For example, baby formula and diapers are typically placed in the same aisle in grocery stores. This technique uncovers hidden patterns in customer purchases, helping businesses optimize product placement for increased sales. It is an unsupervised machine learning method that doesn't involve training on a dataset or predicting outcomes. Instead, it identifies relationships between items and combinations of items that are commonly purchased together (Selvaraj, 2023).

Market Basket Analysis often uses association rule learning to uncover these relationships. Association rule learning, an unsupervised learning technique, identifies dependencies between data items to enhance profitability. It discovers interesting relations among dataset variables based on rules. For instance, if a customer buys bread, they are likely to purchase butter, eggs, or milk, prompting these products to be placed together on shelves. Association rule learning includes three main types of algorithms: Apriori, Eclat (Equivalence Class Clustering and Bottom-Up Lattice Traversal), and FP-Growth (Frequent Pattern Growth) (“Association rule learning,” 2024).

2.1.5 Metrics for Market Basket Analysis

In association rule learning, various metrics are employed to assess the quality and significance of discovered association rules. Support quantifies the frequency of an item or itemset in the dataset, indicating its commonality or rarity. Confidence measures the strength of association between two items, indicating how often one item's presence predicts another's presence. Lift compares the strength of association between two items to the expected association if independent, with values above 1 indicating stronger than expected association (Ali, 2023).

2.1.6 Object Detection

Machine learning and deep learning are subsets of Artificial Intelligence (AI). Machine learning involves creating algorithms and models that allow computers to learn and make decisions without explicit programming. Deep learning, a subset of machine learning, utilizes multi-layered artificial neural networks to understand intricate patterns in data. It emphasizes hierarchical representations and automatic feature extraction, with key features including neural networks, unsupervised and semi-supervised learning, end-to-end learning, and high computational demands. Deep learning excels in image and object recognition, Natural Language Processing (NLP), speech synthesis and recognition, and autonomous driving tasks (Yilmaz, 2023).

In this research, the focus is on using deep learning for object detection, specifically to detect food items and products in customer data from the Mae Fah Luang University’s night market. Object detection is a crucial challenge in computer vision, involving identifying and localizing all objects within an image. It combines both object localization and classification for multiple objects in an image. Traditional object detection methods like Viola-Jones, Histogram of Oriented Gradients (HOG), and Deformable Part Models (DPM) were used before deep learning, but now deep learning methods like Two-Stage and One-Stage Detectors are preferred due to their effectiveness (Pandey et al., 2022).

2.1.7 You Only Look Once (YOLO)

YOLO is an advanced object detection algorithm created in 2015 by Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. Their research paper, “You Only Look Once: Unified, Real-Time Object Detection,” introduced this groundbreaking approach to real-time object detection. The authors approach object detection as a regression problem, not a classification task. They use a single convolutional neural network (CNN) to spatially separate bounding boxes and assign probabilities to detected objects (Keita, 2022).

YOLOv8 is the current state-of-the-art object detection algorithm built on a deep CNN architecture. It introduces enhancements like the Cross Stage Partial Network (CSPNet) backbone for efficiency and accuracy, Feature Pyramid Network (FPN) and Path Aggregation Network (PAN) neck architecture for improved feature aggregation, PAN head architecture for robustness, and a unique training approach that blends supervised and unsupervised learning. YOLOv8 divides the input image into a grid of cells and predicts bounding boxes and class probabilities for each cell. It then uses Non-Maximum Suppression (NMS) to select the most likely bounding boxes for each object, filtering out overlaps (Muhammadshabrozshahab, 2023).

2.1.8 Metrics for Object Detection

Object detection performance is typically evaluated using metrics such as Average Precision (AP), which is the average precision at different recall levels, and mean Average Precision (mAP), which is the mean of AP across different classes in the dataset. Object localization accuracy is measured using the Intersection over Union (IoU) metric, calculated between the ground truth and predicted bounding boxes. If the IoU exceeds a predefined threshold, usually 0.5, the object is considered successfully detected (Pandey et al., 2022).

Through a comprehensive review of these topics, this literature review aimed to provide a solid foundation for understanding state-of-the-art methodologies and their practical applications in customer analysis. By exploring the evolution of these techniques, their theoretical underpinnings, and real-world case studies, this section offered valuable insights into how businesses can leverage these advanced analytical tools to enhance their customer relationship management and overall operational efficiency.

2.2 Related Works

This section reviews studies relevant to this research on night markets, customer segmentation, market basket analysis, and object detection techniques. In this research, the K-means clustering algorithm is employed to discern customer behaviors and patterns for segmentation purposes. Mehta et al. (2021) compared various clustering techniques and found that the K-means algorithm yielded the best Silhouette and Davies-Bouldin scores, concluding its superior efficacy in customer segmentation. For this study, clustering models were developed using customer profile data, particularly their use of Internet banking, to segment customers at XYZ Bank. Both K-means and K-medoid methods were utilized, with K-means exhibiting better performance in terms of intra-cluster distance (Aryuni et al., 2018). Additionally, Gupta et al. (2021) applied the K-means algorithm to group restaurants based on geographic location, successfully dividing them into six clusters, thereby achieving the primary goal of geographic segmentation.

The Taishan Night Market was the subject of this paper (Li et al., 2021), which proposed strategies for its transformation to enhance economic, cultural, and social impacts. The study highlighted the importance of the night market in economic development and residents' quality of life. Key challenges such as poor food hygiene and low economic impact were identified, highlighting the need for urgent transformation. The methodology involved a comparative analysis with Chengdu to derive transformation suggestions. Proposed solutions focused on strengthening management, infrastructure, and cultural aspects, as well as promoting the sharing economy. Bamrongpol et al. (2020) analyzed the influence of visitor satisfaction, service quality, market reputation, and visitor trust on revisit intention in urban night markets using Structural Equation Modeling (SEM). It emphasized the importance of catering to local visitors for market sustainability, highlighting the interrelation of factors and the significance of service quality, market reputation, and visitor trust. Limitations included rejected hypotheses due to low correlation values and potential questionnaire limitations for single visitors.

Association rule learning is utilized in market basket analysis to uncover patterns in customer purchasing behavior, and numerous researchers have focused on tackling the challenges within this area of study. Berlilana et al. (2022) used a quantitative approach to analyze consumer shopping patterns in supermarkets. Transaction data was processed using a hash-based algorithm and KNIME software for association rule learning. Only relevant supermarket transaction data was included, and a confidence level of 90% was used for generating association rules. In another study, Aldino et al. (2021) analyzed transaction data using data mining algorithms such as Apriori and FP-Growth to identify consumer buying patterns. FP-Growth exhibited superior performance compared to Apriori, achieving a higher accuracy of 217% and faster processing time of 6 seconds, as opposed to Apriori's accuracy of 46% and processing time of 30 seconds.

Numerous researchers have tackled the challenges of food image recognition, exploring classification and using various techniques in their work. Tan et al. (2023) conducted a comparative study of deep learning models for food detection and recognition, examining the impact of model architecture and input size on performance across various datasets. YOLOv5 demonstrated superior performance on the UEC Food 100 and UEC FOOD 256 datasets, while Faster Region-Convolutional Neural Network (Faster R-CNN) performed better on the School Lunch dataset. The study underscored the critical role of model architecture and input size in enhancing the performance of food detection systems. The methodology involved a comparative analysis of six deep learning models across three publicly available datasets, evaluating performance using the mAP metric.

Pandey et al. (2022) introduced the IndianFood10 dataset, comprising 10 traditional Indian food classes, and achieved a high mAP score of 91.8% using YOLOv4. Additionally, it extended the dataset to IndianFood20, emphasizing the importance of object detection in traditional Indian cuisine. The main findings included the creation of the IndianFood10 dataset with 10 traditional Indian food classes, achieving a high mAP score of 91.8%, providing an extension dataset IndianFood20, and highlighting the significance of object detection in traditional Indian cuisine. The methodology involved creating two datasets of traditional Indian food items (IndianFood10 and IndianFood20), training a YOLOv4 object detection model on 80%

of the IndianFood10 dataset, and evaluating the model using Precision-Recall curves and AP.

