WEB-BASED INTELLIGENT PACKAGING EVALUATION (WIPE) PLATFORM

By

Mahsa Tavasoli

A THESIS

Submitted to Michigan State University in partial fulfillment of the requirements for the degree of

Packaging-Master of Science

2023

ABSTRACT

Package and product systems are exposed to distribution hazards during distribution. Several packaging evaluations have been proposed to predict damages resulting from these hazards. These evaluation results let the designer know how to improve packaging functions. The packaging industry mostly relies on laboratory evaluations under controlled conditions. The availability of online purchasing platforms and the global market demand rate's continual growth are critical factors for changing traditional distribution to e-commerce distribution. However, ecommerce distribution has more handling points that cause more risk of product damage. Also, ecommerce distribution faces many unexpected hazards that would not be captured during physical tests. Therefore, To overcome these limitations, we proposed a web-based intelligent packaging evaluation (WIPE) platform as an alternative evaluation method to extract packaging defects and find causes of damage without redoing laboratory tests. In this study, we used machine learning algorithms and association rule mining to determine customers' attitudes and the relationship between critical damaged features of products or packaging based on customers' reviews on a real e-commerce platform. The main contributions of this work are: firstly, the automation of the data flow of packaging evaluation based on customer reviews; secondly, the correlation of images and text of online product reviews; and thirdly, the determination of relationships between the most frequent words in customer reviews to predict damages and their causes and effects. The last contribution is the most important contribution of this research, and to the best of our knowledge, this platform is the first that brings a combination of sentiment analysis and association rule mining to the packaging evaluation area. In this study, the WIPE platform's performance was evaluated by considering two case studies, laundry detergent liquid bottles and pods on the Amazon platform.

Copyright by MAHSA TAVASOLI 2023

ACKNOWLEDGMENTS

First and foremost, I would like to thank my thesis supervisor, Dr.Euihark Lee, for all his compassion and excellent guidance on this project over the past two years. This paper would not have been accomplished without his help and dedicated involvement in each phase of the work. I would also like to express my gratitude to the members of my examination committee, Dr. Monireh Mahmoudi and Dr.Qiben Yan. I would also like to thank the School of Packaging faculty for supporting me during my graduate study.

Most importantly, none of this could have happened without my family and friends. I would like to thank my great parents, Fatemeh and Hossein; lovely sisters, Marzieh, Mehrnaz, and Mohadeseh; supportive brothers, Mehrdad and Mohsen; dear friend, Hannah; and my cute niece and nephews, Arian, Aram, Hannah, Ava, Elena, and Ala for always believing in me.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS	vi
CHAPTER 1 INTRODUCTION	. 1
CHAPTER 2 BACKGROUND AND LITERATURE REVIEW	4
CHAPTER 3 PROBLEM STATEMENT AND METHODOLOGY	37
CHAPTER 4 RESULT AND DISCUSSION	52
CHAPTER 5 CONCLUSION	31
BIBLIOGRAPHY	34
APPENDIX A: ASSOCIATION RULES FOR CASE STUDY #1	€
APPENDIX B: ASSOCIATION RULES FOR CASE STUDY #2) 4

LIST OF ABBREVIATIONS

FEM: Finite Element Modeling				
ASTM: American Society for Testing and Materials				
ISO: International Student Orientation				
ISTA: International Safe Transit Association				
TAPPI: Technical Association of the Pulp and Paper Industry				
WIPE: Web-based Intelligent Packaging Evaluation				
DC: Distribution Center				
B&M : Brick and Mortar				
ML: Machine Learning				
AI: Artificial Intelligence				
KNN: K-Nearest Neighbor				
FP-Growth: Frequent Pattern Growth Algorithm				
SVM: Support Vector Machine Learning				
LR: Logistic Regression				
GAN: General Adversarial Networks				
ANNs: Artificial neural networks				
SA: Sentiment Analysis				
NLP: Natural language processing				
LSTM: Long short-term memory				
TF-IDF: Term Frequency-Inverse Document Frequent				

CHAPTER 1 INTRODUCTION

1.1 Introduction

Packaging is a key component in the supply chain as packaging protects products from vibration, pressure changes, temperature changes, and shocks throughout their life cycle. However, to determine if a product's packaging is effective in preserving the items it contains, an evaluation of the packaging is needed. Packaging evaluation measures a product's packaging characteristics based on its response to main packaging functions such as protection, contamination, apportionment, unitization, communication, and convenience.

Many distribution hazards can occur throughout the product's transferring process, which may damage either the product or the package. These distribution hazards can be the result of road conditions, which are often encountered due to long travel distances with countless stress possibilities(Huart et al., 2016). Since one of the goals of packaging is to get a product from start to point to the destination without unacceptable loss or damage economically and practically, failed packaging without proper evaluation can result in reduced customer satisfaction, brand reputation, sales, and profits (Saghir, 2002).

Packaging evaluation methods include field tests, laboratory experiments, and numerical analysis such as Finite Element Modeling (FEM) analysis. The field evaluation method is based on real tests. Therefore, these tests are simulated under real conditions like the real world. However, real tests are costly and time-consuming because they need to provide additional equipment and travel around the world. The next evaluation is laboratory tests. Several standards organizations such as the American Society for Testing and Materials (ASTM), International Student Orientation (ISO), International Safe Transit Association (ISTA), and Technical Association of the Pulp and Paper Industry (TAPPI) publish test methods for laboratory package

evaluations. Since these experiments are performed under controlled settings, they are less correlated with actual environmental factors or customer requirements (Molina-Besch & Pålsson, 2020). The third group of packaging evaluation is numerical analysis for packaging evaluation. These methods are known because they are the most cost and time-effective methods to simulate since they can simulate real conditions without needing test materials or costs. But it has some limitations in getting material properties and would be complex for complicated conditions.

Thus, it is necessary to propose a method to measure package quality based on real conditions, time, and cost-efficiency. Thanks to machine learning technology, another packaging evaluation has been proposed based on artificial intelligence and people reviews on e-commerce. Online product reviews have some benefits. It shows customers' needs and wants because it is based on real experience and real conditions. Also, customers' reviews help designers with informed decision-making through damage prediction and damage causes prediction.

In previous research of our group, a systematic approach was proposed to evaluate package performance by analyzing customer reviews. However, this method has some limitations. It is not automated and only focused on the text of reviews. Also, although the most frequent words in reviews can give us insights into the packaging problems, there is no way to find the relationships between these frequent words.

This thesis looks at packaging evaluation methods and proposes a Web-based Intelligent Packaging Evaluation (WIPE) platform for packaging evaluation based on machine learning approaches through association rule mining. Two packaging evaluation case studies have been studied to evaluate the proposed method's performance.

2

1.2 Objectives of the study

The purpose of our research is to have a more effective packaging evaluation from the reviews on e-commerce websites. To achieve this goal, this thesis seeks to address three specific objectives, which are:

Objective 1: To automate the data flow of packaging evaluation based on customers reviews

Objective 2: To Correlate images and text of online product reviews

Objective 3: To determine relationships between the most frequent words in customer reviews to predict damages and their causes and effects

1.3 Structure of the thesis

The thesis has the following structure. The first chapter of the thesis includes an introduction to show clear focus, purpose, and direction. Chapter 2 focuses on the background of packaging evaluation and a literature review of proposed evaluation methods. Chapter 3 provides a problem statement and a novel packaging evaluation methodology. From there, Chapter 4 will discuss the results and evaluate the packaging of two products as case studies based on the proposed method. To finalize, Chapter 5 will wrap up the proposed method's conclusions and areas of improvement.

CHAPTER 2 BACKGROUND AND LITERATURE REVIEW

2.1 Packaging distribution, distribution hazards of E-commerce distribution

We must first define packaging and its primary functions to comprehend packaging distribution and evaluation importance. Paine (Paine & Paine, 1983) proposed a definition for packaging: a coordinated system of arranging products for transport, distribution, storage, retailing, and final use. Packaging is like a tool for ensuring secure delivery to the final customer in good condition at the lowest cost and with the least amount of environmental effect.

According to Livingstone and Sparks (Livingstone & Sparks, 1994), Lockamy III (Lockamy, 1995), and Obertson (Obertson, 1990), there are six main functions of packaging:

- Protection: packaging should keep the product's content protected and secured. Any damage during distribution could result in protection failure. These damages can be physical or environmental.
- 2. Containment: the preservation and maintenance of the content. If there is a leak or loss, the results can vary from a little inconvenience to a significant and perhaps devastating failure.
- Apportionment: the process of reducing massive, high-volume manufacturing to manageable sizes.
- 4. Unitization: the packaging levels can be unitized to improve material handling and transportation efficiency.
- 5. Communication: this function means identification of the packaging throughout the supply chain and product information. All packaging types must show information such as product name, number of products, and product identification number with the least language barrier.

6. Convenience: to make utilizing products simple and easy. Easy-open and reclose feature is one example of convenience features of packaging.

Primary, secondary, and tertiary packaging are three levels of packaging that must secure the product and fulfill all critical functions from manufacturing to final delivery (Pålsson, 2018). Primary packaging is the first of the three levels of packaging. Primary packaging, often known as a consumer or retail packaging, is the packaging that comes into direct touch with the product itself. Primary packaging's major goals are to protect and/or preserve, contain, and inform the consumer.

Secondary packaging can aggregate several primary packages together or combine numerous components into a single object. The primary purposes of secondary packaging are unitizing, protecting, and bundling. Tertiary packaging has two main goals. Merging many products (that are already contained in the two previous levels of packaging) into a single, larger container eases handling. Secondly, during transit and storage, this third level of packing protects products and the first two packaging levels. Figure 2.1 represents these packaging levels.



Figure 2.1

The packaging Levels: the primary packaging, secondary packaging e.g., several in one pack; tertiary packaging, e.g., a pallet for shipping

Over the supply chain, many potential sources of hazards disrupt these main packaging functions. For example, leakage at the neck and misalignment of a cap would cause inconvenience failure. Also, any damage in the handling stage, such as falling from pallets, can cause protection failure. Packaging design and redesign aim to optimize the six primary packaging functions under these conditions(Emblem & Emblem, 2012). In this study, we focused on distribution hazards and reviewed related works for packaging evaluation for the packaging distribution.

The following potential risks may raise concern when products and packaging are distributed across supply chains (Molina-Besch & Pålsson, 2020), (Brandenburg, 1991):

- Manual handling: manual handling creates hazards for packages during dropping, loading, and unloading trucks, stacked containers in storage, and bouncing off the walls of the transport vehicle.
- Mechanical handling in a warehouse: In warehouse and distribution centers, forklifts and conveyors handling packaged goods create hazards such as shock impacts that cause considerable damage, negatively impacting the lifetime of pallets and packaged goods (Masis et al., 2022).
- 3. Transport vehicle's impact: the starting, stopping, and jolts resulting from vehicle movement cause damage to loads. An example of damaged packages that shifted and dropped while being transported is shown in Figure 2.2-(A).
- 4. Vehicle Vibration: the vibration results from the engine motion, and contact of the vehicle with the road, negatively affects packages.
- 5. Environmental conditions: Climate hazards occur due to temperature, humidity, and air pressure changes over transit. Low temperatures can cause Aqueous liquids to freeze, which also fractures containers. Hot temperatures can also cause adverse effects, such as

increased diffusion coefficients, entry of water vapor causing contamination, hydrolysis, and oxidation. As a result of a decrease in pressure, thin containers may burst, or strip packs may inflate in mountainous regions or during a flight (Mandal et al., 2022). Figure 2.2-(B) presents water damage in palletized products while in distribution.

These five sources cause impacts, vibration, shocks, and improper temperature and humidity changes as the package and product travel through the distribution system shown in Figure 2.2. Package systems can be designed to minimize distribution damage, but sometimes the product must be redesigned to endure (Brandenburg, 1991).



(A)

(B)

Figure 2.2 Damaged packages due to distribution hazard(Logistic Zipline, 2018)

The availability of online purchasing platforms, as well as the world market demand rate's continual growth, are critical driving factors for changing traditional distribution to E-commerce distribution(Y. Wang et al., 2020). E-commerce refers directly to the consumer business. The trading of products and services through the internet is known as e-commerce. Amazon (Amazon, 2022b) predicted that 2.14 billion consumers would make online purchases this year, and more than 150 million Prime members are now shopping at Amazon locations. Figure 2.3 shows the

worldwide current and predicted growth of E-Commerce, based on annual sales over ten years from 2014 to 2024.



Figure 2.3 Global retail e-commerce sales (in billion USD) worldwide over ten years (International trade administration, 2021)

The packaging industry is changing on several levels in e-commerce because protective packaging requirements and demand are particularly impacted by e-commerce distribution hazards.

In traditional retail, an item is produced, packaged, and sent to a distribution center in palletized form—this retail is often known as the Brick and Mortar (B&M) business model. Then, the packaged items are shipped by case to a local shop from the distribution center (DC). After being taken out of their protective case package for the first time, the goods are stocked on the shelves for customers to buy and consume. Figure 2.4 depicts every step in the Brick-and-Mortar distribution model.

In the e-commerce distribution chain, All packages are delivered to e-commerce distribution centers, then kept in DC inventory. Then, some products are safely wrapped for delivery, which results in more handling. The order is then repackaged in single or multiple-item

boxes for shipment. The packages are then transported by truck or airplane to DC from anywhere in the nation or world. The received packages are processed and sorted in DC before rerouting to the local DC. After that, they are transported to the consumers' locations. Figure 2.5 illustrates a product's distribution route from the manufacturer to the consumer in the E-commerce model.



Figure 2.4 Products Traveling route from Manufacturer to Consumer through Brick-and-Mortar

distribution chain(Dunn, 2018)



Figure 2.5 Products Traveling route from Manufacturer to Consumer through E-Commerce retail distribution chain (Dunn, 2018)

By comparing these two distribution types, it can be concluded that the distribution network for e-commerce can contain almost three times as many touch points as the distribution chain for traditional retail. So, there is an extremely significant possibility of such goods getting destroyed along the trip (Dunn, 2018).

2.2 Introduction of packaging evaluation

Since it is essential to guarantee the safety of a product-package system traveling against hazards over the distribution system by freight train, truck, container ship, and so on, several packaging evaluation tests have been proposed. These methods are categorized into field evaluation, laboratory evaluation, and numerical evaluation.

Field tests usually include preparing a sample package and sending it on a long round trip to a location that is expected to be hazardous from the viewpoint of damage. One of the main applications of field tests is to be used as a validation tool to verify expected results from laboratory tests. Another application of field shipping tests is to generate damage statistics for Comparison.

