USING SOCIAL MEDIA BIG DATA WITH MACHINE LEARNING TO IMPROVE CUSTOMER SATISFACTION

HİLAL DEMİR

BOĞAZİÇİ UNIVERSITY

2023

USING SOCIAL MEDIA BIG DATA WITH MACHINE LEARNING TO IMPROVE CUSTOMER SATISFACTION

Thesis submitted to the

Institute for Graduate Studies in Social Sciences

in partial fulfillment of the requirements for the degree of

Master of Arts

in

Business Information Systems

by

Hilal Demir

Boğaziçi University

2023

DECLARATION OF ORIGINALITY

I, Hilal Demir, certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Bogaziçi University, including final revisions required by them.

Signature: .....

Date: 10/05/2023

ABSTRACT

Using Social Media Big Data with Machine Learning to Improve Customer Satisfaction

With the increasing importance of customer relationship management and improving customer support in today's competitive business landscape, there is a growing need to leverage machine learning techniques for gaining insights, forecasts, and better decision-making. Sentiment analysis, in particular, has emerged as a powerful tool for improving customer support services. In this study, we explore the use of three gradient boosting algorithms, XGBoost, CatBoost, and LightGBM, for sentiment classification on Twitter data. We employ ensemble classifications to analyze the sentiment of the data and observe improvements in performance. Our results are compared to other two algorithms that are popularly used in the context of sentiment analysis and show that the ensemble classification of the three algorithms yields the highest accuracy and F1 score. By addressing the gap in understanding how different machine learning algorithms can be used to enhance customer support processes, this research aims to contribute to the improvement of customer satisfaction and loyalty. Specifically, the study aims to improve the accuracy of sentiment classification, thereby enabling businesses to better meet customer expectations for fast and efficient customer support.

ÖZET

Müşteri Memnuniyetini Geliştirmek İçin Sosyal Medya Verilerinin Makine Öğrenimi ile Kullanımı

Günümüzün rekabetçi iş dünyasında, müşteri ilişkileri yönetimi ve müşteri desteğinin artan önemiyle birlikte, makine öğrenimi tekniklerinin kullanılması için giderek artan bir ihtiyaç vardır. Özellikle duygu analizi, müşteri destek hizmetlerinin geliştirilmesi için güçlü bir araç olarak ortaya çıkmıştır. Bu çalışmada, Twitter verileri üzerinde duygu sınıflandırması için XGBoost, CatBoost ve LightGBM olmak üzere üç makine öğrenimi algoritması kullanımı araştırılmaktadır. Aynı zamanda algoritmaların farklı kombinasyonlar halinde sınıflandırmaları kullanılmakta ve performanstaki iyileşmeler gözlemlenmektedir. Sonuçlar duygu analizinde sıkça kullanılan iki algoritmayla daha karşılaştırılmıştır ve üç algoritmanın birlikte sınıflandırmasının en yüksek doğruluk ve F1 skorunu verdiğini göstermektedir. Farklı makine öğrenimi algoritmalarının müşteri destek süreçlerini nasıl geliştirmek için kullanılabileceği konusundaki araştırma boşluğunu ele alarak, bu araştırma müşteri memnuniyeti ve sadakatinin artırılmasına katkıda bulunmayı amaçlamaktadır. Özellikle, çalışma duygu sınıflandırmasının doğruluğunu arttırmayı hedeflemekte ve işletmelerin hızlı ve etkili müşteri desteği için müşteri beklentilerini daha iyi karşılamalarına olanak tanımaktadır.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In recent years, importance of customer relationship management increased across organizations with its goal to establish long term relationships with customers leading to increased loyalty. The increase in importance of customer relationship management encourages repetitive business and increases good reputation for the companies. It is now more crucial than ever for organizations to have a clear CRM strategy due to the escalating market competitiveness. In a study by Reinartz, Krafft, and Hoyer (2004), the authors concluded that the implementation of an effective customer relationship management process is crucial for firms that want to remain competitive in the corporate world of today. In the current digital age, customers have high expectations for fast and efficient customer service and organizations must work to meet these expectations in order to remain competitive (Verhoef, Lemon, Parasuraman, Roggeveen, Tsiros, & Schlesinger, 2009). In fact, The Customer Experience Impact (2010) report showed that 80% of customers would cease doing business with a company after just one negative customer service experience. About 95% of the participants stated that after a negative customer experience they would "take action" ("The Customer Experience Impact Report", 2010). In order to match these expectations and keep customers loyal, businesses are under pressure to enhance their customer support services (Gulfraz, Sufyan, Mustak, Salminen, & Srivastava, 2022). While literature provides various studies examining how to improve customer satisfaction and loyalty like the study by