In summary, previous studies have overlooked several limitations in data collection, including issues with data quality, especially concerning products in plastic bags. Conventional approaches to market basket analysis often rely on data from supermarkets or traditional markets, overlooking the unique challenges of night markets where food items are typically packaged in plastic bags. To address these limitations, this study proposes using advanced data collection techniques, such as utilizing state-of-the-art object detection models. For instance, instead of time-consuming interviews with customers that can take up to 15 minutes per person to ascertain their purchases, this approach offers a more efficient alternative, reducing the effort required for market basket analysis in night market environments or various consumer-oriented industries.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Overall Methodology

Figure 3.1 shows the overall methodology of this study. This research aims to develop a comprehensive approach to customer segmentation and market basket analysis through the application of machine learning and object detection techniques. The methodology involves a series of well-defined steps, starting from data collection and preprocessing to model training and evaluation. Initially, data was collected in two primary forms: customer information and product images. The customer information underwent extensive preprocessing, including data merging, cleaning, selection, exploratory data analysis (EDA), and feature scaling. Concurrently, the product images were annotated, split, preprocessed, and augmented to prepare them for object detection modeling. This dual approach ensured that both datasets were robust and ready for detailed analysis.

The research was conducted through two sequential experiments. In the first experiment, manual data collection for market basket analysis proved to be time-consuming, highlighting a significant limitation in the process. Additionally, analyzing data from night markets, particularly for market basket analysis, poses significant challenges. The first experiment collected images of products from customers at these markets, yet these images often depicted food items and products in plastic bags or boxes, complicating identification. To address this, the second experiment was conducted using deep learning object detection techniques to automatically label food items, allowing for a comparison of the results between the first and second experiments. This automated approach allowed for quicker and more precise data labeling, overcoming the limitations observed in the first experiment. The preprocessed customer information and the automatically detected items were then integrated into a machine-learning pipeline for customer segmentation using the K-means clustering algorithm. Simultaneously, market basket analysis was performed using the Apriori, FP-Growth, and Eclat algorithms to uncover patterns and associations within the transactional data. The final evaluation of the models and algorithms confirmed their effectiveness, providing valuable insights into customer behavior and product associations. This integrated methodology demonstrates the potential of combining machine learning and deep learning techniques to streamline complex data analysis processes in market research. Figure 3.1 demonstrates the overall research methodology.

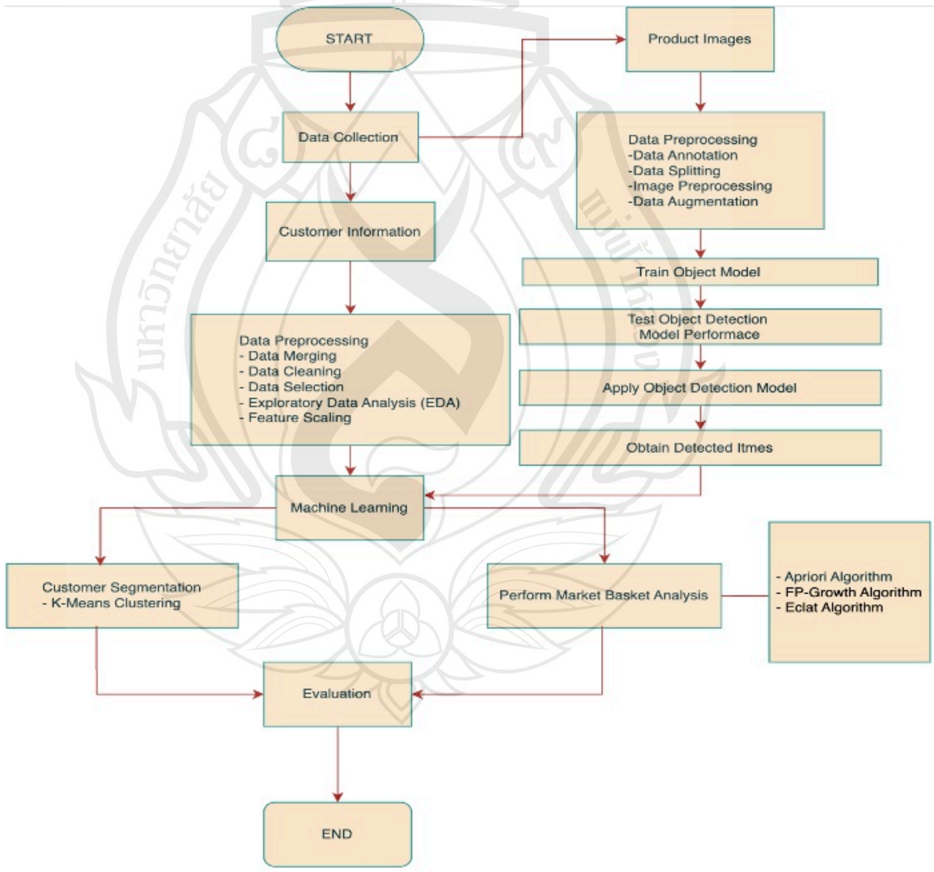

Figure 3.1 Overall Research Methodology

3.2 Data Collection for the First Experiment

The data was gathered from the night market at Mae Fah Luang University. The night market usually operates on Wednesday and Friday in a week. The sellers sell a variety of food products ranging from street food and snacks, beverages, clothing and accessories, handicrafts and souvenirs, electronics and gadgets, toys and novelty items, home décor, beauty and skincare products, art and craftwork, plants and flowers, and local products. A few records of the dataset along with their column names are shown in Table 3.1.

Table 3.1 The Sample Dataset of the Night Market Data

Age	Gender	Occupation	Initial_Budget	Purchased_Items	Spending_ Amount
19	2	1	200	Street food and snacks, Beverages	120
19	1	1	250	Street food and snacks, Fruits	100
21	1	1	100	Street food and snacks, Desserts	90
35	1	2	500	Street food and snacks, Beverages, Fruits	150

3.3 Data Preprocessing for the First Experiment

3.3.1 Data Merging

The research was conducted within a specific timeframe. Due to time constraints, data was collected using Google Forms and personal interviews, aiming to obtain both sets of results simultaneously. The first dataset gathered from a Google Form questionnaire, consisted of 43 responses. This data was initially in range format and was converted to a single representative value by calculating the midpoint of each range. The second dataset was collected through interviews conducted by a team of researchers and was converted into a Comma-Separated Values (CSV) file format. These two datasets were then merged for further analysis. Table 3.2 presents the features or labels of the consumer dataset. Numerical labels were assigned to the features to enhance visualization and usability.

Table 3.2 The Features of the Consumer Dataset

Features	Table Column Head
Age (1)	Age of the consumers.
Gender (2)	Gender of the consumers.
Occupation (3)	Occupation of the consumers.
Initial_Budget (4)	The amount of money allocated at the beginning.
Purchased_Items (5)	All items bought by the consumers.
Spending_Amount (6)	Total amount of items bought by the consumers.
Visit_Frequency (7)	How often do the consumers visit the night market?
Arrival_Hour_PM (8)	The specific time at which the consumers arrive.
Time_Spent (9)	The duration that customers dedicate to the night market.

3.3.2 Data Cleaning

At the same time, a data-cleaning process was implemented, as illustrated in Table 3.3.

Table 3.3 Data Cleaning

Problem	**Solution**
There were some outliers in the 'Spending Amount' Feature	The outliers were eliminated using the Interquartile Range (IQR) method. The original dataset had 102 rows but after removal, it was reduced to 95 rows and 9 features.

3.3.3 Data Selection

The features "Purchased_Items" and "Arrival_Hour_PM" were excluded from the clustering process as they were designated for market basket analysis. Consequently, these two columns were omitted for customer segmentation. Conversely, for association rules learning, only "Purchased_Items" and "Arrival_Hour_PM" were selected, as the remaining features were not pertinent to market basket analysis.

3.3.4 Exploratory Data Analysis (EDA)

This section delves into the EDA component. Table 3.4 showcases the fundamental descriptive statistics of the dataset, providing a summary of its key characteristics and properties. A comprehensive examination of the dataset revealed the absence of missing values. Additionally, a correlation analysis was conducted, as depicted in Figure 3.2. The correlation values calculated for all pairs of features were found to be very low, suggesting minimal to no linear correlation between the features.