A van shipment test was performed by Böröcz et al. (Böröcz & Molnár, 2020) for the typical delivery along the routes shown in Figure 2.6. This study demonstrated the impact of free movement space around the packages during shipping and gave a statistical analysis of the nature of random vibration. The results of this study can be compared with typical vibration testing profiles utilized for packaging testing. Packaging engineers may find this research's collected data helpful as a technical tool for checking parcel packages prior to shipping. Additionally, a new measurement technique was presented in this study to assess and examine the vibration levels in stacked small, packaged products that are not held by installation during transportation (Böröcz & Molnár, 2020).

Laboratory and physical testing are another evaluation method. A laboratory test aims to produce damage comparable to the damage brought on by real shipment. ASTM standards also include well-known standardized tests. As an illustration, the ASTM D1596 Standard Test Method for Shocking Absorbing Characteristics of Package Cushioning Materials is used to experimentally determine a cushion material's shock properties (Hagman et al., 2017). Moreover, Pallet testing standards are often used to simulate the damage caused by forklift handling in laboratory settings. ASTM 1185 and ISO 8611 can both be used to determine the pallet resistance to impacts by forklifts (Böröcz & Molnár, 2020; Kipp, 2000)



Figure 2.6 Route measured for this study (Kipp, 2000)

Numerical methods such as Finite Element Modeling (FEM) analysis are another packaging evaluation method used in developing and evaluating new products and packages. For example, it might be applied in the paper industry to better understand what board qualities are perfect for packaging effectiveness. One of the significant benefits of FEM analysis is the ability to cut costs by lowering the number of destructive tests. Also, this model can be used to conduct parametric studies to assess the impact of a range of factors (Marin et al., 2022). In the packaging industry, FEM analysis has been applied for experimental and numerical verification of 3D modeling (Hagman et al., 2017), brim foaming simulation of cups(Upadhyaya & Nygårds, 2017), deformation of cigarette packages (Gustafsson & Nygårds, 2017), drop testing (Nygårds et al.,

2019), and a gable top package's box compression simulation(Fadiji et al., 2017; Marin et al., 2021). Also, FEM simulation and verification are beneficial for the physical safety assessment of packaging for dangerous products. Marin et al. (Lengas et al., 2022) proposed a method for data collection of drop tests and FEM based model for the simulation of impact loading as a mechanical failure calculator. Figure 2.7 illustrates the maximum stress expressed in MPa applied on a paperboard in a FEM model (A) and a physical experiment (B) and shows the agreement of these methods in figure 2.7 (C).



Figure 2.7 (A) Comparison between FE simulation and (B) physical experiment and (C)Comparison of force compression curves for the experiments and the two FE models for a paperboard(Marin et al., 2022)

Fadiji et al. (2017) investigated the mechanical properties of paperboard packaging material utilized in packaging fresh agricultural products. In this research study, the corrugated paper board's edge compression test was modeled using FEM analysis, and its results were verified with experimental findings. Bahlau et al. (2022) proposed a biodegradable packaging structure to reduce the maximum stress and, as a result, material usage reduction. This study used FEM modeling for the topology optimization process (Bahlau & Lee, 2022) All these analyses have a simple material model for predictions, making this method timesaving and increasing prediction accuracy.

While Field tests, libratory tests, and numerical assessments have benefits in different conditions, there are some drawbacks for each of them.

Real tests are often time-consuming, expensive, and non-repeatable. Also, they have limited resources. Although laboratory tests are efficient and cost-saving, they do not replicate the wide variety of transient events occurring during transportation. That is the reason for developing a method capable of replicating actual transient events in a reproductive way in laboratories (Hagman et al., 2017). Additionally, laboratory tests are limited to real-life conditions simulation.

Concerning FEM analysis-based evaluation methods, if the material model is complex and there are geometrical imperfections, the prediction of FEM evaluation will not have acceptable results.

To overcome these limitations, a faster method is required. Recently researchers employed artificial intelligence-based methods such as machine learning (ML) techniques to evaluate packaging properties without experimental tests. Machine learning models offer packaging engineer analytics and predictive tools, technical parameter optimization, packaging digitization,

13

and packaging evaluation opportunities. The following sections focused on artificial intelligent approaches introduction and their applications in general and the packaging industry.

2.3 Introduction of Artificial intelligence

The idea that machines may show human intelligence was the basic premise behind the term "artificial intelligence (AI)," which was first used in the 1950s(Bini, 2018). AI is a subfield of computer science that studies how intelligent behavior may be simulated in computers and how machines can mimic human cognitive abilities(Mathews, 2019). Several scientists from a variety of areas, such as mathematics, psychology, engineering, economics, and political sciences, started debating the potential of developing an artificial brain in the 1940s and 1950s. In 1956, the academic discipline of the study of artificial intelligence was established(Kaplan, 2022). AI has evolved from a theoretical concept to a practical application on a massive scale in the current era of rapid technological innovation and exponential increases in extraordinarily big datasets, also known as "big data"(Helm et al., 2020). As a result, it is predicted that AI-powered machines will be able to carry out any intellectual work that a human can by the year 2050(Mathews, 2019). AI has several subsets, including deep learning, machine learning, natural language processing, robotics, and neural networks. There has been a lot of interest in how to find and interpret mixed data and information.

In 2006, Hinton put forth the idea of deep learning. The most widely used kind of AI is machine learning (ML), and deep learning is the most well-known subset of ML. There has been a lot of interest in how to find and interpret mixed data and information. Researchers now have a greater perspective of understanding the world due to deep learning. Through the use of massive volumes of training data, deep learning approaches simulate the large-scale neural network of brain architecture and create a complex multi-layer Artificial neural network (ANN). Humans finally

learn several levels of abstraction using multi-level learning of the data to support additional processing strongly. Deep learning aims to increase classification or prediction accuracy by learning more useful features by creating ML models with many hidden layers and a lot of training data.

Machine learning (ML) is an area of computer science and a subset of artificial intelligence (AI) that focuses on using big data and algorithms to simulate human learning processes, progressively increasing the accuracy of those algorithms. Today, data science increasingly depends on machine learning, called the learning problem. Tom Mitchel (Mitchell, 2010) defined a learning problem as a "computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."

A general schema for ML model steps introduced by UC Berkeley is shown in Figure 2.8. Algorithms are trained to discover patterns in training data to find significant insights or predictions using statistical approaches. Using Learning models results in highly accurate predictions.



Figure 2.8 Machine Learning steps

A machine learning algorithm's learning system is of three major components:

1. A decision process: The application of the ML models is to generate predictions and classifications. These algorithms can estimate a pattern in the input data based on unlabeled or labeled data.

- An error function: The model's accuracy is evaluated using an error function. An error function can assess the model's accuracy by comparing it to known examples if there are any.
- 3. A model optimization process: The model's performance can be better based on training data if the value of hyperparameters is changed. The algorithm can select different combinations of values through an "evaluate and optimize" process until an acceptable accuracy level is reached.

The machine learning techniques can be divided into four main groups(IBM Cloud Education, 2021) based on what kind of experience they are allowed to learn:

- 4. Supervised learning
- 5. Unsupervised learning
- 6. Semi-supervised learning
- 7. Reinforcement learning

In supervised machine learning, we are provided with labeled data and believe there is a relationship between input and output. Under supervised machine learning, there are two main categories of problems, classification and regression. Classification problems use a model to classify test data into distinct classes accurately. An example of a classification problem in the real world is filtering spam emails in a separate folder in your email box. Algorithms used for classification frequently include decision trees, support vector machines, Naïve Baye–, K - Nearest Neighbor (KNN), linear classifiers, and random forests. In regression problems, the model predicts a continuous target variable. Polynomial regression, logistic regression, and linear regression are some common regression methods.

Unsupervised machine learning algorithms automatically recognize useful patterns in unlabeled data in contrast to supervised learning. Massive volumes of unlabeled data are now being produced, which makes unsupervised learning necessary. This learning model is useful for data analysis, cross-selling techniques, consumer segmentation, market basket analysis, web mining, social network analysis, recommender system, and picture and pattern recognition since it can find similarities and differences in information. Three main categories of unsupervised learning problems are clustering, association, and dimension reduction (Ghahramani, 2004). The k-means algorithm is one of the common models of clustering algorithms. Also, Apriori and Frequent Pattern Growth (Fp-Growth) algorithms are popular association problem algorithms. For dimension reduction, the PCA algorithm is a well-known method.

Semi-supervised learning has a data set combining labeled and unlabeled data. This combination helps models categorize unlabeled data. This method is particularly useful when relevant feature extraction from data is challenging. A popular semi-supervised method that uses a small set of labeled data is general adversarial networks (GAN)(IBM Cloud Education, 2021).

Reinforcement learning is the same as supervised learning, but it is not trained by training data. This model trains by trial and error without any guidance from the human agent (Goodfellow et al., 2016).

The most popular machine learning algorithms based on their functions and applications will be covered in the following paragraphs(Goodfellow et al., 2016):

<u>K-means clustering</u>: K-means is known as an **unsupervised** learning algorithm. The algorithm name includes the letter "K," which stands for the number of clusters we want to divide our unlabeled data into. as a result, each data instance relates to a specific group with related attributes. Clustering algorithms can find patterns in data to group it via unsupervised learning. In

this algorithm, the input is not labeled and uses gained experience during problem-solving. Clustering is frequently used by retail businesses to find communities of homes that are similar. Also, this algorithm is used in several facility location optimization problems. An optimal placement method was proposed for locating data collectors in a smart grid by using the K-means algorithm (Tavasoli et al., 2016).

Apriori and Fp-growth algorithms: Apriori and Fp-Growth are recognized as **unsupervised** learning techniques because they are frequently applied to uncover or explore interesting patterns and relationships. The Apriori algorithm is developed to operate on databases that contain transactions and construct association rules using item sets. It indicates how strongly or weakly two items are associated using these association rules. This algorithm quickly calculates the itemset associations using a breadth-first search and a hash tree. Since finding the frequent item sets from a huge dataset involves an iterative procedure using Apriori, Fp-growth was proposed. The FP-growth Algorithm is a different approach to finding frequent item sets without requiring candidate generations. It employs a divide-and-conquer technique and a Frequent Pattern Tree to find the most common pattern. To determine how frequently a certain feature appears throughout the data set, this machine learning system first identifies that characteristic.

<u>Linear regression:</u> This algorithm is a **supervised** machine learning algorithm. It uses a linear relationship between several values to predict numerical numbers. It mostly focuses on determining the variable's relationship and prediction. For instance, the method might be applied to forecast housing prices based on local historical data.

Logistic regression: This approach generates predictions for target variable variables, such as yes/no" responses to questions. It is a supervised classification algorithm that calculates the probability value based on a statistical logistic model. Using a dependent variable proceeds through a predictive procedure and then discovers the relationship between the independent variable, say x, and the dependent variable, say f (x). Figure 2.9 displays the values of the logistic regression function f (x) concerning various x values (Dadhich & Thankachan, 2022). It can be applied to tasks like categorizing spam and quality assurance on a manufacturing line.



Figure 2.9 Logistic regression function(Dadhich & Thankachan, 2022)

<u>Neural networks:</u> Neural networks learn as **supervised** learning models. They resemble the human brain functions by having a huge network of interconnected processing nodes. Neural networks are successful in detecting patterns and are useful in applications such as speech recognition, image creation, natural language translation, and image recognition. Figure 2.10 shows the input and output for a convolutional neural network we implemented for packaging image clustering.



Figure 2.10 An example of an artificial neural network application for packaging images clustering

One of the branches of neural network algorithms is the long short-term memory (LSTM) classifier. Neural networks that can learn order dependency in sequence prediction tasks are known as long short-term memory networks. Several tasks, including sentiment analysis, question answering, machine translation, word embedding, and name entity recognition in natural language processing (NLP), as well as image classification, object detection, image segmentation, image generation, and unsupervised feature learning in computer vision, have shown significant improvement from deep learning-based models(Minaee et al., 2019). Many of the traditional methods for tackling various computer vision and natural language processing applications have been enhanced by deep learning models. These methods use end-to-end models to concurrently learn the extracted features extraction and class prediction, as opposed to manually extracting features from text and images and delivering them to a classification model.

Various publications have utilized deep learning techniques for sentiment analysis in recent years. Zhang et al. (2018) published a detailed literature review on the application of deep learning to sentiment analysis. They introduce word embedding methods that deep learning requires. Using a mathematical procedure, words in a lexicon are transformed into vectors of continuous real numbers. As a result, word embedding offers the deep learning model a wealth of new information, resulting in more accurate sentiment analysis (L. Zhang et al., 2018). A sentiment analysis technique based on deep convolutional networks from tweets and short words was developed by Santos et al. Minaee et al. (2019) improved sentiment analysis by proposing an ensemble model to outperform the performance of text analyzing (Minaee et al., 2019)

Decision trees: This model falls under the group of **supervised** learning algorithms. Decision trees can be used to categorize data into categories and regression problems. A tree diagram can be used to show the branching sequence of connected decisions seen in decision trees. In contrast to the neural network's black box, decision trees are simple to validate and verify, which is one of their advantages. Figure 2.11 shows a section of the binary decision tree, including nodes that show a specific question is asked which has a yes or no answer. These questions are based on the value of weight w, length L, and height H of several boxes in the training dataset. Leaves show the predicted compression value based on w, L, and H values.



Figure 2.11 Decision tree for compression strength values

<u>Random forests:</u> These algorithms are supervised learning models used for classification and regression problems. Based on data samples, the random forest algorithm builds decision trees, obtains predictions from each one, and then uses voting to determine the optimal option. It has more accuracy because it averages the results in contrast to a single decision tree

<u>The K-nearest neighbors (KNN)</u>: The KNN classifier is a supervised machine learning approach that can solve classification and regression modeling problems. But in the business world, classification prediction challenges are where it shines. The KNN method predicts the values of new data instances based on "feature similarity," which further indicates that the new data point will be given a value depending on how closely it fits the points in the training set. Figure 2.12 shows two classes in blue and green colors. The inner circle has four neighborhoods, and the outer circle contains bibe items, so their K is four and nine, respectively. When the model receives unknown data, a question mark at the center circles, it calculates its distance from its neighbors.



Figure 2.12 classification with KNN classifier(Dadhich & Thankachan, 2022)

Support Vector Machine: SVM is a supervised machine learning approach used for classification and regression. The SVM algorithm aims to find a subset of features in an N-

dimensional space that clearly classifies the input points. The hyperplane's dimension is affected by the number of features.