Homburg, Jozić, and Kuehnl (2017) that provided an overview of customer experience management and its impact on customer satisfaction and loyalty, a promising solution is to leverage machine learning algorithms to examine various types of customer data like social media data and improve customer support services.

Machine learning has a potential to lead more effective marketing strategies by providing personalized recommendations, predicting customer needs, and automating routine tasks (Siegel, 2013). In the context of customer satisfaction, businesses can consider customer input on social media platforms and forecast user behavior and offer personalized recommendations using volumes of customer data through machine learning algorithms. Businesses, then, can personalize the customer experience, forecast customer churn, and examine customer feedback by applying machine learning to their customer services. For example, Ngai, Xiu, & Chau, (2009) discussed how businesses can use machine learning to analyze customer data, provide personalized recommendations, forecast customer churn and examine customer feedback to improve customer service. By leveraging machine learning techniques, businesses create a customer-centric culture, which would increase customer satisfaction and, in turn, improve sales and profits.

In the context of customer support, machine learning algorithms like natural language processing (NLP), chatbots, and predictive analytics can help to identify patterns in customer data and provide insights into customer behavior. Various applications of machine learning like chatbots can then be used to improve customer support services, which can ultimately lead to increased customer satisfaction and improved customer relationships (Misischia, Poecze, & Strauss, 2022).

There are several machine learning algorithms that can be used to analyze social media data and improve customer support services. XGBoost, CatBoost, and LightGBM are three popular algorithms that have been shown to be effective and going to be used in the context of this research. These algorithms are all gradient boosting methods that use decision trees to make predictions. The XGBoost algorithm is a fast and scalable gradient boosting library that has been widely used in industry for its accuracy and speed. CatBoost is another gradient boosting algorithm that has been shown to perform well in scenarios with categorical data. LightGBM is a gradient boosting framework that accelerates training process by histogram-based algorithms. To the best of our knowledge, there are limited research that use these three algorithms in different fields and none in the purpose of increasing customer satisfaction in organizations' support systems with the use of social media big data.

The objective of this research is to explore how these machine learning algorithms can be used in a customer relationships management context. While the field of customer relationship management has seen a significant transformation with the advent of machine learning algorithms, there is still a lack of understanding of how different machine learning algorithms can be used to improve customer support services (Akter et al., 2022). Furthermore, while previous studies have examined the use of machine learning algorithms in customer support, most of them have focused on a single algorithm or a limited range of algorithms (Zhang, Srivastava, Sharma & Eachempati, 202). Therefore, there is a need to explore how different algorithms can be used together to provide more accurate insights into customer behavior (Odusami, Abayomi-Alli, Misra, Abayomi-Alli, & Sharma, 2021).

In this context, we will examine the use of XGBoost, CatBoost, and LightGBM, three widely used machine learning algorithms, to analyze data coming from Twitter account of a large global organization as a case and provide insights into customer behavior (Chen, & Guestrin, 2016; Prokhorenkova, Gusev, Vorobev, Dorogush, & Gulin, 2018; Ke et al., 2017). Reasoning behind the selection of algorithms are, XGBoost algorithm offers resistance to overfitting and makes the model more robust, CatBoost algorithm improves generalization and cand handle categorical data and LightGBM is efficient in large-scale datasets (Abbasniya, Sheikholeslamzadeh, Nasiri & Emami, 2022). Then these insights will be compared other popular machine learning algorithms for sentiment analysis and discussed about their possible effects on the improvements of the customer support services and maintain customer loyalty. By addressing this gap, the research aims to contribute to the growing use of machine learning algorithms in customer support and provide practical insights for organizations seeking to improve their customer support services.