Table 3.4 Basic Descriptive Statistics

	count	mean	std	min	25%	50%	75%	max
Age	95	22.02	4.294	18	19	21	23	43
Gender	95	1.75	0.460	1	1	2	2	3
Occupation	95	1.11	0.341	1	1	1	1	3
Initial_Budget	95	140.11	72.387	50	100	120	150	500
Spending_Amount	95	118.21	58.955	10	75	105	150	250
Visit_Frequency	95	1.63	0.484	1	1	2	2	2
Time_Spent_Minutes	95	30.73	15.280	5	20	30	30	90

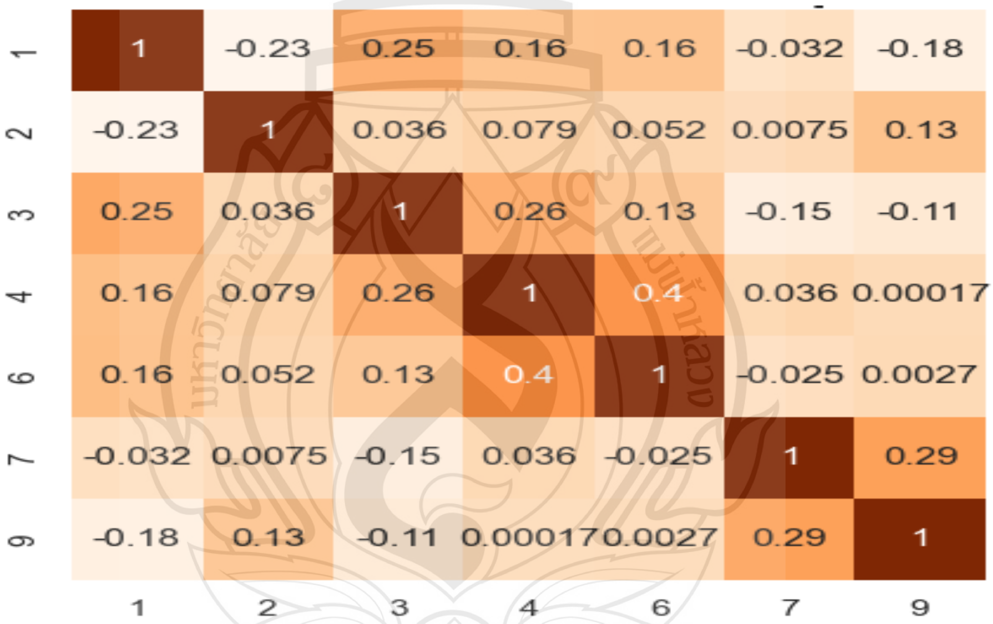

Figure 3.2 Correlation Heatmap of Features

3.3.5 Feature Scaling

Feature scaling is an essential step in data preprocessing designed to align all numerical features of a dataset to a common scale or range. Consequently, standardization also referred to as z-score normalization, was applied to all the features.

3.4 Machine Learning for the First Experiment

3.4.1 K-Means Clustering

The research utilized the K-means clustering algorithm to segment consumers based on their features from the collected dataset. The K-means clustering algorithm was employed for customer segmentation using the dataset's features.

3.4.2 Apriori Algorithm

Additionally, the Apriori algorithm was used to uncover associations and identify frequently purchased items by consumers, focusing on the features "Purchased_Items" and "Arrival_Hour_PM" from the dataset. The three key measures applied were: Support at 0.01%, Confidence at 0.5%, and a minimum Lift threshold of 1.3 for rule selection and evaluation. A low support threshold was used due to the limited amount of data.

3.5 Data Preparation for the Second Experiment

3.5.1 Data Collection

Following the survey conducted at the night market, we meticulously prepared the dataset for training the object detection model. We identified the most popular items sold both with and without plastic bags: apples, baked chicken, boiled corn, chicken steak, chicken rice, coconut water, egg tarts, egg fried rice, grilled eggs, mulberry juice, oranges, pumpkin seeds, sausage sandwich, and sweet crepes for items sold with plastic bags, and bananas, chicken steak, cookies, egg tarts, grilled eggs, kimchi soup, mulberry juice, orange juice, oranges, and rice berry for items sold without plastic bags. Subsequently, we collected a comprehensive array of food and product images corresponding to each item from market patrons. This meticulously curated dataset serves as the groundwork for training our object detection model, facilitating precise identification and classification of food items in subsequent analyses.

3.5.2 Dataset

The dataset comprises 1,000 annotated images, each accompanied by its respective annotation file. Half of the dataset, totaling 500 images, contains multi-dish compositions, while the remaining 500 are single-dish images. Single-dish images are annotated with the respective food item of interest, while multi-dish images are annotated to capture multiple unique food classes present within the image. Additionally, there are 500 images depicting food items sold with plastic bags and 500 images without plastic bags. The dataset encompasses 14 distinct classes for food items sold with plastic bags and 10 classes for those without plastic bags, resulting in a total of 24 food classes.

Due to the diverse nature of night market food, each dish often pairs with multiple others. For example, rice might accompany chicken or vegetables, showcasing the variety of culinary combinations available. For market basket analysis, it's observed that individuals often purchase meals along with beverages, fruits, or desserts, and occasionally opt for combinations of all these elements. Consequently, our dataset encompasses numerous images showcasing these combinations, along with corresponding annotation files for both food items sold with and without plastic bags.

3.5.3 Annotation

We utilized the LabelImg software (Tzutalin, n.d.), an open-source software to annotate the images in our dataset. Each raw input image was uploaded, and we manually created bounding boxes around each food item of interest, labeling each box with its corresponding food class. Subsequently, for every image in the dataset, a text file was produced in YOLOv8 format, containing information regarding the coordinates of the bounding boxes delineating each food item and the assigned food class number. Figure 3.3 showcases annotated night market food items such as meals, fruits, beverages, and desserts.

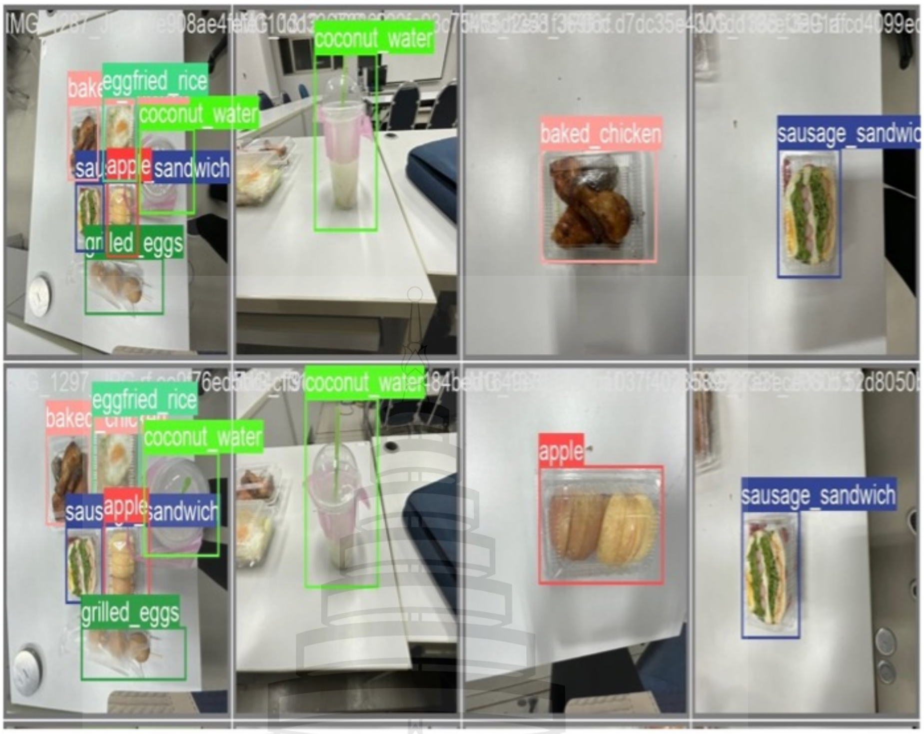

Figure 3.3 An Example of an Annotated Image of Food, Fruits and Beverages

3.5.4 Approach

In the approach section, we employed three distinct object detection models: one tailored for detecting food items sold with plastic bags, another for those sold without, and a third model capable of detecting both types of food items, regardless of packaging. The dataset was partitioned into three subsets for training, validation, and testing purposes. Specifically, 80% of the dataset was allocated to the training set, while 13% was set aside for validation, and the remaining 7% was designated as the testing set. Before training, image preprocessing techniques were applied. This included auto-orienting the images, resizing them to a standard size of 640 x 640 pixels, and employing data augmentation methods such as horizontal flipping to enhance the diversity of the training data. Post-training, we conducted a comprehensive comparative evaluation of all models, assessing their performance and effectiveness in detecting food items with and without plastic bags, respectively.

After obtaining the results from the object detection models, we utilized the outputs to perform market basket analysis. This phase involved scrutinizing the detected food items to discern patterns and associations in consumer purchasing behavior. To accomplish this, we employed three association rule learning algorithms: Apriori, Eclat, and FP-Growth. By examining the co-occurrence of various food items in the detected scenes using these algorithms, we aimed to extract meaningful insights into the combinations of items frequently purchased together by patrons at the night market.