<u>Naïve Bayes (NB)</u>: One of the quickest, easiest, and most efficient supervised classification algorithms currently in Naïve Bayes. It is based on the Bayes Theorem and probability and makes the assumption that predictors are independent. The Naïve Bayes classifier assumes that a feature's inclusion in a class has nothing to do with any other features. It assigns the instance probabilities for a specific instance to be classified. For example, Figure 2.13 shows an NB algorithm with three predefined classes. Any unknown data will be sorted into one of these three categories if it is received. This approach is for categorizing problems with binary and multiple classes.



Figure 2.13 Three categories in the form of green triangles, red squares, and blue circles(Dadhich & Thankachan, 2022)

Natural language processing (NLP) is one of the AI technologies that makes it more userfriendly and significantly impacts real-world AI applications. Machines can now understand human language thanks to a field of artificial intelligence (AI) called natural language processing (NLP). NLP techniques date back to the 1950s when computer science's branch on artificial intelligence initially attracted interest. The Georgetown experiment from 1954 talked about the vast, automatic text translation between languages and how it might expand. Early NLP methods relied on grammatical and heuristic stemming rules and rule-based models. Since then, quick developments have boosted NLP to leadership as a driver of AI development, particularly in the twenty-first century (Shankar & Parsana, 2022). Its purpose is to develop computer programs that interpret language and undertake automatic tasks, including topic classification, translation, and spell-checking. There are various branches of NLP: speech processing, natural language understanding, natural language generation, knowledge base building, dialogue management systems, sentiment analysis, text mining, text analytics, and other related concepts(Moreno & Redondo, 2016). At the most fundamental level, sentiment analysis approaches include categorization-based and user-based techniques.

Lexicon and learning paradigms form the categorization approach. Dictionary-based, huge corpus-based, semantic-based, and statistically-based algorithms are examples of lexicon-based strategies. In the dictionary-based lexicon technique, a set of control words with well-known attitudes is repeatedly expanded by adding related words. AFINN, NRC, and Bing are four sentiment lexicons with identified word meanings that are frequently utilized in the literature. According to the AFINN lexicon, words are given ratings ranging from -5 to 5, with lower scores denoting negative sentiment and higher ones indicating positive sentiment (Z. Zhang, 2018). The NRC lexicon divides words into positive and negative attitudes as well as eight different emotions, including anger, disgust, fear, anticipation, joy, trust, sadness, and surprise(Mohammad, 2017). The Bing Lexicon classifies words into positive and negative classes in a binary approach(B. Liu, 2020). Table 2.1 shows examples of each lexicon dictionary for different scores. In this study, we utilized a rule-based algorithm based on the AFINN lexicon for sentiment analysis (more explanations are provided in section 3.2.2).

Lexicon	Word	Score
AFINN	outstanding	5
	abandon	-2
	abhors	-3
NRC	abandon	fear
	abandoned	anger
	abandoned	sadness
Bing	abnormal	negative
	break	negative
	afford	positive
	agreeableness	positive

Table 2.1: Examples of scores for specific words for AFIN, NRC, and Bing lexicon

The review words with context-specific attitudes are discovered using the corpus-based lexicon approach. The sentiments of aspects, phrases, and texts are determined using the extracted words and sentences via a machine-learning technique called the semantic lexicon method. The statistical-based lexicon method is a machine-learning technique that analyzes extracted semantic features to identify the attitudes of aspects, sentences, and texts(Dadhich & Thankachan, 2022).

Learning paradigms include supervised, unsupervised, and semi-supervised learning techniques. The supervised learning method classes the unknown data into predetermined categories after training the known dataset. The most frequently supervised algorithms that are used for sentiment analysis are SVM, NB, KNN, and LSTM. The unsupervised learning approach categorizes the reviews without using any prior knowledge, which is the reverse of the supervised learning method. The integration of supervised and unsupervised learning methods is known as semi-supervised learning.

A supervised learning technique called Naïve Bayes (NB) identifies the distinct independence between the obtained characteristics. It also calculates the variance of the features using this property. It is a likelihood classifier that applies the Bayes theorem and probability

principles. Using these ideas, it first trained the known reviews before predicting the sentiment of the unexpected ones. A branch of ANN-based mode for sentiment analysis is the long short-term memory (LSTM) classifier

2.4 General artificial intelligence applications

In this section, we give a few general AI application examples that you might see in real life:

Speech recognition: Speech recognition is a common feature in mobile devices that allows voice search and messaging accessibility (like Siri). In artificial intelligence science, speech recognition is a structured sequence classification problem. It is a technology that converts spoken words into a word sequence. It is also called automatic speech recognition, computer voice recognition, or speech-to-text. Deng et al. (Deng & Li, 2013) provided a variety of ASR approaches into a set of well-known ML algorithms such as SVM and Bayes learning algorithms.

<u>Customer service</u>: Customers are at the core of any business's success, which is why companies are more concerned with the need to obtain client satisfaction. Sabah's study ((Sabbeh, 2018)) established a benchmark for the most frequent ML algorithms (Random Forest and Adaboost Learning) for customer churn classification. Moreover, Robert et al. (Ireland & Liu, 2018) provided invaluable insights into product designs based on customers over E-commerce. They used Naïve Bayes learning and Apriori algorithms to find customers' opinions towards a product. Along the customer journey, online chatbots are taking the place of human customer service representatives, which has altered how we view consumer involvement on websites and social media. Chatbots provide tips and advice, cross-sell products, and provide sizing recommendations for consumers. They respond to frequently asked questions concerning subjects like shipping. Virtual agents on e-commerce websites, messaging bots for Slack and Facebook

Messenger, and duties typically carried out by virtual assistants and voice assistants are a few examples.

<u>Computer vision:</u> This artificial intelligence technology allows computers to extract meaningful information from digital photos, movies, and other visual inputs and make smart decisions. This research area, powered by machine learning, such as convolutional neural networks and SVM, is used for self-driving cars in the automotive industry, radiological imaging in healthcare, and photo tagging on social media. Also, Khan et al. (Khan & Al-Habsi, 2020) categorized machine learning approaches for computer vision based on application areas such as the food industry, environment, health, and maintenance issues.

<u>Recommendation engines:</u> AI algorithms can assist in finding data trends that can be employed to create more effective cross-selling strategies by using historical consumption behavior data.

Online businesses employ this strategy to present clients with important and relevant product recommendations throughout the purchase process. Akbar et al. (Akbar et al. 2022) reviewed two machine learning algorithms, such as neural networks and decision-based algorithms for recommendation engines.

<u>Fraud detection:</u> Machine learning can be used by banks and other financial institutions to identify unauthorized transactions. A model can be trained using data regarding recognized fraudulent transactions through supervised learning. Transactions deserving of additional inquiry can be found via anomaly detection. Dornadula et al. research (Dornadula & Geetha, 2019) examined the key supervised and semi-supervised learning categories and their fraud detection problems.

27

2.5 Artificial intelligence applications in the packaging industry

Artificial intelligence has established itself as a verified approach to guide the industry's next evolution from manufacturing to packaging and distribution. The packaging industry's implementation of AI methods is mainly driven by the rise in demand for environmentally friendly packaging, consumer goods, the circular economy, and reduced packaging and product damage. This section investigates artificial intelligence applications in the packaging industry and organizes them according to the goals of the presented methods.

Packaging planning: Since product packaging design can have a strong and direct impact on consumers' attitudes toward a product, effective packaging planning is required to determine whether the package succeeds or fails. Knoll et al. (Knoll et al., 2019) proposed a two-step machine learning model for an automated packaging planning process to select suitable packaging based on product features. This research used a supervised method, a Random Forest (RF) Learning classifier, and a regression model to identify relevant features and the fill rate of parts per packaging. This model received features such as dimensions, shape, weight, quality, price, and dangerous goods and classified parts based on packaging categories, including small load carrier, large load carrier, heavy load, and special loads. Then by using the classification method, feature importance was calculated. Based on RF's result, the part price of the product with high importance was selected as a quality control indicator. Then they used regression classification for packaging planning.

Silva et al. (da Silva e Silva et al., 2021) research project applied the KNN algorithm to categorize and select biodegradable active packaging made from fish gelatin with the addition of palm oils and essential oils of clove and oregano. The KNN model received features of different films, including Solubility, Thickness, Water activity, Humidity, Tensile strength, Elongation, and

Water vapor permeability. Then, based on the received input, selected the most reliable packaging method with the highest antioxidant activity, tensile strength, and elongation values. Wei et al. proposed an evaluation method for optimal packaging selection for different food based on unsupervised machine learning algorithms. In this research, packaging characteristics were extracted using fuzzy directivity classification. Then a decision tree algorithm was trained by the best food characteristics data to assign the best food packaging to specific foods(Tian & Song, 2020).

Packaging delivery optimization: Packaging delivery guarantees that the products arrive at their destination in a satisfactory quality, saving both time and cost and improving sustainability. A machine learning-based sustainable delivery approach was proposed for the green logistics location routing problem. This method used gaussian mixture clustering to improve simultaneous location-routing decision-making (Y. Wang et al., 2020) . One of the most important events in delivery systems is determining an optimal delivery path. Several recent research has attempted to tackle the problems of truck-based routing, drone-based routing, and truck /drone combination routing for delivery purposes. From the drone-truck combination delivery aspect, Chang et al. (Chang & Lee, 2018) proposed an optimal delivery path method based on k-means and the Traveling Salesman Problem to discover a delivery path that reduces total delivery time for a truck/drone delivery system.

Figure 2.14 –(A) illustrates three steps of the optimal drone-truck delivery model, including clustering delivery locations, routing the centers of clusters, and finding shift weights.

Figure 2.14 –(B) shows Delivery locations grouped into K clusters using the K-means clustering model.



Figure 2.14 (A) Optimal delivery in three steps (B) delivery location clustering

Packaging defect detection: Product packaging defects are one of the serious determinants that affect sales and profit. Wu et al. (Wu & Lu, 2019) present a product packaging box damage detection technique. In this study, Support Vector Machine Learning (SVM) and image processing were used together to identify defects in the package. The predictor model was trained by 8100 product images labeled in six defect types of groups. The output was to detect flaws detection. For the evaluation of model performance, the models' results were compared with the real data. Also, Yang et al. (X. Yang et al., 2020) developed an image capture process system that detects logistics packaging box defects using the SVM learning model. The model can detect two common small and medium defects in logistics packaging boxes with higher accuracy and lower system costs.

An innovative package evaluation was proposed by Esfahanian et al. (Esfahanian & Lee, 2022) based on ML approaches and packaging review filtration. Firstly, they created a word list
called Pack-List based on the most frequent words related to packaging in reviews with low ratings. For example, for mirror products, the Pack list is shown in Table 2.2.

Pack-List word for mirror							
Break Fail Scratch Damage							
Failure	Protect	Crack	Defect				
Protective	Deliver	Defective	Shatter				
Delivery	Distort	Fragile	Destroy				

Table 2.2: Pack-List word for mirror (Esfahanian & Lee, 2022)

Then by using this word list, all reviews related to packaging problems were extracted. By having packaging-related reviews, they calculated positive and negative packaging rates by utilizing the Naïve Bayes classification. Figure 2.15 (A) shows percentages of positive and negative reviews for three TV products, A, B, and C. For comparing. Therefore, they could calculate the packaging failure rate based on negative reviews for these three TVs. Their packaging failure rates over the years are shown in Figure 2.15 (B). And Finally, the most frequent product part and damages are shown in Figure 2.15 (C). Also, Sarah et al. proposed an intelligent packaging evaluation based on an artificial neural network (Holland et al., 2022). The proposed model analysis images uploaded by customers on the e-commerce platform to identify packaging failures.



(C) Figure 2.15 (A) Percentage of negative and positive reviews for three TV brands (B) Percent of Negative Reviews over time for three TV brands(Esfahanian & Lee, 2022)(C) Word clouds for

three televisions (TVs)

Packaging maintenance: Koca et al. proposed scheduled Maintenance prediction for unplanned downtime in accordance with packaging robots (Koca et al., 2020). This study used Multilayer Perceptron (MLP) Machine Learning model Neural Network trained with the failure data. Failure data attributes included a season of failure, machine number, failure area, operation time, downtime, production at failure, and conversion parameters. The trained model was able to predict upcoming failures and estimate machine number, failure area, and operation time for packaging robots. Their results showed that production efficiency increased, and unplanned production downtime costs immensely decreased.

Packaging design: Several researchers use machine learning algorithms for designing packages. From a graphic design aspect, Juárez-Varón et al. proposed a machine learning-based method for applying to neuromarketing analysis. They predicted consumer behavior regarding key elements of the packaging design of an educational toy by using supervised decision tree learning (Juárez-Varón et al., 2020). This study has identified the most characteristics of educational toy packaging, how consumers judge the value of a brand and a product based on packaging, and how packaging influences how much fun kids will have. This study has helped to reconsider the way that scientific literature on consumer behavior models and the packaging for educational toys is addressed.

<u>Chemical evaluation:</u> 3D printing packaging has opened up opportunities for innovative packaging designs. One of the challenges of the 3D printing industry is the thermal degradation of 3D-printed Polylactic acid (PLA) specimens. Zhang's research (Agrawal & Srikant, 1994) was a novel classification method for predicting the thermal degradation of heat-treated Polylactic Acid (PLA) materials. This research study covered three main topics. Firstly, five categories of polymeric degradation were introduced: thermal degradation, photodegradation, chemical

degradation, biological degradation, and mechanical degradation. Secondly, various testing methods for thermal degradation prediction were briefly described in a table grouped by degradation category and polymer type. Lastly, four machine learning models were applied to four datasets to classify degraded 3D printed PLA materials. These four classification methods included multi-class decision forest, multi-class decision Forest, multi-class logistic regression, and ANNs (Artificial neural networks) and trained on four Fourier Transform Infrared Spectroscopy Data sets. The results showed that the multi-class logistic regression and multi-class neural network models, which are trained by the first and fourth datasets, had a higher accuracy than other models. Zhang et al. conducted a comprehensive study that shows how Machine Learning solutions can be helpful in predicting the chemical durability of oxide glasses. They used two metrics to evaluate the chemical durability of glass, including weight loss and visual appearance change. The authors aimed to understand a specific aspect of the material that is difficult to probe by experiments. For the first task, weight loss prediction, they first collected a big dataset on the chemical durability of about 1400 oxide glass compositions under three chemical testing conditions, Acid-Base, and HF solutions. The following table shows the details of the dataset. Then They used the Random Forest (RF)model and Artificial Neural network (ANN) to predict weight loss based on compositions of oxide glasses. Table 2.3 Shows their accuracy. So, you can find out that the ANN model predicted weight loss more accurately than the RF model.