This study aims to provide insight where there is a gap in knowledge for a quantitative study for use of machine learning algorithms in the purpose of increasing customer satisfaction in organizations' support systems with the use of social media data. The literature review for the study aims to present an understanding of big data and machine learning concepts followed by customer relationship management and customer support concepts. Then, in order to capture the practices in the customer satisfaction-related studies with the use of machine learning algorithms, concepts like customer segmentation, sentiment analysis, and natural language processing are investigated. Also, description and review of the XGBoost, CatBoost, and LightGBM algorithms are provided.

In the materials and methodology, explanation of the dataset that is used in this research and the detailed descriptions and implementation details of the XGBoost, CatBoost, and LightGBM algorithms are provided. In the result and discussion section, evaluation metrics, findings on the performances of these algorithms and their potential effects in customer support services will be presented.

This study aims to answer following research question:

RQ: Which ML algorithms can be used on social media data in CRM context?

CHAPTER 2

LITERATURE REVIEW

The literature research was conducted in order to comprehend the significance of the study topic, the components of the research question, and to examine studies relating to customer satisfaction and machine learning algorithms. First, general concepts like big data, machine learning and customer relationship management are examined. Then, concepts like customer segmentation, sentiment analysis and natural language processing are examined in order to capture the practices in the customer satisfaction related studies where machine learning algorithms are used. Later, the previous studies related to XGBoost, CatBoost and LightGBM algorithms are examined in order to justify the reasoning behind the selection of these algorithms with their specific features and improvements.

2.1 Big data and machine learning

Big data has recently risen to the top of the list of essential issues in both business and computer science. Big data is referred when the data has volume, variety and velocity and large and complicated datasets that cannot be analyzed using traditional data processing techniques (Weerasinghe, Pauleen, Scahill, & Taskin, 2018). Davenport (2014) defines big data as large, complex, rapidly changing datasets that cannot be analyzed with traditional data processing methods or tools. Big data's rise has given companies new opportunity to obtain knowledge, make wise choices, and enhance their operations.

Big data has a variety of applications in different fields, including business, healthcare, education, and government. In the business sector, big data is used to improve decision-making, develop new products and services, and increase revenue. For example, Walmart, the world's largest retailer, uses big data to optimize its supply chain and reduce costs. By analyzing data from its stores, warehouses, and suppliers, Walmart can make better decisions on inventory management, pricing, and promotions (Davenport & Harris, 2007). In healthcare, big data is used to improve patient outcomes, reduce costs, and enhance clinical research. Electronic health records, sensor data, and genomic data are some of the sources of big data in healthcare. By analyzing these datasets, healthcare organizations can identify patterns and trends, personalize treatments, and predict disease outbreaks (Chen, Chiang, & Storey, 2012).

Big data has its challenges associated with its collection, storage, processing, and analysis. Data quality, or the accuracy, completeness, and consistency of data, are among the most common problems in big data and they need to be addressed in big data analytics projects (Bell & Shankar, 2019). Big data is often unstructured and noisy, which can affect the quality of analysis and decision-making (Marr, 2015). Another challenge is privacy and security: as big data often contains sensitive information that can be exploited by malicious actors (Chen & Zhao, 2012). Moreover, the sheer volume and velocity of big data can overwhelm traditional data processing systems, leading to performance and scalability issues (Buyya, Calheiros, & Dastjerdi, 2016).