CHAPTER 4

RESEARCH RESULT

4.1 Research Outcome for the First Experiment

Based on the research conducted, the K-means clustering algorithm separated the clusters into five sections according to the features based on the elbow method as described in Figure 4.1. The clustering validation metrics such as Silhouette, Davies Bouldin, and Calinski-Harabasz were also considered when choosing the ideal value of K. The Silhouette score ranges from -1 to 1, with a higher value indicating better-defined and more separate clusters while Davies Bouldin metric does not have a specific predefined range although a lower value indicates a better clustering. The Calinski-Harabasz index also does not have a specific predefined range with a higher value indicating better clustering. Figure 4.2 explores the relationship between the numbers of K and the Silhouette Coefficient.

Eventually, Principal Component Analysis (PCA) was employed to reduce the feature dimensionality to two for visualization, allowing for a clear representation of both the two features and the number of clusters as described in Figure 4.3. Upon examination of Figure 4.3, it becomes evident that clusters 3 and 4 are distinctly separated, whereas there is less clear separation between clusters 0, 1, and 2.

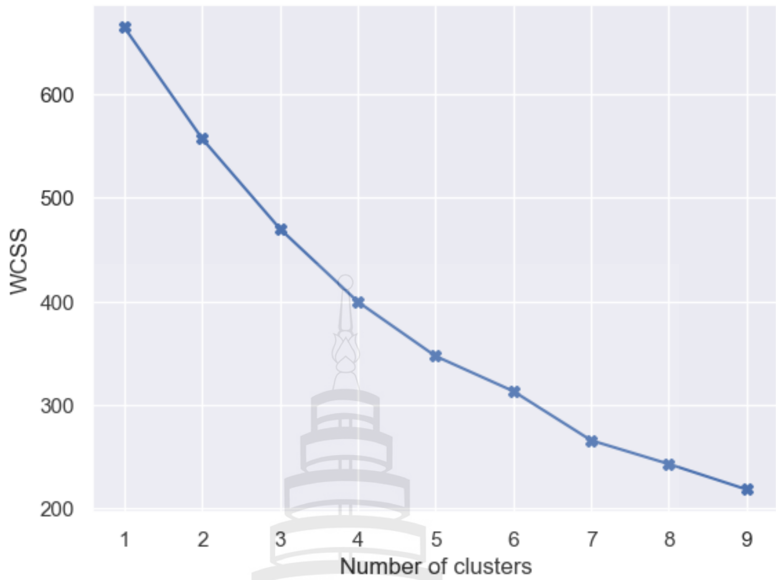

Figure 4.1 Within Cluster Sum of Squares and Number of Clusters

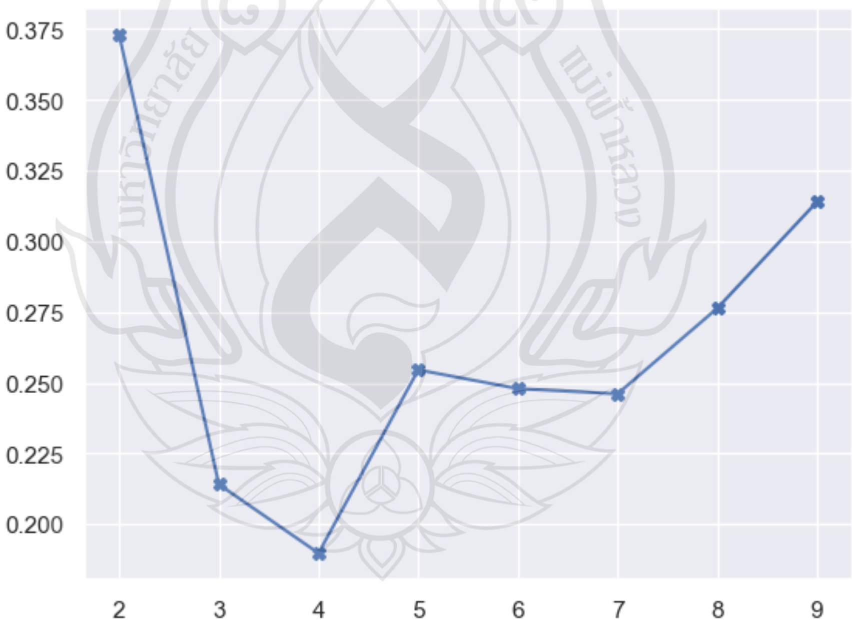

Figure 4.2 Silhouette Score and Number of Clusters (K)

Figure 4.3 Number of Clusters

Table 4.1 Clusters Interpretation

Clusters	Male	Female	Avg_Initial Budget	Min_Spent	Max_Spent	Avg_Spent
Cluster 0	9	34	141	30	200	104
Cluster 1	0	2	125	80	110	95
Cluster 2	1	6	117	40	110	58
Cluster 3	9	11	142	55	250	141
Cluster 4	6	16	146	75	250	143

Table 4.1 outlines the characteristics of five clusters. Cluster 0 has the highest proportion of females, with spending ranging from 30 to 200 Thai Bahts. Cluster 1 has no male members and the second lowest average spending. Cluster 2 has one male member and the lowest average expenditures. Clusters 3 and 4 have the highest average initial budget, maximum, and average spending. The experiment suggests that organizers prioritize attracting Clusters 3 and 4 to the night market for higher spending potential.

Using the Apriori Algorithm, four association rules were generated, each accompanied by its respective support, confidence, and lift metrics. Table 4.2 explains these rules, with the antecedent predicting the purchase of another item and the consequent being the item predicted to be purchased.

Table 4.2 Association Rules Interpretation

Antecedent	**Consequent**	**Support**	**Confidence**	**Lift**
Plants and flowers	Fruits	0.017	1.000	3.471
Street food and snacks	Beverages	0.017	1.000	2.185
Desserts	Beverages	0.068	0.667	1.457
Fruits	Beverages	0.084	0.625	1.366

Table 4.3 Association Rules Prediction

Antecedent	**Consequent**	**Correct**	**Incorrect**
Plants and flowers	Fruits	0	20
Street food and snacks	Beverages	10	5
Desserts	Beverages	6	6
Fruits	Beverages	4	7

The results from Table 4.3 indicate that the rules are effective in predicting and revealing the characteristics of features. The second rule demonstrates strong predictive power, with 10 people buying street food snacks and beverages together out of 20 observations, while 5 people bought only street food and snacks, excluding beverages. However, the first rule was not as effective due to customers not purchasing plants and flowers with fruits together. The third and fourth rules are the second and third most reliable, with 6 customers buying desserts and beverages together, while 6 bought only desserts without purchasing beverages. Additionally, 4 consumers out of 20 observations bought fruits and beverages together, while 7 customers bought only fruits without beverages. Since street food and snacks, as well as beverages, are frequently purchased items, it would be beneficial to place shops or vendors offering these products close to each other.

4.2 Research Outcome for the Second Experiment

We evaluated the performance of three object detection models trained to identify food items with and without plastic bags, employing standard evaluation metrics such as Precision-Recall curves and AP for the second survey. Precision measures the accuracy of predictions, indicating the proportion of correct predictions. Recall quantifies the effectiveness of identifying positive instances, representing the proportion of correctly identified positives (Hui, 2018).

In Figures 4.4 and 4.5, we illustrate the training and validation curves for the object detection models trained to identify food items, with and without plastic bags, respectively. Figures 4.4 and 4.5 display the training and validation loss curves, encompassing box loss, class loss, and focal loss. Additionally, Figures 4.6 and 4.7 display the mAP scores at IoU thresholds of 0.5 (mAP50) and from 0.5 to 0.95 (mAP50-95), alongside performance metrics such as precision and recall for both training and validation datasets for both models. These figures offer insights into the models' loss optimization and evaluation metrics across epochs, providing a comprehensive view of their training and validation performances.

Furthermore, in Figures 4.8 and 4.9, a comprehensive view is provided of the training and validation performances of the combined model, which detects food items with and without plastic bags, offering insights into its loss optimization and evaluation metrics across epochs.