Then, they measured the correlation between weight loss and 13 different compositions, including SiO₂, B₂O₃, Al₂O₃, Na₂O, CaO, MgO, K₂O, P₂O₅,BaO, Li₂O, ZnO, SrO, and ZrO₂ by using the RF model. Their results showed that key features for weight loss prediction are SiO₂, Na₂O, and P₂O₅ for acid, basic, and HF solutions.

Models	Prediction accuracy in the Acid solution	Prediction accuracy in basic solution	Prediction accuracy in HF solution	
ANN	0.95	0.97	0.94	
RF	0.87	0.91	0.92	

Table 2.3: Accuracy comparison of ANN and Random Forest (RF) models

For the second task of their research study, they used these 13 oxides as data features for classifying the surface appearance change rating, which indicates surface damage resulting from exposure to different chemical solutions. Then, after model training, their result illustrates RF prediction is more accurate than the ANN model. Their accuracy is shown in the following table:

Table 2.4 : Accuracy comparison of ANN and Random Forest (RF) models with 13 oxides data

features

Models	Prediction accuracy	Prediction accuracy	Prediction accuracy				
withers	in the Acid solution	in basic solution	in HF solution				
ANN	0.96	0.88	0.86				
RF	0.91	0.83	0.80				

Sustainability aspect: The current COVID-19 outbreak, the e-commerce boom, and the increased usage of flexible, affordable, and easy-to-use packaging are the key factors driving waste management investment. Different techniques have been proposed for plastic recycling. However, plastic/ polymer clustering is one of the challenges in mechanical and chemical recycling due to the degradation of polymers in different conditions. High temperature or oxygen levels during mechanical recycling may lead to some polymers' degradation. So, the same plastic waste should be clustered in the same group. Yang et l. (Y. Yang et al., 2020) proposed a system with different steps for classifying plastic waste (transparent and white). The contributions of this study are data collection, utilizing NIR for information extraction, using the PCA method for features reduction, selecting the optimal machine learning model (Artificial Neural Network and SVM) for

transparent and white plastic classification, finding the best model, Artificial Neural Network for large volume spectral data classification having high accuracy.

CHAPTER 3 PROBLEM STATEMENT AND METHODOLOGY

3.1 Problem statement

Packaging development is different from product development. Though products' primary function is to meet customer requirements, packaging facilitates a product's potential to provide value to the supply chain. Since this role involves multiple packaging functions, the supply chain participants have various packaging requirements, making the packaging development decision-making process more complicated. Tools and procedures specifically designed for packaging development are required as generic product development tools and methods do not evaluate the key roles of packaging (Molina-Besch & Pålsson, 2020). Several mechanisms aim at the evaluation of all relevant requirements on the packaging. So, these approaches help companies evaluate and consider packaging alternatives throughout the latter stages of packaging development. These methods include field tests, laboratory tests, and numerical evaluations.

As discussed in the previous chapter, each category of the evaluation method has some limitations. Field tests are generally costly, time-consuming, and difficult to repeat. While laboratory testing is cost-effective and efficient compared to real tests, it cannot accurately reproduce the extensive range of events that may arise during transportation and distribution. Furthermore, laboratory tests have limits for actual modeling conditions. In terms of numerical evaluation, such as FEM analysis-based modeling, if the material model has complex properties, the prediction of FEM evaluation will not have acceptable results or will be the most difficult to interpret. Therefore, a quicker approach is necessary to get around these limits. Early studies used Machine Learning (ML) approaches to determine packaging characteristics without performing tests based on customer reviews published on online shopping platforms(Esfahanian & Lee, 2022).

The number of customer reviews that a product or service receives rises quickly, along with the growth of e-commerce and user-generated content. Customers can write personal experiences about a wide range of brands, products, or services on online shopping websites, blogs, and forums such as Amazon. These online product reviews offer a new crucial information source for product and packaging design and evaluation. Hence, opinion mining is invaluable to other customers and extremely important for e-commerce companies. This collected information is advantageous because it reveals what consumers want from a product, it's packaging and the potential design problems.

In the past, surveys, interviews, focus groups, etc., were used to manually gather and interpret consumer opinions on products and packaging. However, these manual approaches were generally conducted based on a small number of customers and required an experienced designer, making them costly and time-consuming. Therefore, several methods have been proposed for the opinion mining of customer reviews on online shopping websites.

Sentiment analysis is the process of computationally recognizing and classifying opinions stated in a written comment, especially to establish whether the customer has a positive, negative, or neutral attitude toward a given topic (such as product, package, etc.)(Ding et al., 2007; Medhat et al., 2014). Commercial enterprises utilize this approach to understand customers better, evaluate brand reputation, and detect sentiment in social data. This intelligent data analytics of customer feedback helps companies learn which packaging and container problems make customers happy or unhappy to address the issue immediately and redesign the packaging system.

As it was shown in the previous research, Esfahanian et al.(Esfahanian & Lee, 2022) proposed a novel packaging evaluation based on sentiment analysis of customers' feedback published on the Amazon website. The process steps of this approach are shown in Figure 3.1,

and the data flow was used in CSV files. The first innovative aspect of this study is the Pack List (packaging word lists) being collected to detect the packaging-related reviews accurately. After web scrapping of customer reviews on Amazon, extracted reviews were saved in an offline dataset. Those concerning packaging issues were filtered using this Pack List. Then, sentiment analysis (SA) was applied to determine whether customer reviews had a positive or negative sentiment score. The SA calculation results were then saved in a local file, and the packaging failure percentage was compared through different months and years to determine how time impacts packaging failure. Moreover, a word cloud of negative sentences is then made to display the most frequent problems.

Although this technique offers a novel meticulous evaluation method to identify concerns by looking at customer reviews on defective packaging and helps with defect detection at early stages, it has its drawbacks. First of all, disconnected data follow and manual data analysis processes lead to many back-and-forth delays, making the processing time difficult to change. Although the text contains valuable insights, the method neglects informative uploaded images by customers, which come with reviews. Hence, considering both text and graphics in package evaluation would be more effective and accurate.



Figure 3.1 Manual process of systematic evaluation method

Figure 3.2 shows a word cloud on the left side. Word cloud is a visualization tool to show the frequency of words in reviews. Although packaging failure and word cloud may identify bold concerns, they can not show the relationship between the damage and where it occurred. Hence, the packaging failure rate is not enough to get damaged parts. Another area of improvement for manual packaging evaluation could be association rule mining of the most frequent problems and why and how they occurred. This study aims to use customers reviews to satisfy the following objectives:

Objective 1: To automate the data flow of packaging evaluation

Objective 2: To Correlate images and text of online product reviews

Objective 3: To determine relationships between the most frequent words in customer reviews to predict damages and their causes and effects



Figure 3.2 Drawback of a word cloud for defect identification

3.2 Model development for Web-based Intelligent Packaging Evaluation (WIPE) platform

In this section, we present our suggested web-based automated intelligent approach for identifying packaging failures and their connections using sentiment analysis of customer evaluations. We break down the model into four main modules:

- Web scrapper
- Sentiment analysis module
- Association rule mining
- Data analysis dashboard

Each module is presented in relation to the main contributions of this research, which include data process automation, automated embedded image downloading, and automatic association rule mining.

3.2.1 Web scrapper

One of the leading e-commerce sites in the world, Amazon.com, was selected as the study's data source for online reviews. Amazon.com has many products and millions of online product reviews with helpful information such as reviewer name, credibility, rating, date and time, helpfulness, and the ability to edit reviews, as shown in Figure *3.3* (Y. Liu et al., 2013).

There are a number of free online data sets related to Amazon reviews. Ni et al.(Ni et al., 2019) built the most recent dataset with high quality, called Ref2Seq, which included personalized and relevant justifications of reviews. Also, Amazon (Amazon, 2019) provided several review data collections, such as "The Multilingual Amazon Reviews Corpus" and "Helpful Sentences from Reviews," that offer includes the review's connection to a verified purchase. This still lacks a lot of information, such as the reviewer's name and rating. Although

these data sets contain a wide range of products and categories, it lacks some of the new review quality features established by Amazon, such as verified purchases. Also, it is outdated (Njunge et al., 2022).

In this section, we present how an embedded web scraper was implemented inside the packaging evaluation application and how it allowed all modules to have a connected flow of data.

Recently, open-source web content scraping from the internet has become more practical. Data extraction from websites is a process called web scraping. We programmed a web scraper using JavaScript in a visual code environment. A few web scrapping services are for purchasing and offer functionality comparable to the developed module in this research. The actual data pipeline, however, is inaccessible, making it impossible to see how the data is being converted to structured data. The web scraper procedure is able to find reviews related to packaging concerns that are expressed by customers. Also, it can extract all required data, such as the reviewer's name, rating, review page number, and images that were not assessable in previous works. This information can also be kept in a database or another storage system for analysis. While manual data collection from webpages is possible, web scraping often refers to an automated procedure.

This web scraper was embedded in our program. Whenever it receives an Amazon link for a specific product, it returns all reviews' text and images related to that product.



Figure 3.3 A representative of an online product review

For web programming, several libraries were used. The main libraries are Express, Cheerio, and Axios. Express is a node JavaScript framework for web applications that offers various functionalities, such as creating Single-page, multi-page, and hybrid web apps. A layer was added on top of Node js to help manage servers and routes. Axios's library was used to extract the HTML script of web pages with a specific pattern. Then, to parse the generated HTML and retrieve the reviews, we utilized the Cheerio library we had installed. Finally, the extracted data were filtered based on packaging word lists, and related reviews were transformed into structured data types and sent to the sentiment analysis module. The main advantage of this embedded web scrapping is to make an automated data flow through the application to save time and increase the speed of the application. A visual comparison of the disconnected and connected flow of data through manual and automated methods is shown in Figure 3.4.



Figure 3.4 Connected data flow in comparison to disconnected data flow

The second contribution of this study is to extend web scrapping capability by downloading customer-uploaded images and text. The HTTPS library was used for getting images and their attributes, like URLs. The embedded web scrapping flow is shown in Figure 3.5.



Figure 3.5 Embedded web scrapping steps

As Figure 3.6 shows, this application converted unstructured data to structured data. This means that we could extract and calculate the following features from the product web page: Identification number of reviews, links of images, most frequent word sets, packaging-related

sentences, the whole text of the review, sentiment analysis score, customer name, publishing date and review's images in separate folders for each review as shown in Figure 3.6.

asin	link	pageNuml thumbnai	images	itemsets	senList	pkgList	reviewTex	nlpAnalysi custome		-		
B085V	5PP https://w	v 59 ["https://i	["https://images-na.ssl-images-amazon.com/images/I/717f8UZJ7sL.jpg"]	0	[["found",	["damage	When I op	-0.42857 IhateAm	1.45		100	
B085V	5PP https://w	v 65 ["https://i	["https://images-na.ssl-images-amazon.com/images/I/718qFhJTWWL.jpg"]	0	[["package	["package	This is the	-0.5 3567744	1			
B085V	5PP https://w	63 ["https://i	["https://images-na.ssl-images-amazon.com/images/I/71PcQ8IK57L.jpg","https://images	-0	[["item","o	["damage	' I love tide	-1.5 2excolleg		1.14		
B085V	5PP https://w	67 ["https://i	["https://images-na.ssl-images-amazon.com/images/I/71Y-sUxipaL.jpg"]	0	[["package	["damage	As seen fre	-0.42857 Reshma				
B085V	5PP https://w	v 58 ["https://i	["https://images-na.ssl-images-amazon.com/images/I/81JOejkG3nL.jpg", "https://images	-0	[["lid","bre	["break"]	The value	-0.33333 Heatherl	9-8	14-9	22-8	1
B085V	5PP https://w	v 58 ["https://i	["https://images-na.ssl-images-amazon.com/images/I/81PH+IP4VaL.jpg"]	0	0	["package	Received p	0 Randall				_
B085V	5PP https://w	v 29 ["https://i	["https://images-na.ssl-images-amazon.com/images/I/919n1fn+mLL.jpg"]	0	[["issue","	["damage	The only is	-0.5 Kindle Ci	A DECEMBER OF			
B085V	5PP https://w	v 20 ["https://r	["https://m.media-amazon.com/images/I/61oBhsyWKDL.jpg","https://m.media-amazon.	0	[["try","pr	["damage	Nothing m	-0.5 Tugcebu				
B085V	5PP https://w	63 ["https://r	["https://m.media-amazon.com/images/I/71bsBJfaKmL.jpg","https://m.media-amazon.co	0	[["apparer	["package	Apparenth	-0.4 Bao Hu	1 100		A attain	
B085V	5PP https://w	v 61 ["https://r	["https://m.media-amazon.com/images/I/71K2mEEldaL.jpg"]	0	[["duct","c	["damage	Product w	-0.5 james j h				
8085V	5PP https://w	31 ["https://r	["https://m.media-amazon.com/images/I/81PZ-z+GPEL.ipg"]	0	[["last","co	["damage	As far as ti	-0.54167 kim	54-10	56-10	59-5	
									54 10	50 10	55 5	

Figure 3.6 Relevant images' features extracted from a web page and saved in separate folders based on review ID and page number

In the following section, we demonstrate how to extract computational features such as sentiment analysis score, frequent packaging word sets, and the relationship between each packaging word from structured data.

3.2.2 Sentiment analysis design

This research's main objectives were to explore the sentiments expressed in Amazon product reviews from the packaging damages aspect and find the "reason" and "causes" behind customer reviews. In this study, sentiment analysis was adopted as a methodology. We compared the accuracy rate of three NLP methods: LSTM, Lexicon-based, and Naïve Bayes models, to calculate reviews' sentiment analysis. Table 3.1: shows the accuracy of these three methods for negative and positive reviews by comparing them with the actual attitudes of specific reviews. Based on the results, the LSTM model accurately predicts customers' attitudes in comparison to NB and lexicon methods because the LSTM model could identify more negative reviews. The accuracy of the lexicon-based model was higher than the NB model; however, it had lots of neutral reviews. The reason was that the model could not find some words in the lexicon, so it estimated the sentiment as neutral. Based on this comparison, we used the LSTM method to calculate the sentiment behind online reviews, which helped us identify packaging damages.

Method	Accuracy %		
LSTM sentiment analysis	86%		
Lexicon-based sentiment analysis	66%		
Naïve Bayes sentiment analysis	62%		

Table 3.1: LSTM, Lexicon-based, and Naïve Bayes sentiment analysis methods accuracies

To reveal the "reasons" and "causes" of damages stated in reviews, we used association rule mining which extracted the relationship between relevant concerns. This method is covered in the next section.