Future research and development in big data can go in several different areas. Machine learning, which is the use of algorithms and statistical models to analyze and learn from data, is one of the primary areas of concentration. For example, in the study by Jordan and Mitchell (2015), current state and future trends of machine learning is

discussed and it has been concluded that machine learning has the potential to revolutionize fields such as healthcare, finance, and transportation. According to Provost and Fawcett (2013), machine learning is an essential technique to address some of the challenges of big data, such as data accuracy and analysis complexity. Another area of focus is the new technology and tools advancement for big data processing and analysis, such as distributed computing, cloud computing, and data visualization (Buyya et al., 2016).

Machine learning (ML) has become a strong tool for businesses to use their data for insights, forecasts, and better decision-making. Businesses can now analyze massive volumes of data, spot patterns and trends, and make predictions that would be challenging for humans to make. For example in their study, Linden, Smith and York (2003) showed that Amazon's recommendation engine uses ML algorithms to offer personalized product recommendations based on a customer's purchase history. ML algorithms can detect fraud in real-time, enabling organizations to prevent financial losses. For example, a study conducted by Rtayli and Enneya (2020) founded that ML algorithms could detect credit card fraud in real-time with high accuracy, cutting down the number of false positives.

The combination of big data with machine learning has several applications across industries. As it can be seen in Table 1, in the healthcare industry, big data analytics with machine learning has the potential to improve patient outcomes by predicting diseases, identifying high-risk patients, and creating personalized treatment plans (Supriya & Deepa, 2020). In the retail industry, big data with machine learning can be used for customer segmentation, personalized marketing, and demand forecasting (Gandomi & Haider, 2015). Additionally, the ethical implications of using big data with

machine learning, such as privacy concerns and bias, need to be addressed to ensure that the benefits of this combination are not outweighed by the potential harm (Narayanan & Shmatikov, 2010). Despite these kind of challenges, big data with machine learning can help gain valuable insights from large volumes of data that would be difficult or impossible to analyze using traditional methods (Gandomi & Haider, 2015). Machine learning algorithms can also learn and adapt to new data, allowing organizations to continuously improve their processes and decision-making (Jane & Ganesh., 2019). Additionally, big data with machine learning can improve operational efficiency by automating repetitive tasks and reducing errors (Watson, 2014).

Big data and machine learning uses and advantages are anticipated to grow as technology advances, necessitating investments in this field from businesses if they want to remain competitive. However, big data also poses significant challenges, such as data accuracy, privacy, and security. New technologies and approaches are always being developed in the big data industry to address these issues and allow for more efficient analysis and decision-making.

Table 1. Summary of the Applications and Challenges of Big Data with Machine Learning Across Industries

Industry	Application	Citations
Healthcare	Predicting diseases, identifying high-risk patients, creating personalized treatment plans	Supriya & Deepa, 2020
Finance	Fraud detection, credit risk assessment, investment analysis	Rtayli & Enneya, 2020
Retail	Customer segmentation, personalized marketing, demand forecasting	Gandomi & Haider, 2015
All Industries	Ethical implications such as privacy concerns and bias	Narayanan & Shmatikov, 2010
All Industries	Gain valuable insights from large volumes of data, learn and adapt to new data, improve operational efficiency	Gandomi & Haider, 2015; Rtayli & Enneya, 2020

2.2 Customer relationship management

In recent years, customer relationship management (CRM) has grown in importance as a component of company strategy. CRM's goal is to establish enduring connections with customers in order to increase customer loyalty, encourage repeat business, and spread good word of mouth (Grönroos, 2007). It is now more crucial than ever for organizations to have a clear CRM strategy due to the escalating market competitiveness.

CRM's primary objective is to establish and sustain customer relationships, which ultimately results in more profitability and sales. According to Gummesson (2017), CRM is a customer-centric approach to business that fosters mutually beneficial

long-term relationships with customers. Therefore, CRM plays a critical role in creating a customer-centric business culture.