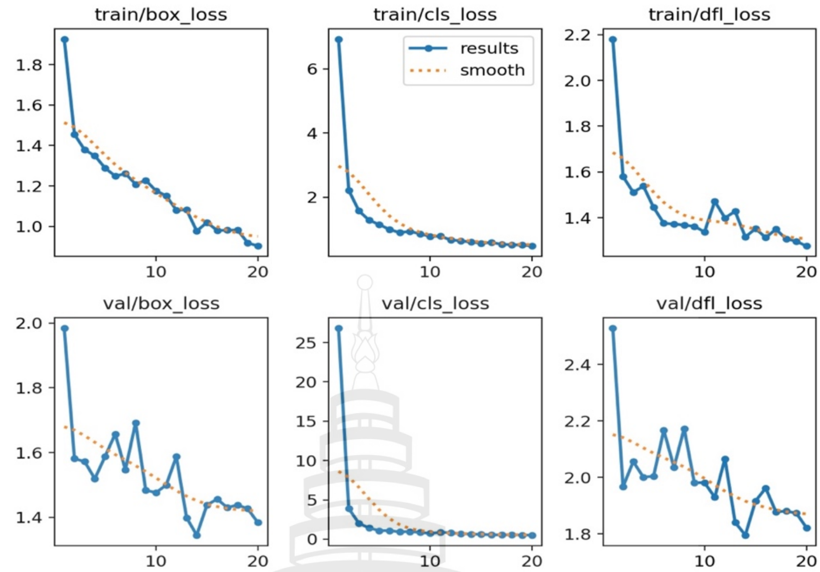

Figure 4.4 Training and Validation Curves for without Plastic Bags Model

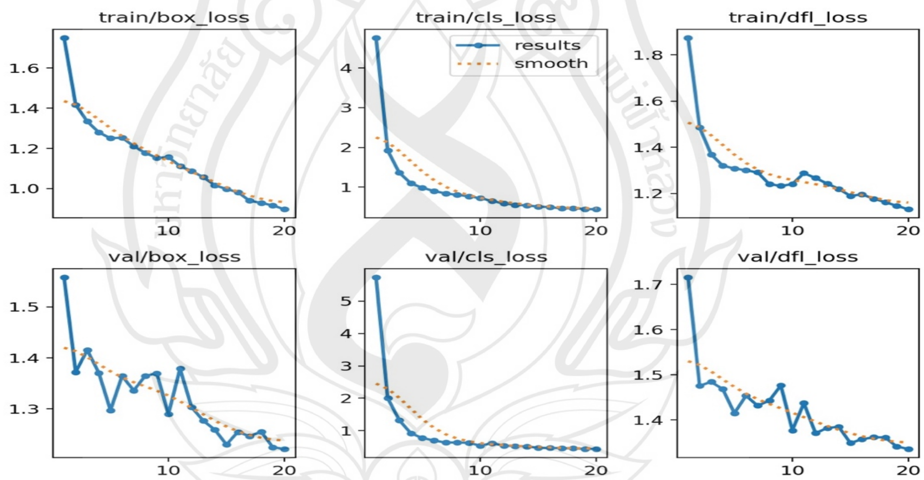

Figure 4.5 Training and Validation Curves for with Plastic Bags Model

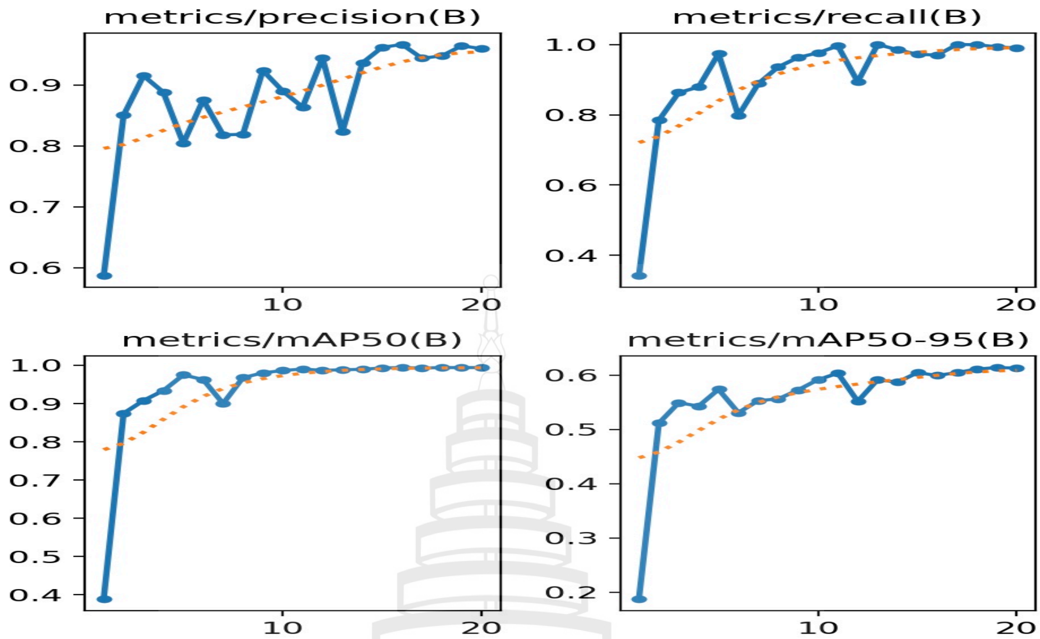

Figure 4.6 Evaluation Metrics for without Plastic Bags Model

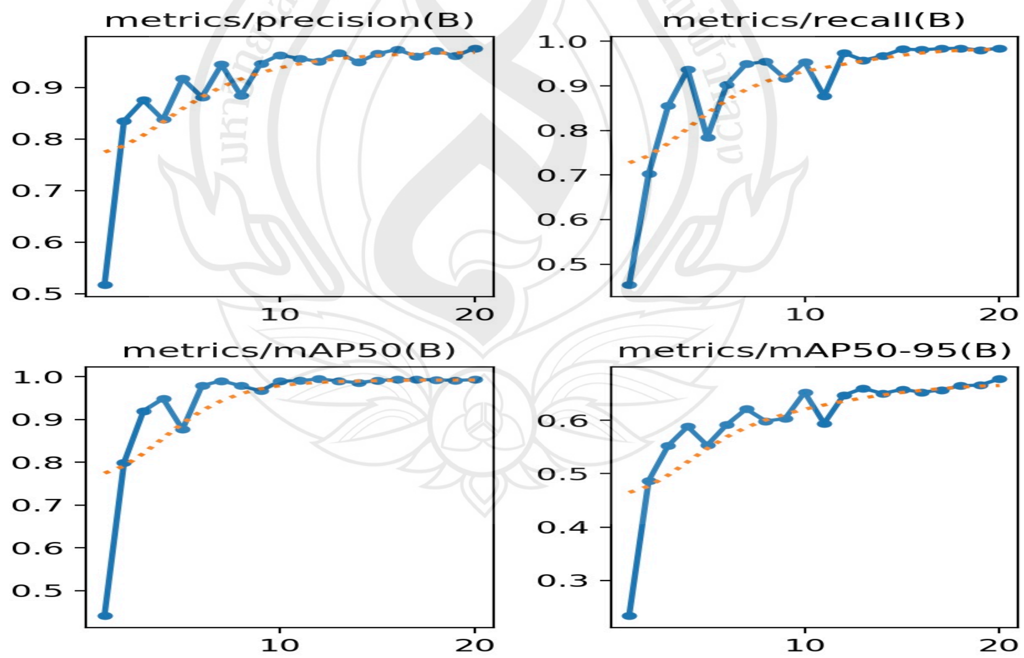

Figure 4.7 Evaluation Metrics for with Plastic Bags Model

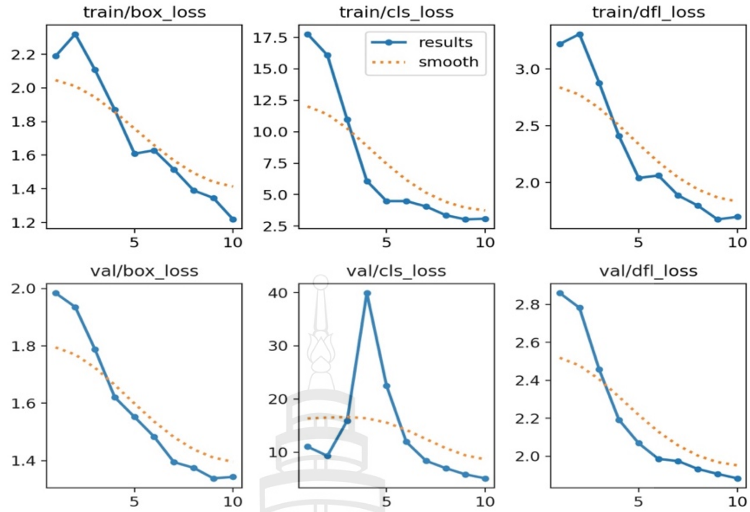

Figure 4.8 Training and Validation Curves for Combined Model

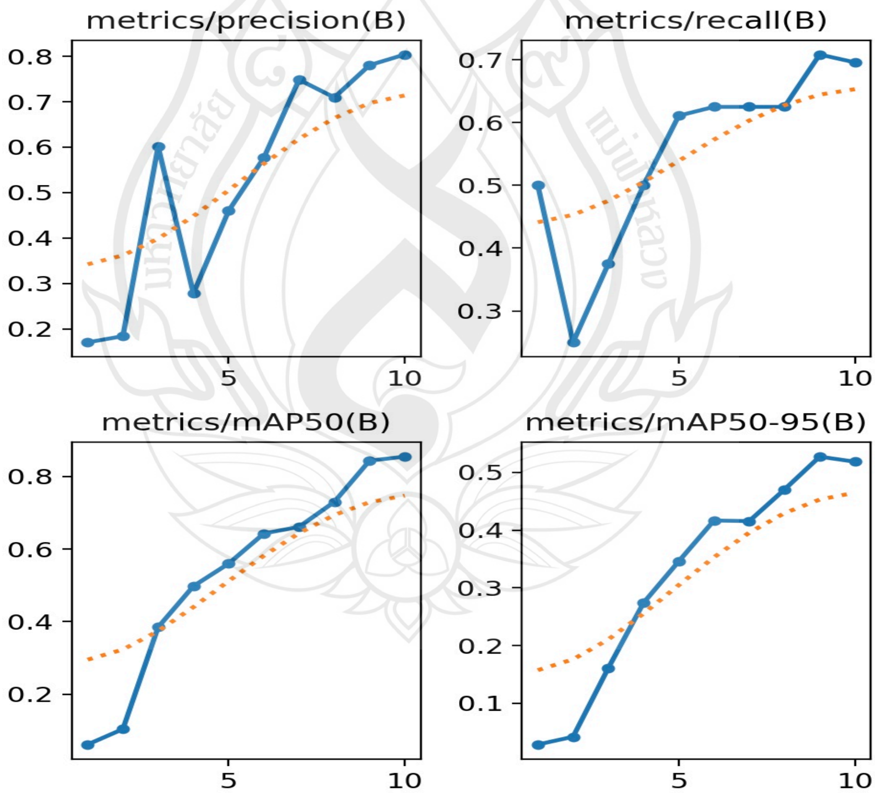

Figure 4.9 Evaluation Metrics for Combined Model

Using an IoU threshold of 0.5, the model designed for detecting food items with plastic bags achieved an mAP score of 99.3%. Conversely, the model targeting food items without plastic bags achieved a slightly higher mAP score of 99.4%. Additionally, the third model, capable of detecting both types of food items, achieved an mAP of 84.4%. Detailed metrics results are presented in Table 4.4. These results demonstrated commendable performance for all three models within their respective domains, with a marginal superiority observed for the detection of items sold without plastic bags.

Table 4.4 Object Detection Models Performance

Model	**mAP**	**Precision**	**Recall**
Food Without Plastic Bags	99.4 %	96.4 %	99.3 %
Food with Plastic Bags	99.3%	97.6 %	98.4 %
Combined Model	84.4%	78.1 %	70.8 %

In Figures 4.10 and 4.11, the predictions of detected classes for object detection models trained to identify food items with and without plastic bags are depicted. Additionally, in Figure 4.12, the predictions generated by the third model, which is the combined model capable of detecting food items with and without plastic bags, are showcased. Each figure showcases the bounding boxes and class labels generated by the respective models, providing visual insights into the accuracy and effectiveness of the detection process for each category.