NLP is interested in how computers and human (natural) languages interact, particularly how to design computers to handle and analyze significant amounts of natural language data. There are several applications for NLP, including running a search engine, voice recognition to increase security, email filters, smart assistants, predictive text, data analysis, sentimental analysis, entity recognition, voice-based apps, and chatbots(Bahja, 2021). Bahja (Bahja, 2021) concluded that sentiment analysis is the main NLP application in E-commerce. Sentiment analysis of reviews can forecast product damage and defective packaging systems, allowing the packaging designer to make the best possible decisions at the right time. Several supervised learning techniques are used for sentiment analysis and categorization. The objective information from the facts and the subjective information from the consumer comments are both enhanced by machine learning. The reviews are then rated as either positive, negative, or neutral. Dadhich and Thankach provided a classification scheme(Dadhich & Thankachan, 2022)es (Dadhich & Thankachan, 2022).In this study, the WIPE application sentiment analysis process went through four main steps to achieve the best results outline, which is shown in Figure 3.7.



Figure 3.7 Sentiment analysis framework

Step 1: the TF-IDF machine learning technique is advantageous in assessing the relative value of words in the text. We used reviews with a rating lower than two stars, and then packaging relevant word lists were generated via the Term Frequency-Inverse Document Frequency (TF-

IDF) algorithm. It gives a term more weight when it appears more frequently in a single document and often occurs in a wide number of documents. In order to do this, two metrics must be multiplied: the number of times a word appears in a document, called **Term frequency (TF)**, and the word's inverse document frequency across a group of documents, called **Inverse document frequency (IDF)**(H. Wu et al., 2008). The simplest method of computation of a word's frequency is the number of times it appears in a document. The frequency is then adjusted based on the document's length. By computing the logarithm of the total number of documents divided by the number of documents containing a keyword, the inverse frequency of a word is determined. So TF-IDF is calculated via Eq.(3.1):

$$TF - IDF = TF \times IDF \tag{3.1}$$

Where *TF* and *IDF* are calculated by following formulas:

$$TF = \frac{f_i}{f_r} \tag{3.2}$$

Where f_i is the number of times the term "*i*" appears in a review, f_r is the total number of terms in every review

$$IDF = Log\left(\frac{N}{N_i}\right) \tag{3.3}$$

Where N is the total number of reviews N_i is the number of reviews, including the term "i"

While one of the disadvantages of TF-IDF is that when most evaluations discuss the same problem, TF-IDF can sometimes ignore the significance of identical terms, it has some applications such as information retrieval and keyword extraction.

Step 2: In the next step sentence the whole review is divided into separate sentences Packaging relevant sentences are filtered, and finally, the data pre-processing process is applied to packaging sentences. An example of a review separation is shown in

Figure 3.8. the reason that we go with sentiment separation is that the majority of applications might find document-level sentiment classification to be too basic. A common presumption made by researchers about sentence-level analysis is that a sentence often only contains one opinion, but a document generally comprises a variety of viewpoints(B. Liu, 2012).



Figure 3.8 Sentence breakdown of a review and retrieved packaging-related sentences

Since the raw data review text received from a customer is frequently noisy and probably contains numerous mistakes, it must be transformed into a format that our sentiment analysis model can interpret and use. Data preparation is the process for this purpose shown in Figure 3.9.



Figure 3.9 Automated data preprocessing for each sentence

Converting contractions into the standard lexicon: Firstly, we transform contractions (such as I'm) to their standard vocabulary to keep the uniform structure in our text data (such as I am).

Converting uppercase letters to lowercase: Another problem for sentiment analysis is capitalization. For instance, if the pronoun "us" were written in capital letters in the review, it would be assumed as the country "USA." So, we convert all letters to lowercase to make all words in a uniform format

Removing non-alphabetical and special characters: Special characters and numerical tokens will be removed because they don't affect sentiment, enhancing the precision with which we can categorize the user's sentiment.

Tokenization: Tokenization breaks the raw text into meaningful units. In our project, units are words as a token of a sentence. For example, the sentence "box get damaged" can be tokenized as 'box',' get' and 'damaged' tokens. By examining the word order in the text, tokenization aids in comprehending the text's meaning.

Correcting misspelled words: Typographical errors are very likely to appear in customer reviews because actual customers write them. Therefore, in this phase, we fix spelling mistakes.

For instance, if a customer accidentally types "damag" it will be changed to "damage."

Removing "stop words": Stop words include often-used words like "we," "he," "she," and "them". Most of the time, removing these terms boosts a model's performance without affecting the text's sentiment.

Step 3 Sentiment Analysis: As we mentioned at the beginning of this section, we compared three sentiment analyses: LSTM deep learning, Lexicon-based, and Naïve Bayes machine learning models, for sentiment analysis of negative sentences. The lexicon-based method assigned a sentiment score between minus five and plus five. The data can be classified into three groups positive, negative, and neutral categories. In this study, we used the Natural library(JavaScript sentiment analysis libraries) for sentiment analysis of reviews. The Natural Library's sentiment analysis method is a rule-based supervised method trained on a predefined AFINN vocabulary word list, which improves the accuracy of the sentiment analysis of reviews. Koto et al. provided a comparative study using four different datasets to show that the AFINN lexicon is currently one of the two most effective methods for sentiment analysis (Koto & Adriani, 2015). AFINN contains more than 3300 English words with polarity ratings between minus five (negative) and plus five (positive). For instance, the word "like" has a polarity of 2, while "damage" has a polarity of -3. The technique computes the total polarity of each word in a text and normalizes it with the length of a sentence. The equation for sentiment analysis score calculation of a sentence is shown in Eq. (3.4). If the algorithm returns a negative number, it indicates a negative sentiment; if it returns a positive number, it indicates a positive sentiment; if it returns 0, the sentiment shows a neutral feeling.

$$SA = \frac{\sum polarity \ rating \ of \ positive \ wi}{W}$$
(3.4)

Where SA is the sentiment analysis of a sentence, wi is a word of the sentence that is included in the AFINN dictionary, and W is the total number of words in the sentence.

If a review includes more than one sentence related to packaging, overall sentiment analysis will be the average of sentiment analysis of sentences, Eq. (3.5).

Overall sentiment analysis score
$$=\frac{\sum SA \ score \ s_i}{S}$$
 (3.5)

Where s_i is a sentence, SA is a sentiment analysis score, and S is the number of packaging-related sentences.

However, the rule-based model results in many neutral sentiments that were incorrect in many cases. So, the next model was implemented by Naïve Bayes, trained by reviews with ratings 1 and 2. After training, model evaluation was done on test data. The result showed the lowest accuracy of 60 percent. So, the third model, based on an artificial neural network, was used for sentiment analysis. This model was an LSTM sentiment package of Node.js, which was trained by a set of tweets. This open-source module develops LSTM networks and character and word-level embeddings to determine if a given text is "positive" or "negative.". The LSTM-based model is compact and sufficiently run everywhere, including web applications or portable platforms, because of the use of characters. Since WIPE is a web-based application, using character and word-level embedding didn't reduce the speed of the application and also increased the sentiment accuracy of sentiment analysis from 66% to 86% by finding more negative reviews.

Step 4 Data Analytics: calculated sentiment scores of reviews should be converted the results of the analysis into meaningful information: The final step is to convert the results of the analysis into meaningful information. The results will be displayed on Word cloud, bar charts, pie charts, and line charts. The data tool used in this project is the Chart library of node JavaScript.

For example, the word cloud is a way to show frequent words inside documents based on the TF-IDF parameter. To find the common problem, we can filter negative reviews and then apply the TF-IDF method to find the most frequent words in reviews. Figure 3.10 shows that the most common words repeated in reviews can be leak, package, damage, and box.



Figure 3.10 Frequent packaging words

3.2.3 Association rule mining design

In the previous chapter, we explained how the sentiment analysis of a review is calculated and what problems are common in negative reviews. However, the relationship between common problems was not discovered; for example, we could not find which sentence was relevant: **the box got damaged** vs. **the product got damaged**.

To achieve this research's third objective, we use association rule mining approaches to explore the relationship between these most common words to know their importance in mentioned problems.

Objective 3: To determine relationships between the most frequent words in customer reviews to predict damages and their causes and effects

Sentiment analysis can show whether users are dissatisfied with a product, but it cannot explain why. Despite the simplicity of categorizing findings into binary categories, positive and negative, a sentiment's applicability to engineers is reduced when it is detached from its context (Jin et al., 2016). Despite significant advances in sentiment analysis, sentiment does not offer engineers an entire context of what is causing the sentiment. For example, a review may express dissatisfaction with a "cap" feature but fails to specify what problem the cap causes. The discovery of "association rules" often describes the most frequent items seen in reviews and those items' relationship identification. We used association rule mining to identify where the most frequent damage happened or the relationships between frequent damages.

The association rule learning approach uses rules to extract meaningful relationships between variables in a large database. The association rule learning mechanism was initially used for market basket analysis to find how items purchased by customers are related(Kaur & Kang, 2016). An example of a market basket analysis problem shows in Figure 3.11. Assume we have five items available on store shelves. We want to uncover the association between the previously bought items to allow retailers to identify the relationship between more frequently bought items, Figure 3.11 (A), (B). The interesting relationship can be discovered by using association rule mining and represented in the form of association rules, as shown in Figure 3.11 (C). The association rule shows a strong association between bread and butter. It demonstrates that many people purchase milk and butter together. These rules can assist retailers in identifying the purchasing trends of their customers.



Figure 3.11 Association rule mining example in market basket analysis problem (C) Association rule: Bread→Butter

The concept of association rule mining can be divided into two parts in the following section.

- Data concepts
- Rules concept

We explain each of them based on transaction data like a set of reviews shown in Table 3.2.

Table 3.2 Transaction data example

Transactions	Items		
t_1	${A}{B}{C}{D}$		
t_2	${A}{B}{C}$		
t_3	$\{A\}\{B\}$		
t_4	${E}{D}{C}$		
t_5	${A}{D}{E}$		

Concerning data concepts, I is a set of all items/words which are defined as Eq.(3.6).

$$I = \{i_1, i_2, \dots, i_N\}$$
(3.6)

Where i is an item and N is the number of items. For example, for transaction data presented in, I

$$= \{A, B, C, D\}.$$

Transaction (*T*) is a set of items and $t \subseteq I$. Transaction Database (*T*) is a set of transactions as shown in Eq.(3.7):

$$T = \{t_1, t_2, \dots, t_n\}$$
(3.7)

A transaction *t* contains *X*, a set of items (itemset) in *I* if $X \subseteq t$.

An association rule is an implication of the form, as follows:

$$A \to B$$
, where $A, B \subset I$ (3.8)

Where A is the antecedent and B is the consequent. An antecedent is something that can be discovered in the data. A topic that is discovered along with the antecedent is called a consequent.

Association rules are constructed by looking for common if-then patterns in the data and utilizing the support and confidence criterion (defined in the following section) to identify the most crucial associations.

An **itemset** is a set of items. For example, $X = \{A, B\}$ is an itemset. A *k*-itemset is an itemset with *k* items. For example, $X = \{A, B\}$ is a 2-itemset

If we have two reviews, 'Screen got a crack' and 'Box damaged'. The data concept can be defined as shown in Figure 3.12.



Figure 3.12 Data concepts of association rule mining for packaging-related reviews

Association rule mining belongs to unsupervised automated learning fields. This learning rule mining method was used to extract key information from big databases. They are represented by the form $X \rightarrow Y$, where X is an item or itemset that indicates the antecedent and Y is an item or itemset referred to as the consequent. We can say that consequent items and antecedent items frequently co-occur. Hence, extracting relationships that appear to be hidden can be achieved using association rules. Support and confidence are the two parameters commonly used to assess the validity of association rules. We will see how these measurements can be defined.

Definition 1: Support is defined as the rule holds with support *sup* in *T* (the transaction data set) if *sup*% of transactions contain $X \cup Y$. Support *sup* is calculated using Eq. (

$$Sup = \text{Probability}(X \cup Y) = \frac{\text{No. of Transactions with } X \cup Y}{\text{Total no. of transactions}}$$
(3.9)

For example, considering reviews shown in Figure 3.13 for calculating the support parameter for Association Rule *X*: {box } \rightarrow *Y*: {damage}, we have :

• Sup {box, damage} =
$$\frac{2}{3} * 100 \cong 66\%$$

Sup {box, damage} shows how frequently a set is accrued. Its value means that {box, damage} were seen together in 2 out of 3 sentences.



Figure 3.13 Data concepts for three reviews

Definition 2: Confidence is shown as $Conf(X \rightarrow Y)$. An association rule $X \rightarrow Y$ is a pattern that states when X occurs, Y occurs with a certain probability called confidence. The rule holds in T with confidence *conf* if *conf*% of transactions that contain X also contain Y. It is calculated by Eq.(3.10):

$$Conf(X \rightarrow Y) = Probability (Y / X) \text{ or Probability } (X \rightarrow Y) = = \frac{Sup X \cup Y}{Sup X}$$
 (3.10)

For example, for association rule $X : \{ box \} \rightarrow Y : \{ damage \}$, Confidence is calculated as the following formula:

Confidence {damage|box} =
$$\frac{\frac{2}{3}}{\frac{2}{3}} * 100 \cong 100\%$$

This means that it is 100% likely that a {damage} will be seen if {box} was seen.

Utilizing Confidence in association rule mining is an effective way to raise awareness of data relationships. Its main advantage is that it highlights the relationship between specific items

within the set by comparing co-occurrences of items to the overall occurrence of the antecedent in the specified rule.

There are several algorithms for obtaining association rule mining. The most popular algorithms are Apriori, proposed by Agrawal and Srikant (1994), and Fp-growth, offered by Han et al. (Han et al., 2000). Garg and Kumar provided a performance comparison study to evaluate the scalability of each algorithm as the dataset size increases (Garg & Kumar D, 2013). They conclude that while Apriori is easily understandable, it is time-consuming, reduces speed for large data sets, and has scalability issues(Islamiyah et al., 2019). On the other hand, Fp-growth is a faster algorithm because it uses a compact data structure called a tree(K. Wang et al., 2002). Therefore, we chose Fp-growth for this thesis's purpose to identify the relationship between frequent word sets of customers review. The common challenge for these algorithms is that if the size of the dataset increases, the number of rules increases exponentially. The solution is a pruning process that defines minimum value for Support and Confidence of rules. Therefore, rules with support and confidence higher than the minimum criteria are driven by many combinations. However, a specific optimal minimum Support is not defined, and Confidence for all applications has not been defined. So, the optimal values should be calculated by testing different values. We also only consider 2-itemsets for this research and only item sets that include a word from PACK-List settings for our purpose. Finally, after setting association rule mining parameters, the framework is similar to Figure 3.14. Association rule mining is applied to negative review results from the sentiment analysis step. As a result, most frequent words and their relationship is driven and used for data analytics.



Figure 3.14 Web-based Intelligent Packaging Evaluation (WIPE) platform architecture

There are enormous applications for association rule mining. The original application was in market baskets. However, association rule mining has uses beyond the market baskets. Fauzan et al. (Fauzan et al., 2020) developed a course recommender system based on association rule mining and evaluated its performance using the Canvas Network dataset and the HarvardX-MITx dataset as case studies. Table 3.3 shows some examples of the results of their model applied to the HarvardX-MITx dataset. Their model results show that it is 79% likely that {CS50x } course will be taken if { 6.00x } course was taken.

Table 3.3: Results of rule formation for course recommender

Antecedent	Consequent	support	confidence
{6.00x}	{CS50x}	0.42	0.79
{CS50x}	$\{6.00x\}$	0.42	0.70

From the medical application aspect, since diagnostic procedures for cardiac disease are time-consuming, expensive, and error-prone, a group of methods based on computational intelligence were suggested for identifying cardiac disease(Nahar et al., 2013). By using a rule-mining-based methodology, significant risk factors for heart disease were discovered for both men and women. It is discovered from the list of healthy rules that being "female" is one of the elements of a healthy heart condition.

Social media mining is another field where association rules have become more widespread recently. Diaz-Garcia et al. proposed an association rule-based social media mining system that can reduce a big collection of tweets into a smaller list of rules. This set of rules can be used to group the emotions connected to a certain person, place, or thing at a specific time (Diaz-Garcia et al., 2019).

Another application of association rule mining is in the E-commerce recommender system. Personalized and more accurate recommendations are provided to online users by recommender systems, which also increase e-commerce revenues and customer loyalty. One of the main recommender systems is provided by Amazon, as shown in Figure 3.15.





Figure 3.15 Amazon recommendations while purchasing a product

CHAPTER 4 RESULT AND DISCUSSION

Two case studies, "Tide laundry detergent liquid soap" and "laundry detergent soap pod," were conducted to validate the WIPE application performance proposed in this thesis. Although the WIPE platform was developed to apply and be practical to various products, these two detergent products were chosen for the case studies for several reasons. First, there are two different designs for the same product function and brand. Therefore, it is possible to compare its packaging problems. Secondly, these products contain many reviews and ratings with a variety of positive and negative reviews to consider.

4.1 Case study #1: Characterizing the packaging evaluation of Tide laundry detergent liquid soap

The first product chosen for packaging evaluation was Tide laundry detergent liquid soap in a plastic bottle (Amazon, 2022). The product photo and its features are shown in Figure 4.1. This product has 31,259 ratings and 922 reviews(in 93 pages), and its average rating was 4.8 out of 5. This product is from the first American detergent brand. It cleans deep and smells fresh. Its volume is 92 Fl oz, and the product has a rigid plastic bottle that includes a handle and a cap for measuring the proper amount for medium and large loads. Based on the research conducted in the OECD report (OECD, 2021), the main packaging requirements of a detergent bottle should be as follows:

- Eye-catching presentation of the detergent at the point of sale
- Displaying details about the detergent's producer and ingredients
- Keeping the detergent's functional impact;
- Successful in making the detergent simple to stack for storage and transportation
- Protection from the detergent 's leaking during distribution hazards



Figure 4.1 product photo and its features

From the data preparation aspect, we scrapped all reviews related to the packaging word

list. Some of the examples of raw data are shown in Figure 4.2.



Figure 4.2 Raw data for analysis

A packaging word list was generated by applying the TF-IDF method on reviews with ratings between 1 and 2. The result is shown as a word cloud in Figure 4.3. The final packaging word list includes 'leak', 'package', 'product',' damage', 'deliver', 'box', 'bottle', 'packaging', 'cap', and 'broken'.



Figure 4.3 Packaging word list generation

The result of the sentiment analysis step in the web-based intelligent packaging evaluation model is shown in Figure 4.4, Figure 4.5, Figure 4.6, Figure 4.7, , Figure 4.9, Figure 4.10 and Figure 4.11 for the case study detergent bottle.

From Figure 4.4 (A), packaging failure and packaging success rates were calculated, which are 16% and 3.3%, respectively. Packaging failure and success rates were calculated by (4.1) and Eq.(4.2):

Packaging failure rate =
$$\frac{No. \ of \ negative \ reviews}{Total \ no \ of \ reviews} * 100$$
 (4.1)

Packaging success rate =
$$\frac{No. \ of \ positive \ reviews}{Total \ no \ of \ reviews} * 100$$
 (4.2)

The packaging failure rate helps designers rethink the packaging features, and if it is above their assurance level in manufacturing, they should redesign the package and address its problems. However, if the packaging failure rate is lower than the assurance level, it would be acceptable without major design changes. Moreover, Figure 4.4 (B) shows the distribution of sentiment in three classes, negative and positive reviews. Therefore, a negative rate of 82.68 % shows that the number of negative reviews dissatisfied with packaging and container design is more than positive.



Figure 4.4 (A) Packaging Failure and success rates (B) sentiment Distribution over filtered packaging-related reviews

As shown in Figure 4.5, the packaging failure rate can be demonstrated over time. The monthly packaging failure rate is shown monthly for 2020, 2021, and 2022 years. We can understand that months like August 2020, October 2020, and January 2022 show the highest problems related to packaging. The designer can extract reviews of these times and study the discussed main problems.



Figure 4.5 packaging failure rate over time

From Figure 4.6, the WIPE application's result for the case study highlights how necessary it is to read every review because the majority of negative ones are found after page 5. If there are more than ten reviews, they are listed on a separate page. Consequently, since this case study has over 922 reviews, it has 93 pages of reviews. Customers may not desire or have time to read every review on every page. As shown in Figure 4.7, "leak", "box", "bottle", "tide", "product", "detergent"," package", "spill", "arrive," and "cap" were all regarded as the most negative features of the package-product system.



Figure 4.6 Negative reviews distribution on Amazon pages




Figure 4.7 Word cloud for most frequent words inside negative reviews and most frequent

The results show that there should be some problems with the box, bottle, cap, and package. But what are the relationships between these problems? Which parts of the package/product got damaged? To answer these questions, the Fp-growth association rule mining algorithm with specific minimum support and confidence(0.01 and 0.01) values was applied to negative words, and approximately 110 frequent 2- itemsets were derived. The complete list of 2-itemsets for the case study detergent bottle is shown in Appendix A. After reviewing them and separating meaningful itemsets, 91 of the 2-itemsets remained. The top frequent 2-itemsets are shown in , sorted based on support value. as it is shown, the most frequent complaints are related to "box," "leakage," and "cap." These 2-itemsets can be divided into separate groups based on consequences. Each category indicates the relationship between the most frequent words and packaging word lists. For example, Figure 4.9 shows the relationship between the box and other frequent words. It can be inferred that the words: "leak"," spill"," open"," duct", "cap," and "soak" are accrued frequently with box, respectively.



Figure 4.8 Frequent 2-itemsets



Figure 4.9 related words result in a box problem

From the damaged parts identification aspect, we used a graph showing which words are connected to packaging word lists (directed edges' weight show support value and vertices are frequent words in reviews. In Figure 4.10, the weight of edges is the support value of a rule between two vertices(words), and directed edges between words show that there is an association between rules between ancestors and consequences. If the edges are thicker, there are strict rules between frequent item sets. For example, boxes are related to the frequent words are shown in Table 4.1 (these words are ordered based on their frequencies).

Antecedent	Consequent	Confidence	Support
leak	box	0.357143	0.160428
detergent	box	0.481481	0.069519
bottle	box	0.222222	0.064171
spill	box	0.666667	0.064171
inside	box	0.857143	0.064171
open	box	0.6875	0.058824
arrive	box	0.473684	0.048128
duct	box	0.333333	0.042781
cap	box	0.285714	0.042781
item	box	0.636364	0.037433
tide	box	0.3	0.032086
when	box	0.3125	0.026738
contain	box	0.625	0.026738
ship	box	0.5	0.026738
bag	box	0.833333	0.026738
mess	box	0.444444	0.02139
plastic	box	0.4	0.02139
tape	box	0.333333	0.02139
receive	box	0.266667	0.02139
seal	box	0.571429	0.02139
wrap	box	1	0.02139
soak	box	1	0.02139
delivery	box	0.5	0.016043
lid	box	0.25	0.016043
package	box	0.1	0.016043

Table 4.1: Association rules for consequent of "box"

Therefore, the main concern of boxes is leakage. As shown in Figure 4.10, association rules that result in leakage (as a word in the packaging word list) include {open \rightarrow leak}, {cap \rightarrow leak}, {crack \rightarrow leak}, {loose \rightarrow leak}, {top \rightarrow leak}. These item sets were accrued frequently together through negative reviews related to packaging word lists. Therefore, consumers expressed that improper sealing or wrapping, a loose cap, or an open box caused an unsatisfied box. So, it can be concluded that there are some serious design problems associated with the detergent liquid that, if improved, would satisfy more customers. All these leak-causing factors can affect the efficient containment of the product. After identifying these factors, we can conclude that the package would result in containment failure. Hence, leak testing should be redesigned at the development stage, focusing on specific components like the cap.

Using confidence in association rule mining is an excellent technique to highlight the relationships between data. By comparing items' co-occurrences to the antecedent's overall occurrence in the given rule, the confidence parameter highlights the relationship between specific items to one another inside the set, which is its main advantage. Figure 4.11 shows associations between frequent itemsets. In this figure, the weight of edges shows the confidence value for the association rule between two adjacent vertices. Hence, {shipment \rightarrow box}, {wrap \rightarrow box}, {open \rightarrow box}, {spill \rightarrow box}rules are sorted based on confidence, and it can be said that the probability of accrued leakage is high when the crack is seen in the product and package system.



Figure 4.10 Relationships between words based on Support value; edges' weight shows support value and vertices are frequent words in reviews



Figure 4.11 Relationships between words based on confidence value; edges' weight shows confidence value and vertices are frequent words in reviews

4.2 Case study #2 Characterizing the packaging evaluation of detergent pod

In this section, two detergent designs were selected to demonstrate a comparison of different packaging performances. The two detergent designs have the same brand, liquid detergent and pod detergent. Liquid detergent has a measurement cap that helps customers measure the liquid volume needed for a specific load. This product got an 85 % five-star rating on Amazon and has 922 reviews and 31,259 ratings. Its size is 92 Fl Oz (pack of 1) which is poured into a plastic bottle. The second design is a detergent pod. It got an 88% 5-star rating and includes 81 pods. Its number of reviews is 5,650, with 91,706 total ratings. Both products have free shipping on orders over \$25.00 shipped by Amazon. The designs of these products and brief descriptions are demonstrated in Figure 4.12.



Figure 4.12 Two detergent designs and Amazon descriptions

We used frequent packaging-related words in reviews with one and two ratings, and general terms related to packaging failure for the packaging word list. This list includes "break", "crack", "destroy", "defect", "defective", "scratch", "protect", "protective", "failure", "fail", "delivery", "leak", "package", "damage", "delivery", "box", "bottle", "packaging", "cap", and "lid".

Once packaging evaluation was applied by the WIPE platform based on the packaging word list on these two products, their packaging failure rates were calculated. The liquid detergent bottle got the highest packaging failure, 16 %, and the detergent pod got the lowest 4% failure rate, shown in Figure 4.13. This result showed that detergent pod performs better in packaging functions such as protection, containment, and convenience. Figure 4.13 also shows that a designer can compare the packaging failure rates of different designs for the same products with an assurance level of 10%. Therefore, the liquid detergent bottle's design needs to be reconsidered by designers because its failure rate is as much as the assurance level.



Figure 4.13 Packaging failure comparison for detergent pods and liquid bottles

Additionally, the WIPE results illustrated what concerns appeared frequently for these two designs. Hence, Figure 4.14 demonstrates that the major packaging issues included "pod", "leak"," break", "box"," lid", "package", "open", "stick", "bottle", "duct," and" cap". These common phrases imply that the pod has some significant design flaws that, if fixed and improved, would result in more satisfied customers.



Figure 4.14 Frequency of words in negative reviews of detergent pod

As discussed in the previous chapter, WIPE can predict the relationship between packaging word lists and the most frequent concerns in reviews. Hence, Table 4.2 shows the result of associate rule mining results following findings related to the detergent pod: **Pod** concerns, **Package** concerns, **Box** concerns, **Leakage** concerns, **Lid** concerns, **Bottle** concerns, and **Cap** concerns(complete association rule list is given in Appendix B).

Antecedent	Consequent	confidence	support
package	pod	0.350877	0.071942
box	contain	0.254902	0.046763
contain	box	0.342105	0.046763
leak	contain	0.206897	0.043165
tide	package	0.333333	0.035971
bottle	no	0.375	0.032374
not	leak	0.214286	0.032374
stick	leak	0.333333	0.028777
no	leak	0.333333	0.02518
soap	package	0.266667	0.014388
mess	package	0.266667	0.014388
think	package	0.4	0.014388
need	package	0.285714	0.014388
wash	package	0.25	0.014388
lid	off	0.212121	0.02518
off	lid	0.777778	0.02518
arrive	package	0.428571	0.021583
load	package	0.428571	0.021583
together	leak	0.315789	0.021583
time	package	0.357143	0.017986
carry	bottle	0.714286	0.017986
cap	tide	0.266667	0.014388
easy	package	0.444444	0.014388
big	box	0.4	0.014388
wet	box	0.8	0.014388
child	cap	0.571429	0.014388
easier	bottle	0.571429	0.014388

Table 4.2: Most frequent 2-itemsets sorted by support and confidence

For example, suppose we want to know what concern frequently occurred with the package; we can extract all antecedents that resulted in package complaints, as shown in Figure 4.15 (A). in this figure. In that case, a vertex is related words to the package, and the weight of the edge between vertices shows the support value. Therefore, it can be deduced that customers may not satisfy with: "Soap package", "Difficult package", "Package load", "Mess package", "Sticky package", "Package size", "Damaged package"





values, and vertices are frequent words in reviews

Another concern with this product was related to the cap of the container, shown in Figure 4.15-(B). This graph shows that lock caps and child-proof concerns are the most frequent items mentioned in reviews. For more details about this problem, we can correlate images of those reviews, including frequent item sets. For example, Figure 4.16 shows a change related to the

childproof of the cap. While this change improves child resistance, it makes opening the cap difficult for adults. Therefore, the designers should reconsider the cap design that satisfies adults and prevents the child from opening it. Thus, the current design shows convenience failure because, based on the convenience function of the package (as pointed out in section 2.1), the pack should be picked up, opened, and unpacked without potential damage to the content and consumer.