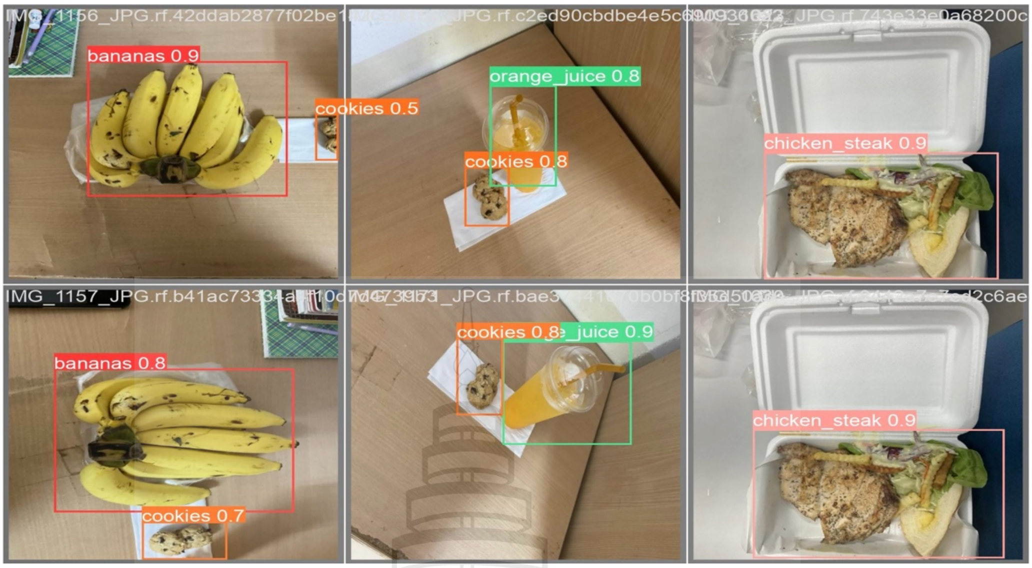

Figure 4.10 Predictions for Food without Plastic Bags Model

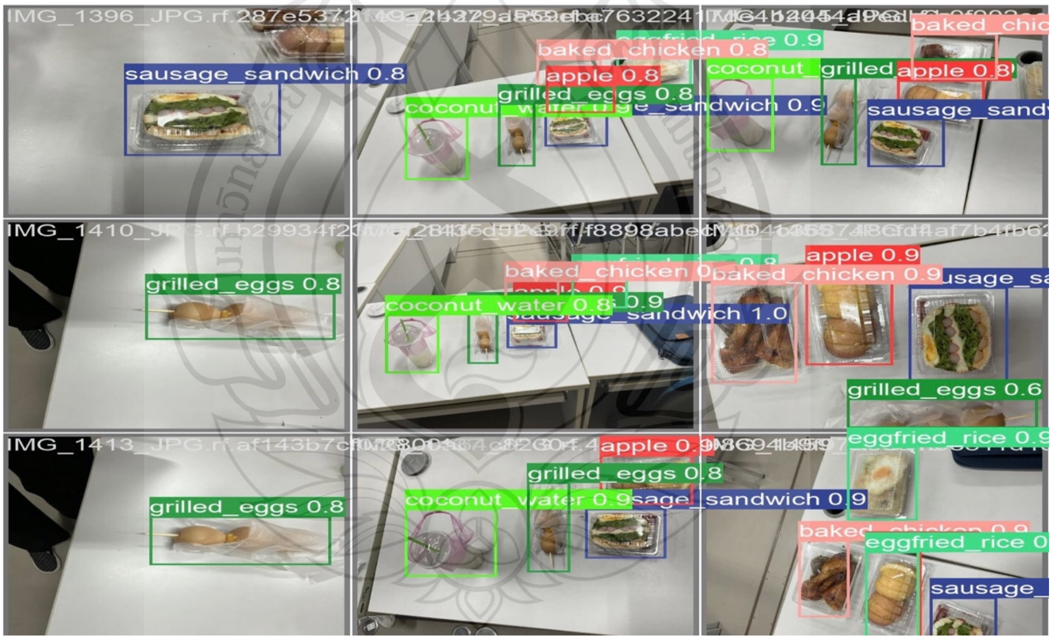

Figure 4.11 Predictions for Food with Plastic Bags Model

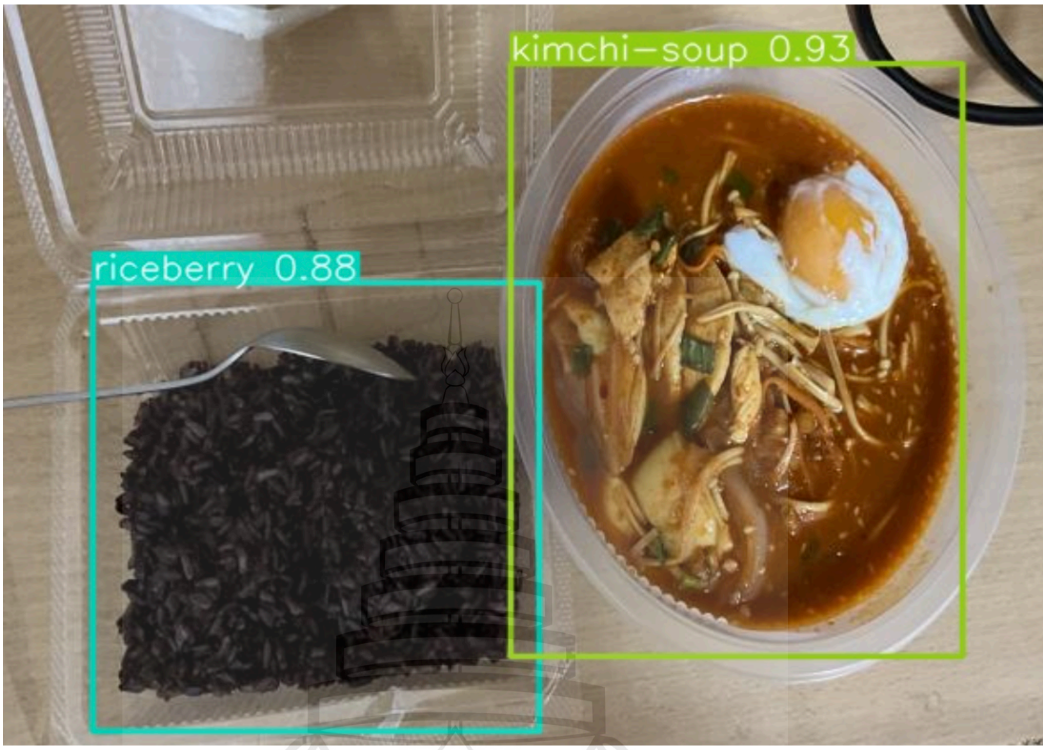

Figure 4.12 Predictions for Combined Model

After obtaining the correctly predicted classes from both object detection models, we utilized three association rule learning algorithms—Apriori, FP-Growth, and Eclat for market basket analysis to analyze the detected food items. The results of each algorithm provided valuable insights into consumer purchasing behavior at the night market. Apriori revealed frequent itemsets and association rules based on item co-occurrence patterns, while Eclat emphasized popular item combinations regardless of size. FP-Growth efficiently handled large datasets, uncovering nuanced patterns in consumer preferences. These findings offer actionable insights for optimizing product offerings and enhancing the shopping experience.

Table 4.5 Association Rules Evaluation

Algorithms	**Rules**	**Support**	**Confidence**	**Lift**
Apriori	8	20 %	50 %	2
FP-Growth	10	15 %	50 %	2
Eclat	13	10 %	50 %	2.5

We evaluated three association rule learning algorithms -Apriori, FP-Growth, and Eclat—based on their ability to generate rules with high support, considering the minimum thresholds required for rule generation. The results are summarized in Table 4.5. Among these, the Apriori algorithm produced the highest support value of 20%, followed by FP-Growth with 15%, and Eclat with 10%. It is important to note that these support values represent the minimum thresholds for generating association rules. If the support is increased beyond these thresholds, rules may not be obtained due to the stricter criteria for rule generation. When comparing algorithms, priority is given to selecting the one with the maximum support to ensure robust rule generation. The selection of Apriori aligns with the objective of maximizing support, aiming to address previously observed low support values. Subsequent analysis focused on the generated rules and their accuracy to gain insights into consumer behavior and product associations.

The association rules derived from the analysis of food items provide valuable insights into consumer behavior and preferences as shown in Table 4.6. These rules revealed meaningful associations between different categories of food items, including meals, beverages, fruits, and desserts. For instance, common food pairings such as chicken steak with coconut water and sausage sandwiches with bananas highlighted prevalent consumer choices. Moreover, specific meal combinations such as egg fried rice with orange juice and mulberry juice, along with diverse food choices like baked chicken with apples and pumpkin seeds, reflected the variety in consumer preferences. Subsequently, we delved into the accuracy of these rules on the newly collected test dataset.

Table 4.6 Food Item Associations from Apriori Algorithm

Rules	Antecedent	Consequent
1	chicken steak, sausage sandwich	coconut water, egg tarts
2	sausage sandwich, coconut water	bananas, sweet crepes
3	coconut water, sweet crepes	kimchi soup, chicken rice
4	egg fried rice, orange juice, mulberry juice	cookies, sausage sandwich
5	bananas, oranges, egg tarts	chicken steak
6	egg fried rice, sweet crepes	oranges
7	grilled eggs, coconut water	oranges
8	baked chicken, apples, kimchi soup, orange juice, egg tarts	pumpkin seeds

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \qquad (1)$$

Equation (1) calculates accuracy by dividing the number of correct predictions by the total number of predictions. By applying this equation to our test dataset, we obtained the accuracy of the predicted rules as demonstrated in Table 4.7. Rule 8 obtained 0 accuracy because there were no transactions in the test dataset that included all the items in the antecedent (baked chicken, apples, kimchi soup, orange juice, and egg tarts) resulting in the purchase of the consequent item (pumpkin seeds). As a result, the rule did not make any correct predictions in the test dataset, leading to 0 accuracy. The average accuracy of these association rules on the new dataset, consisting of 20 transactions, was calculated to be 84%. This indicates that the association rules correctly predicted the outcome for 84% of the transactions in the test dataset.

Table 4.7 Accuracy of the Rules on the New Test Dataset

Rules	Transactions	Accuracy
1	20	1.0
2	20	1.0
3	20	1.0
4	20	1.0
5	20	0.85
6	20	1.0
7	20	0.9
8	20	0.0
	Average Accuracy =	**0.84**

These results further confirm the effectiveness of the association rules in capturing meaningful associations between items in transaction data, providing valuable insights into consumer behavior and preferences in the context of market basket analysis. These findings align with real-world scenarios, indicating sensible and meaningful food combinations likely to be observed among consumers. Overall, these association rules offer actionable insights for menu planning, product placement, and marketing strategies, enabling businesses to better understand and cater to consumer needs and preferences in the food industry.

CHAPTER 5

CONCLUSION

5.1 Research Conclusion

The study sought to identify distinct customer groups and understand their purchasing patterns and behaviors at the night market. In the first experiment of the study, K-means clustering was applied for customer segmentation, successfully dividing customers into five distinct groups using the elbow method and clustering validation metrics such as Silhouette, Davies Bouldin, and Calinski-Harabasz. For market basket analysis, the research established four association rules between products based on their support, confidence, and lift, utilizing the Apriori algorithm.

The application of these association rules to a new group of 20 customers for predictive purposes revealed that the rules are well accepted. Particularly, the second rule, which relates street food and snacks to beverages, emerged as the strongest exhibited rule, correctly predicting 10 out of 20 customers among the four rules in association rule mining. These findings underscore the effectiveness of machine learning techniques in extracting meaningful insights from customer behavior data in market settings.

In the first experiment, manual data collection for market basket analysis proved to be time-consuming and inefficient, highlighting a significant limitation in the process. Additionally, analyzing data from night markets, particularly for market basket analysis, poses significant challenges. The first experiment collected images of products from customers at these markets, yet these images often depicted food items and products in plastic bags or boxes, complicating identification. To address this limitation, a second experiment was conducted. The second experiment utilized object detection techniques to automatically label food items and products, thereby enhancing the accuracy and efficiency of market basket analysis. The effectiveness of these techniques was evaluated using standard evaluation metrics such as Precision-Recall curves and mean Average Precision (mAP). The results demonstrated commendable performance, with the mAP score of 99.4% for food items without plastic bags and 99.3% for those with plastic bags. Notably, the combined model, capable of detecting both types of food items, achieved the mAP of 84.4%.

Additionally, three association rule learning algorithms -Apriori, FP-Growth, and Eclat—were utilized for market basket analysis to uncover meaningful associations among food categories. Among these algorithms, the Apriori algorithm produced the highest support value of 20% and confidence of 50%, generating a total of 8 rules. The accuracy of these association rules on a new dataset, comprising 20 transactions, was calculated to be 84%.

These findings offer actionable insights for businesses in the food industry, empowering them to tailor marketing strategies and product offerings to better align with consumer needs and preferences. For instance, night market organizers can implement joint promotions, such as offering discounts when customers purchase certain items together, encouraging larger purchases. Additionally, organizers can recommend complementary products when a customer buys one item. Collaborative marketing strategies, including social media campaigns and joint advertisements, can also be employed to highlight the appeal of combining these products.