Figure 4.16 Old and new designs of caps based on images attached to online reviews

CHAPTER 5 CONCLUSION

5.1 Conclusion

The packaging evaluation process is vital because it prevents potential damage and predicts how the package provides main packaging functions: protection, containment, apportionment, unitization, communication, and convenience. As worldwide e-commerce is growing, distribution hazards are increasing by manual and mechanical handling, impact and vibration from transport vehicles, and environmental hazards. Laboratory evaluations, field tests, and numerical solutions have been performed for packaging evaluation. Since they have some drawbacks, an artificial intelligent-based solution was proposed.

In order to overcome these drawbacks, we developed our proposed packaging evaluation solution, the WIPE platform. The developed web application significantly improves packaging evaluation and provides intelligent decision-making for packaging designers because it works on real experiences. This study created a web-based application for package evaluation by identifying the most important relationships between frequent problems discussed in customer reviews on ecommerce websites. This software helps transform unstructured data into structured data, including association rules, sentiment analysis scores, and packaging failure rates.

This thesis aimed to achieve three main objectives by using artificial intelligence techniques:

Firstly, the WIPE application automated the data flow of packaging evaluation using customers' reviews. This application included an embedded web scrapper, which allows it to download reviews and images of reviews using a packaging word list. The primary benefit of implementing embedded web scraping is to create an automated data flow across the program, saving time and accelerating application speed. Secondly, we could validate the result of association rule mining by correlating images with frequent packaging issues. Since related works focused only on reviews' text, they could not extract information from images. By utilizing the WIPE application, all images uploaded by customers can be downloaded for each review for a specific product in a separate folder.

And finally, we could identify the relationship between frequent problems accrued in reviews and find what causes damage and their side effects. Hence, the designer can reconsider their designs and address the issue.

In conclusion, the WIPE platform enhances packaging evaluation by automating the process of evaluation, reducing the costs of field tests, minimizing errors, and improving packaging qualities. With the assistance of AI, the WIPE application scrapped approximately large amounts of reviews and interpreted the text and images to discover customers attitude' meaningful patterns and packaging concerns. It accelerates the problem-solving process by using client feedback collected over time. The significance of the WIPE application was raised when the model was applied to two different designs for the same function in order to benchmark. Therefore, we can compare packaging failure rates and success reasons of a design. Also, issue identification using real customer experience helps companies to enhance their product performance. This framework can be used for several products on the Amazon platform to identify trends and realize how packaging and product systems perform as they are changed.

5.2 Future works

This study improved packaging evaluation using web programming, natural language, and association rule mining on online customer reviews. However, future works should be focused on some limitations of natural language processing techniques and interpreting images.

82

In future works, we may improve sentiment accuracy because sentiment analysis results are the input of association rule mining. So more accurate sentiment analysis brings more meaningful and reliable rules. For this purpose, we may implement a vote-based model based on machine learning models such as LSTM, Naïve based, and rule-based models, as shown in Figure 5.1. Also, Image sentiment analysis for interpreting images would increase the accuracy of the model. Additionally, the current platform considers English reviews, so we may go over multilingual sentiment analysis. And finally, we may increase the sentiment score of packaging words in reviews to have packaging-related sentiment analysis. In this case, training data would be labeled by sentiment values of the AFINN or Bing lexicons, and then the machine learning models would be trained by this labeled data while observing a designer and applied to unseen data to predict their sentiment classification.



Figure 5.1 Multilingual vote-based sentiment analysis model

BIBLIOGRAPHY

Akbar, A., Agarwal, P., & Obaid, A. J. (2022). Recommendation engines-neural embedding to graph-based: Techniques and evaluations. International Journal of Nonlinear Analysis and Applications, 13(1), 2411–2423. https://doi.org/10.22075/IJNAA.2022.5941

Amazon. (2019). Open Data on AWS. https://registry.opendata.aws/amazon-reviews-ml/

- Amazon. (2022). Amazon.com: Tide Laundry Detergent Liquid Soap, High Efficiency (HE), Original Scent, 64 Loads: Health & Household. https://www.amazon.com/Tide-Laundry-Detergent-Liquid-Original/dp/B085V5PPP8/ref=sr_1_3_sspa?keywords=liquid%2Bdetergent&qid=16690134 84&sr=8-3-spons&sp_csd=d2lkZ2V0TmFtZT1zcF9hdGY&th=1
- Bahja, M. (2021). Natural Language Processing Applications in Business. E-Business Higher Education and Intelligence Applications. https://doi.org/10.5772/INTECHOPEN.92203
- Bahlau, J., & Lee, E. (2022). Designing moulded pulp packaging using a topology optimization and superimpose method. Packaging Technology and Science, 35(5), 415–423. https://doi.org/10.1002/PTS.2639
- Bini, S. A. (2018). Artificial Intelligence, Machine Learning, Deep Learning, and Cognitive Computing: What Do These Terms Mean and How Will They Impact Health Care? The Journal of Arthroplasty, 33(8), 2358–2361. https://doi.org/10.1016/J.ARTH.2018.02.067
- Böröcz, P., & Molnár, B. (2020). Measurement and analysis of vibration levels in stacked small package shipments in delivery vans as a function of free movement space. Applied Sciences (Switzerland), 10(21), 1–13. https://doi.org/10.3390/app10217821
- Brandenburg, R. K. (Richard K. (1991). Fundamentals of packaging dynamics / by Richard K. Brandenburg and Julian June-Ling Lee.
- Chang, Y. S., & Lee, H. J. (2018). Optimal delivery routing with wider drone-delivery areas along a shorter truck-route. Expert Systems with Applications, 104, 307–317. https://doi.org/10.1016/J.ESWA.2018.03.032
- da Silva e Silva, N., de Souza Farias, F., dos Santos Freitas, M. M., Pino Hernández, E. J. G., Dantas, V. V., Enê Chaves Oliveira, M., Joele, M. R. S. P., & de Fátima Henriques Lourenço, L. (2021). Artificial intelligence application for classification and selection of fish gelatin packaging film produced with incorporation of palm oil and plant essential oils. Food Packaging and Shelf Life, 27, 100611. https://doi.org/10.1016/J.FPSL.2020.100611
- Dadhich, A., & Thankachan, B. (2022). Sentiment Analysis of Amazon Product Reviews Using Hybrid Rule-Based Approach. Smart Innovation, Systems and Technologies, 235, 173–193. https://doi.org/10.1007/978-981-16-2877-1_17

- Deng, L., & Li, X. (2013). Machine learning paradigms for speech recognition: An overview. IEEE Transactions on Audio, Speech and Language Processing, 21(5), 1060–1089. https://doi.org/10.1109/TASL.2013.2244083
- Diaz-Garcia, J. A., Ruiz, M. D., & Martin-Bautista, M. J. (2019). Generalized Association Rules for Sentiment Analysis in Twitter. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11529 LNAI, 166–175. https://doi.org/10.1007/978-3-030-27629-4_17/COVER
- Ding, X., ACM, B. L.-P. of the 30th annual international, & 2007, undefined. (2007). The utility of linguistic rules in opinion mining. Dl.Acm.Org, 811–812. https://doi.org/10.1145/1277741.1277921
- Dornadula, V. N., & Geetha, S. (2019). Credit Card Fraud Detection using Machine Learning Algorithms. Procedia Computer Science, 165, 631–641. https://doi.org/10.1016/J.PROCS.2020.01.057
- Dunn, S. (2018). E-Commerce Packaging Strategy: Design with the End in Mind. https://www.linkedin.com/pulse/e-commerce-packaging-strategy-design-end-mind-sara-shumpert-dunn
- Emblem, A., & Emblem, H. (2012). Packaging technology Fundamentals, materials and processes Edited by Cambridge Philadelphia New Delhi.
- Esfahanian, S., & Lee, E. (2022). A novel packaging evaluation method using sentiment analysis of customer reviews. Packaging Technology and Science. https://doi.org/10.1002/pts.2686
- Fadiji, T., Berry, T., Coetzee, C., & Opara, L. (2017). Investigating the Mechanical Properties of Paperboard Packaging Material for Handling Fresh Produce Under Different Environmental Conditions: Experimental Analysis and Finite Element Modelling. Journal of Applied Packaging Research, 9(2). https://scholarworks.rit.edu/japr/vol9/iss2/3
- Fauzan, F., ... D. N.-I. J. on, & 2020, undefined. (2020). Apriori association rule for course recommender system. Socj.Telkomuniversity.Ac.Id. https://doi.org/10.21108/indojc.2020.5.2.434
- Garg, K., & Kumar D. (2013). Comparing the performance of frequent pattern mining algorithms. International Journal of Computer. https://www.academia.edu/download/81067088/download.pdf
- Ghahramani, Z. (2004). Unsupervised learning. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 3176, 72–112. https://doi.org/10.1007/978-3-540-28650-9_5/COVER
- Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). Deep learning (Vol. 1, No). Cambridge: MIT Press.

- Gustafsson, J.-E., & Nygårds, M. (2017). Loading and deformation of cigarette packages. 409. http://urn.kb.se/resolve?urn=urn:nbn:se:ri:diva-32967
- Hagman, A., Timmermann, B., Nygårds, M., Lundin, A., Barbier, C., F. M., & Östlund, S. (2017). Experimental and numerical verification of 3D forming. . . In Transactions of the 16th Fundamental Research Symposium, 3–26. https://doi.org/10.15376/frc.2017.1.3
- Han, J., Pei, J., & Yin, Y. (2000). Mining frequent patterns without candidate generation. ACM SIGMOD Record, 29(2), 1–12. https://doi.org/10.1145/335191.335372
- Helm, J. M., Swiergosz, A. M., Haeberle, H. S., Karnuta, J. M., Schaffer, J. L., Krebs, V. E., Spitzer, A. I., & Ramkumar, P. N. (2020). Machine Learning and Artificial Intelligence: Definitions, Applications, and Future Directions. Current Reviews in Musculoskeletal Medicine, 13(1), 69–76. https://doi.org/10.1007/S12178-020-09600-8/FIGURES/1
- Holland, S., Tavasoli, M., & Lee, E. (2022). Evaluating Consumer Review Images for Package Failure Using Machine Learning . The 23rd IAPRI World Packaging Conference.
- Huart, V., Nolot, J. B., Candore, J. C., Pellot, J., Krajka, N., Odof, S., & Erre, D. (2016). A damage estimation method for packaging systems based on power spectrum densities using spectral moments. Packaging Technology and Science, 29(6), 303–321. https://doi.org/10.1002/pts.2211
- IBM Cloud Education. (2021). What is Machine Learning? | IBM. https://www.ibm.com/cloud/learn/machine-learning
- International trade administration. (2021). eCommerce Sales & Size Forecast. https://www.trade.gov/ecommerce-sales-size-forecast
- Ireland, R., & Liu, A. (2018). Application of data analytics for product design: Sentiment analysis of online product reviews. CIRP Journal of Manufacturing Science and Technology, 23, 128– 144. https://doi.org/10.1016/j.cirpj.2018.06.003
- Islamiyah, Ginting, P. L., Dengen, N., & Taruk, M. (2019). Comparison of Priori and FP-Growth Algorithms in Determining Association Rules. ICEEIE 2019 - International Conference on Electrical, Electronics and Information Engineering: Emerging Innovative Technology for Sustainable Future, 320–323. https://doi.org/10.1109/ICEEIE47180.2019.8981438
- Jin, J., Ji, P., & Kwong, C. K. (2016). What makes consumers unsatisfied with your products: Review analysis at a fine-grained level. Engineering Applications of Artificial Intelligence, 47, 38–48. https://doi.org/10.1016/J.ENGAPPAI.2015.05.006
- Juárez-Varón, D., Tur-Viñes, V., Rabasa-Dolado, A., & Polotskaya, K. (2020). An Adaptive Machine Learning Methodology Applied to Neuromarketing Analysis: Prediction of Consumer Behaviour Regarding the Key Elements of the Packaging Design of an Educational

Toy. Social Sciences 2020, Vol. 9, Page 162, 9(9), 162. https://doi.org/10.3390/SOCSCI9090162

- Kaplan, A. (2022). Artificial intelligence, business and civilization : our fate made in machines. https://www.routledge.com/Artificial-Intelligence-Business-and-Civilization-Our-Fate-Made-in-Machines/Kaplan/p/book/9781032155319
- Kaur, M., & Kang, S. (2016). Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining. Procedia Computer Science, 85, 78–85. https://doi.org/10.1016/J.PROCS.2016.05.180
- Khan, A. I., & Al-Habsi, S. (2020). Machine Learning in Computer Vision. Procedia Computer Science, 167, 1444–1451. https://doi.org/10.1016/J.PROCS.2020.03.355
- Kipp, W. I. (2000). Developments in testing products for distribution[†]. Packaging Technology and Science, 89–98. https://doi.org/10.1002/9780470414835.CH6
- Knoll, D., Neumeier, D., Prüglmeier, M., & Reinhart, G. (2019). An automated packaging planning approach using machine learning. Procedia CIRP, 81, 576–581. https://doi.org/10.1016/J.PROCIR.2019.03.158
- Koca, O., Kaymakci, O. T., & Mercimek, M. (2020). Advanced Predictive Maintenance with Machine Learning Failure Estimation in Industrial Packaging Robots. 2020 15th International Conference on Development and Application Systems, DAS 2020 - Proceedings, 1–6. https://doi.org/10.1109/DAS49615.2020.9108913
- Lengas, N., Müller, K., Schlick-Hasper, E., Neitsch, M., Johann, S., & Zehn, M. (2022). Development of an analysis and testing concept for the evaluation of impact targets in the mechanical safety testing of dangerous goods packagings. Packaging Technology and Science, 35(9), 689–700. https://doi.org/10.1002/PTS.2656
- Liu, B. (2012). Sentiment Analysis and Opinion Mining.
- Liu, B. (2020). Sentiment Analysis, Opinion Mining, Emotion Classification (2nd ed.). Cambridge University Press. https://www.cs.uic.edu/~liub/FBS/sentiment-opinion-emotionanalysis.html
- Liu, Y., Jin, J., Ji, P., Harding, J. A., & Fung, R. Y. K. (2013). Identifying helpful online reviews: A product designer's perspective. Computer-Aided Design, 45(2), 180–194. https://doi.org/10.1016/J.CAD.2012.07.008
- Livingstone, S., & Sparks, L. (1994). The new German packaging laws: effects on firms exporting to Germany. International Journal of Physical Distribution & Logistics Management.
- Lockamy, A. (1995). A conceptual framework for assessing strategic packaging decisions. The International Journal of Logistics Management.