This predictive analysis allows for making inferences or predictions based on learned associations, providing insights into the relationships and dependencies within a new group of consumers. By understanding consumer preferences and identifying prevalent food associations, night market organizers and sellers can optimize their marketing strategies, product placements, and menu offerings to better cater to consumer needs, ultimately enhancing customer satisfaction and driving business success. In conclusion, this study contributes to a deeper understanding of consumer behavior and product associations in the food industry, paving the way for future research endeavors in this domain.

5.2 Research Discussion

The significance of the findings lies in the use of object detection techniques, which have significantly improved the efficiency and accuracy of market basket analysis in the context of night markets. The association rules generated provide valuable insights that can drive marketing strategies and business decisions. Compared to previous studies that relied on manual data collection methods, this study demonstrates substantial advancements in the identification of food items and the accuracy of market basket analysis. The implications for the industry are considerable, as businesses can implement these findings to enhance their marketing strategies, leading to potential economic benefits for night market organizers and food vendors.

5.3 Research Limitations

Several limitations should be acknowledged in this study. The dataset is limited to specific night markets or regions, which might not be representative of all night markets, and there may be potential biases in the collected images, such as variations in lighting and image quality. The generalizability of the object detection model to other contexts or types of markets may be limited, and there could be challenges in applying the model to different types of food items or packaging. Additionally, the performance of the association rule learning algorithms is dependent on the dataset size and composition, and the accuracy of the generated rules may need improvement for better predictions.

5.4 Future Work

Future research can explore several directions to address these limitations. Expanding the dataset by collecting more diverse datasets from various night markets and regions can enhance the model's robustness and generalizability. Incorporating more types of food items and packaging can also improve the model's versatility. Advanced analytical techniques, such as more sophisticated machine learning and deep learning methods, can be explored to further improve the accuracy of object detection and market basket analysis. Investigating hybrid models that combine multiple algorithms may also yield better performance.

Linking customer segmentation and market basket analysis enables the identification of distinct buying patterns within each customer group. Conducting market basket analysis for each segment uncovers specific rules and associations that characterize their shopping behaviors. This allows businesses to tailor marketing strategies and product placements to better meet the needs of each segment, enhancing customer satisfaction and increasing sales. Leveraging these insights also aids in effective inventory management and stock optimization, ensuring product availability and reducing waste. Overall, integrating customer segmentation with market basket analysis creates a more personalized and efficient shopping experience, driving customer loyalty and boosting revenue.

Developing real-time market basket analysis systems that provide immediate insights to vendors and organizers can be a valuable direction for future work. Integrating these systems with point-of-sale (POS) systems can facilitate seamless data collection and analysis. Additionally, conducting in-depth studies on consumer behavior using the generated association rules can uncover deeper insights, and exploring the impact of various marketing strategies based on these rules on consumer purchasing behavior can provide actionable recommendations for businesses. By addressing these areas, the study can offer a comprehensive overview of the research, its implications, limitations, and future directions.

REFERENCES

REFERENCES

Aldino, A. A., Pratiwi, E. D., Setiawansyah, Sintaro, S., & Dwi Putra, A. (2021). Comparison of market basket analysis to determine consumer purchasing patterns using FP-Growth and Apriori algorithm [Paper presentation]. In *Proceedings of the 2021 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE)* (pp. 29-34). https://doi.org/10.1109/ICOMITEE53461.2021.9650317

Ali, M. (2023, January). *Association rule mining in python tutorial*. Data Camp. https://www.datacamp.com/tutorial/association-rule-mining-python

Aryuni, M., Madyatmadja, E. D., & Miranda, E. (2018). Customer segmentation in XYZ Bank using k-means and k-medoids clustering. In *Proceedings of the 2018 International Conference on Information Management and Technology (ICIMTech)* (pp. 412-416). https://doi.org/10.1109/ICIMTech.2018.8528086

Bamrongpol, D., Sornsaruht, P., & Deebhijarn, S. (2020). Antecedents to Thai night market visitor revisit intention. *Asia-Pacific Social Science Review, 20(3),* 182-19.

Berlilana, T., Hariguna, T., & Hananto, A. R. (2022). Analysis of customer product interests using the market basket analysis model with hash-based algorithm and association rules. In *Proceedings of the 2022 4th International Conference on Cybernetics and Intelligent System (ICORIS)* (pp. 1-5). https://doi.org/10.1109/ICORIS56080.2022.10031409

El Ghezzaz, A. (2023, August 8). *Unveiling the power of data mining techniques: Types of data, methods, applications*. Medium. https://medium.com/@aelghezzaz/unveiling-the-power-of-data-mining-techniques-types-of-data-methods-applications-98b989c05a7a

Gupta, R., Verma, A., & Topal, H. O. (2021). Customer segmentation of Indian restaurants on the basis of geographical locations using machine learning. In *Proceedings of the 2021 International Conference on Technological Advancements and Innovations (ICTAI)* (pp. 382-387). https://doi.org/10.1109/ICTAI53825.2021.9673153

Hui, J. (2018, March 7). *mAP (mean Average Precision) for Object Detection.* Medium. https://shorturl.at/qDMUW

Ishak, N. K., Aziz, K. A., & Ahmad, A. (2012). *Dynamism of a night market. Journal of Case Research in Business and Economics, 4*, 1-10.

JavaTpoint. (2023, January). *Association rule learning.* https://www.javatpoint.com/association-rule-learning

Keita, Z. (2022, September). *YOLO object detection explained: A beginner's guide.* Data Camp. https://shorturl.at/kpGKL

Li, R., Wang, X., & Wang, Y. (2021). Night market cultural transformation and upgrading. *Journal of Service Science and Management, 14*(4), 412–428.

Mazlan, M. S., Meran, N. E. S., Kamal, M. H. M., & Ramli, N. (2017). Decision to visit night market from the Malaysian customer perspective. *Journal of Tourism, Hospitality & Culinary Arts, 9*(2), 143-151.

Mehta, V., Mehra, R., & Verma, S. S. (2021). A survey on customer segmentation using machine learning algorithms to find prospective clients. In *Proceedings of the 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)* (pp. 1-4). https://doi.org/10.1109/ICRITO51393.2021.9596118

Muhammadshabrozshahab. (2023, October 8). *YOLO V8. Working principle.* Medium. https://medium.com/@muhammadshabrozshahab/yolo-v8-104f1375242c

Pandey, D. (2022). Object detection in Indian food platters using transfer learning with YOLOv4. In *Proceedings of the 2022 IEEE 38th International Conference on Data Engineering Workshops (ICDEW)* (pp. 101-106). https://doi.org/10.1109/ICDEW55742.2022.00021

Phyo, T., & Uttama, S. (2023). Unveiling Patterns in the Night Market: A Machine Learning Approach to Customer Segmentation and Market Basket Analysis. In *Proceedings of the 7th International Conference on Information Technology (InCIT)* (pp. 514-519). https://doi.org/10.1109/InCIT60207.2023.10413047

Rentech Digital. (2024, January 9*). List of night markets in Asia.* https://rentechdigital.com/smartscraper/business-report-details/asia/night-markets

Selvaraj, N. (2023, April 24). *How to perform market basket analysis in Python*. 365 Data Science. https://365datascience.com/tutorials/python-tutorials/market-basket-analysis/

Sharma, A., & Babbar, H. (2023). Analysis of data mining algorithms in market basket analysis. In *Proceedings of the 2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT)* (pp. 275-280). https://doi.org/10.1109/InCACCT57535.2023.10141816

Singh, J., Jaiswal, K., Singh, M., Sama, M., & Singhal, S. (2023). Market segmentation using ML. In *Proceedings of the 2023 International Conference on Disruptive Technologies (ICDT)* (pp. 703-707). https://doi.org/10.1109/ICDT57929.2023.10150639

Statista. (2022, July 12). *Seoul night market visitor number South Korea 2015-2017*. https://www.statista.com/statistics/933930/south-korea-seoul-night-market-visitor-number/

Tan, S. W., Lee, C. P., Lim, K. M., & Lim, J. Y. (2023). Food detection and recognition with deep learning: A comparative study. In *Proceedings of the 2023 11th International Conference on Information and Communication Technology (ICoICT)* (pp. 283-288). https://doi.org/10.1109/ICoICT58202.2023.10262523

Tzutalin, T. (n.d.). *LabelImg.* GitHub Repository. Retrieved March 19, 2024, from https://github.com/tzutalin/labelImg

Yilmaz, A. O. (2023, July 3). *Understanding the differences between deep learning and machine learning.* Medium. https://aoyilmaz.medium.com/understanding-the-differences-between-deep-learning-and-machine-learning-eb41d64f1732