- Logistic Zipline. (2018). Freight Claims 101: Practical Advice for Quick Issue Resolution. https://ziplinelogistics.com/blog/freight-claims-101/
- Mandal, P., Khanam, J., Karmakar, S., Pal, T. K., Barma, S., Chakraborty, S., Bera, R., & Poddar, S. (2022). An Audit on Design of Pharmaceutical Packaging. Journal of Packaging Technology and Research, 6(3), 167–185. https://doi.org/10.1007/S41783-022-00141-8
- Marin, G., Hagman, A., Östlund, S., & Nygårds, M. (2022). Torsional and compression loading of paperboard packages: Experimental and FE analysis. Packaging Technology and Science. https://doi.org/10.1002/pts.2693
- Marin, G., Srinivasa, P., Nygårds, M., & Östlund, S. (2021). Experimental and finite element simulated box compression tests on paperboard packages at different moisture levels. Packaging Technology and Science, 34(4), 229–243. https://doi.org/10.1002/pts.2554
- Masis, J., Horvath, L., & Böröcz, P. (2022). The Effect of Forklift Type, Pallet Design, Entry Speed, and Top Load on the Horizontal Shock Impacts Exerted during the Interactions between Pallet and Forklift. Applied Sciences (Switzerland), 12(14). https://doi.org/10.3390/APP12147035
- Mathews, S. M. (2019). Explainable Artificial Intelligence Applications in NLP, Biomedical, and Malware Classification: A Literature Review. Advances in Intelligent Systems and Computing, 998, 1269–1292. https://doi.org/10.1007/978-3-030-22868-2_90/FIGURES/10
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal, 5(4), 1093–1113. https://doi.org/10.1016/J.ASEJ.2014.04.011
- Minaee, S., Azimi, E., & Abdolrashidi, A. (2019). Deep-Sentiment: Sentiment Analysis Using Ensemble of CNN and Bi-LSTM Models. https://doi.org/10.48550/arxiv.1904.04206
- Mitchell, T. M. (2010). CHAPTER 1 GENERATIVE AND DISCRIMINATIVE CLASSIFIERS : NAIVE BAYES AND LOGISTIC REGRESSION Learning Classifiers based on Bayes Rule. Machine Learning, 1(Pt 1-2), 1–17. https://doi.org/10.1093/bioinformatics/btq112
- Mohammad, S. M. (2017). Word Affect Intensities. LREC 2018 11th International Conference on Language Resources and Evaluation, 174–183. https://doi.org/10.48550/arxiv.1704.08798
- Molina-Besch, K., & Pålsson, H. (2020). A simplified environmental evaluation tool for food packaging to support decision-making in packaging development. Packaging Technology and Science, 33(4–5), 141–157. https://doi.org/10.1002/PTS.2484
- Moreno, A., & Redondo, T. (2016). Text Analytics: the convergence of Big Data and Artificial Intelligence. IJIMAI, ISSN-e 1989-1660, Vol. 3, No. 6, 2016, Págs. 57-64, 3(6), 57–64. https://doi.org/10.9781/ijimai.2016.369

- Nahar, J., Imam, T., Tickle, K. S., & Chen, Y. P. P. (2013). Association rule mining to detect factors which contribute to heart disease in males and females. Expert Systems with Applications, 40(4), 1086–1093. https://doi.org/10.1016/J.ESWA.2012.08.028
- Ni, J., Li, J., on, J. M.-P. of the 2019 conference, & 2019, undefined. (2019). Justifying recommendations using distantly-labeled reviews and fine-grained aspects. Aclanthology.Org. https://aclanthology.org/D19-1018/
- Njunge, C., Witman, P. D., Holmberg Joel Canacoo, P., Stewart, J. C., Morris University Alan Davis, R. G., Woodall, R., & Ghosh, B. (2022). A cloud-based system for scraping data from amazon product reviews at scale. Jisar.Org, 15. https://jisar.org/2022-15/n3/JISARv15n3.pdf#page=24
- Nygårds, M., Sjökvist, S., Marin, G., & Sundström, J. (2019). Simulation and experimental verification of a drop test and compression test of a gable top package. Packaging Technology and Science, 32(7), 325–333. https://doi.org/10.1002/PTS.2441
- Obertson, G. L. (1990). Good and bad packaging: who decides? International Journal of Physical Distribution & Logistics Management.
- OECD. (2021). Case study on detergent bottles | Enhanced Reader.
- Paine, F. A., & Paine, H. Y. (1983). A Handbook of Food Packaging. https://doi.org/10.1007/978-1-4615-2810-4
- Pålsson, H. (2018). Packaging Logistics: Understanding and managing the economic and environmental impacts of packaging in supply chains. 248. https://books.google.com/books?hl=nl&lr=&id=WyxdDwAAQBAJ&oi=fnd&pg=PP1&dq= Bulk+packaging+suppliers+focused+on+efficiency+and+environment&ots=I5EhTrYceu&s ig=jp9bkSvMFtJMSHQRZMMjVpxMdgw
- Sabbeh, S. F. (2018). Machine-Learning Techniques for Customer Retention: A Comparative Study. Undefined, 9(2), 273–281. https://doi.org/10.14569/IJACSA.2018.090238
- Saghir, M. (2002). Packaging information needed for evaluation in the supply chain: The case of the Swedish grocery retail industry. Packaging Technology and Science, 15(1), 37–46. https://doi.org/10.1002/pts.565
- Shankar, V., & Parsana, S. (2022). An overview and empirical comparison of natural language processing (NLP) models and an introduction to and empirical application of autoencoder models in marketing. Journal of the Academy of Marketing Science, 50(6), 1324–1350. https://doi.org/10.1007/S11747-022-00840-3/FIGURES/3
- Tavasoli, M., Yaghmaee, M. H., & Mohajerzadeh, A. H. (2016). Optimal placement of data aggregators in smart grid on hybrid wireless and wired communication. 2016 4th IEEE

International Conference on Smart Energy Grid Engineering, SEGE 2016, 332–336. https://doi.org/10.1109/SEGE.2016.7589547

- Tian, W., & Song, X. (2020). Selection of Optimal Packaging Methods for Different Food Based on Big Data Analysis. Proceedings of 2020 IEEE International Conference on Power, Intelligent Computing and Systems, ICPICS 2020, 558–561. https://doi.org/10.1109/ICPICS50287.2020.9202379
- Upadhyaya, M., & Nygårds, M. (2017). A finite element model to simulate brim forming of paperboard. 395–408. http://urn.kb.se/resolve?urn=urn:nbn:se:ri:diva-30296
- Wang, K., Tang, L., Han, J., & Liu, J. (2002). Top down FP-growth for association rule mining. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2336, 334–340. https://doi.org/10.1007/3-540-47887-6_34
- Wang, Y., Peng, S., Zhou, X., Mahmoudi, M., & Zhen, L. (2020). Green logistics location-routing problem with eco-packages. Transportation Research Part E: Logistics and Transportation Review, 143. https://doi.org/10.1016/J.TRE.2020.102118
- Wu, H., Luk, R., Wong, K., on, K. K.-A. T., & 2008, undefined. (2008). Interpreting tf-idf term weights as making relevance decisions. Dl.Acm.Org, 26(3). https://doi.org/10.1145/1361684.1361686
- Wu, Y., & Lu, Y. (2019). An intelligent machine vision system for detecting surface defects on packing boxes based on support vector machine. Https://Doi.Org/10.1177/0020294019858175, 52(7–8), 1102–1110. https://doi.org/10.1177/0020294019858175
- Yang, X., Han, M., Tang, H., Li, Q., & Luo, X. (2020). Detecting Defects with Support Vector Machine in Logistics Packaging Boxes for Edge Computing. IEEE Access, 8, 64002–64010. https://doi.org/10.1109/ACCESS.2020.2984539
- Yang, Y., Zhang, X., Yin, J., & Yu, X. (2020). Rapid and Nondestructive On-Site Classification Method for Consumer-Grade Plastics Based on Portable NIR Spectrometer and Machine Learning. Journal of Spectroscopy, 2020. https://doi.org/10.1155/2020/6631234
- Zhang, L., Wang, S., & Liu, B. (2018). Deep Learning for Sentiment Analysis : A Survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4). https://doi.org/10.48550/arxiv.1801.07883
- Zhang, Z. (2018). Text Mining for Social and Behavioral Research Using R. https://books.psychstat.org/textmining/sentiment-analysis.html

APPENDIX A: ASSOCIATION RULES FOR CASE STUDY #1

Table 3 Association rules for case study #1 detergent bottle with minimum support of 0.01 and

Antecedent	Consequent	Confidence	Support
leak	box	0.357143	0.160428
detergent	box	0.481481	0.069519
bottle	box	0.222222	0.064171
spill	box	0.666667	0.064171
inside	box	0.857143	0.064171
open	box	0.6875	0.058824
arrive	box	0.473684	0.048128
box	arrive	0.140625	0.048128
duct	box	0.333333	0.042781
cap	box	0.285714	0.042781
item	box	0.636364	0.037433
tide	box	0.3	0.032086
when	box	0.3125	0.026738
contain	box	0.625	0.026738
ship	box	0.5	0.026738
bag	box	0.833333	0.026738
mess	box	0.44444	0.02139
plastic	box	0.4	0.02139
tape	box	0.333333	0.02139
receive	box	0.266667	0.02139
seal	box	0.571429	0.02139
wrap	box	1	0.02139
soak	box	1	0.02139
half	leak	1	0.02139
great	leak	1	0.02139
delivery	box	0.5	0.016043
lid	box	0.25	0.016043
package	box	0.1	0.016043
recieve	box	1	0.016043
liquid	box	0.428571	0.016043
side	box	0.6	0.016043
top	box	0.6	0.016043
loose	box	0.6	0.016043
turn	box	0.5	0.016043
fine	box	0.5	0.016043

minimum confidence of 0.01

Table 3 (cont'd)

time	box	0.375	0.016043
look	box	0.75	0.016043
interior	box	1	0.016043
pack	box	0.75	0.016043
soap	box	0.428571	0.016043
shipment	box	0.75	0.016043
buy	leak	0.333333	0.016043
laundry	leak	0.5	0.016043
everything	leak	0.6	0.016043
least	leak	0.75	0.016043
able	package	1	0.016043
away	bottle	1	0.016043
even	box	0.25	0.010695
damage	box	0.2	0.010695
purchase	box	0.4	0.010695
expencive	box	1	0.010695
order	box	0.333333	0.010695
ruin	box	0.5	0.010695
crack	box	0.4	0.010695
close	box	0.5	0.010695
big	box	0.285714	0.010695
content	box	1	0.010695
issue	box	1	0.010695
except	box	0.666667	0.010695
deliver	box	0.333333	0.010695
store	box	1	0.010695
stick	leak	0.5	0.010695
fortune	leak	1	0.010695
exactly	leak	1	0.010695
value	leak	1	0.010695
properly	leak	0.666667	0.010695
sure	leak	0.5	0.010695
during	leak	1	0.010695
seem	leak	1	0.010695
secure	leak	0.666667	0.010695
want	leak	0.5	0.010695
everywhere	leak	0.666667	0.010695
full	leak	0.5	0.010695
happy	leak	0.666667	0.010695

Table 3 (cont'd)

lose	leak	0.666667	0.010695
almost	leak	0.666667	0.010695
went	leak	0.666667	0.010695
bad	leak	1	0.010695
measure	lid	1	0.010695
place	lid	0.4	0.010695
send	package	1	0.010695
see	package	0.666667	0.010695
waste	package	1	0.010695
avoid	damage	1	0.010695
run	bottle	1	0.010695
differ	bottle	1	0.010695
pour	bottle	0.666667	0.010695
unable	bottle	1	0.010695
try	bottle	0.666667	0.010695
few	bottle	1	0.010695
cloth	bottle	1	0.010695
tire	bottle	1	0.010695
throw	bottle	0.666667	0.010695
fail	tide	0.5	0.010695
hold	cap	1	0.010695

APPENDIX B: ASSOCIATION RULES FOR CASE STUDY #2

Table 4 Association rules for case study #2 detergent pod with minimum support of 0.01 and

Antecedent	Consequent	Confidence	Support
package	pod	0.350877	0.071942
box	contain	0.254902	0.046763
contain	box	0.342105	0.046763
leak	contain	0.206897	0.043165
tide	package	0.333333	0.035971
bottle	no	0.375	0.032374
not	leak	0.214286	0.032374
stick	leak	0.333333	0.028777
no	leak	0.333333	0.02518
lid	off	0.212121	0.02518
off	lid	0.777778	0.02518
will	package	0.4	0.021583
arrive	package	0.428571	0.021583
load	package	0.428571	0.021583
together	leak	0.315789	0.021583
time	package	0.357143	0.017986
SO	box	0.227273	0.017986
carry	bottle	0.714286	0.017986
cap	tide	0.266667	0.014388
soap	package	0.266667	0.014388
mess	package	0.266667	0.014388
think	package	0.4	0.014388
need	package	0.285714	0.014388
wash	package	0.25	0.014388
easy	package	0.44444	0.014388
big	box	0.4	0.014388
wet	box	0.8	0.014388
child	cap	0.571429	0.014388
easier	bottle	0.571429	0.014388
wish	package	0.75	0.010791
dont	package	0.3	0.010791
damage	package	0.333333	0.010791
sticky	package	0.333333	0.010791
difficult	package	0.5	0.010791
item	package	0.375	0.010791
size	package	1	0.010791

minimum confidence of 0.01

Table 4 (cont'd)

great	package	0.333333	0.010791
last	package	0.75	0.010791
receive	package	0.428571	0.010791
powder	box	0.75	0.010791
small	box	0.5	0.010791
amazon	box	0.75	0.010791
ship	box	0.75	0.010791
around	box	0.75	0.010791
top	leak	0.375	0.010791
smell	leak	0.428571	0.010791
deliver	when	1	0.010791
lock	cap	1	0.010791
safety	lid	0.6	0.010791
lug	bottle	0.75	0.010791
destroy	cloth	0.75	0.010791