The integration of different data sources is one of the main challenges in CRM. Businesses gather client data from a variety of sources, including surveys, social media, and website analytics. Integrating this data into a single CRM system that can be utilized to build thorough customer profiles is the difficult part (Anshari, Almunawar, Lim, & Al-Mudimigh, 2019). According to Chen and Popovich (2003), data integration is one of the critical challenges businesses face in implementing CRM. Because it is a new concept in customer management, another challenge is the lack of customer engagement in a digital environment (Verhoef, Reinartz, & Krafft, 2010). CRM requires constant communication with clients, which can be time- and resource-consuming. Customers must feel involved in the CRM process and that their requirements are being met. According to Chen, Chiang, and Storey (2012), customer engagement is crucial for the success of CRM.

Businesses can use a variety of strategies to increase customer satisfaction. Utilizing social media channels to read client reviews is one such strategy. Businesses may monitor social media platforms for mentions of their brand and quickly respond customer inquiries (Gupta & Kabadayi, 2010). Customers would see this as a sign that the company values their needs and is prepared to fix their issues. Another strategy is that businesses can position themselves as thought leaders in their field by posting interesting and informative material on social media networks (Kim & Ko, 2010). This would improve customer engagement and help businesses to attract new customers. Social media networks can be used by businesses to reward loyal clients. Businesses can

encourage repeat purchases and boost consumer loyalty by giving their social media followers access to special discounts and promotions (Chen, Fay, & Wang, 2011).

Machine learning has the potential to transform how businesses approach CRM. Large volumes of data can be analyzed by ML algorithms to detect patterns and trends that would be difficult or impossible for humans to identify. Additionally, it can be used to forecast consumer behavior and offer recommendations, both of which would prove the customer experience. For example, businesses can use ML algorithms for customer service. ML algorithms can analyze customer inquiries and provide automated responses, which would reduce the response time and therefore can improve customer satisfaction (Gnewuch, Morana, Adam, & Maedche, 2018). For example, a study conducted by Liu, Hu, Yan, & Lin (2023) founded a correlation between using ML-based chatbots and increase in customer satisfaction.

Businesses can employ various strategies to improve customer satisfaction, including listening to customer input on social media platforms, giving customers useful material, and rewarding royal customers. The customer experience would be enhanced by the use of ML algorithms, which can analyze large volumes of customer data, forecast user behavior, and offer personalized recommendations (Liu et al., 2023). Businesses can use techniques like personalizing the customer experience, applying ML to customer service, forecasting customer churn, and examining customer feedback. By this way, businesses create a customer-centric culture, which would increase customer satisfaction and, in turn, improve sales and profits.

2.3 Customer support in organizations

Any organization's operations depend heavily on its customer service and it involves assisting and directing consumers with regard to the goods or services that the business offers (Marino & Presti, 2018). Effective customer service can help businesses in retaining current clients and attracting new ones. Customer support has changed as a consequence of the development of digital technology, moving away from traditional call centers and toward online platforms like live chat, social media, and email (Barnes & Mattson, 2009).

Customer support is important for several reasons. Firstly, it helps organizations retain existing customers. According to a study by Gupta et al. (2010), customer retention has a significant positive effect on a firm's value, as measured by its market capitalization. Second, it helps businesses in gaining new clients by encouraging favorable word-of-mouth advertising (Hennig-Thurau, Thorsten, Gwinner, & Gremler, 2002). A happy consumer is more likely to recommend the business to their friends and family, which could result in attracting new clients. Finally, according to a report by PwC on customer experience, "by fixing problems the first time, you can build a stronger emotional connection with your customers, one that drives loyalty and growth" The report emphasizes the importance of using customer feedback and support to identify and address issues with products and services. (“Experience is everything: Here's how to get it right.”, 2018, p.11)

A variety of technologies can be applied to enhance customer service. One such technology is live chat. Customers can communicate with a support person in real-time using live chat. Customers may find this to be a quick and simple approach to get their questions answered and issues fixed. According to McLean, Osei-Frimpong, Wilson,

and Pitardi (2020), human touch in have a positive effect on customer to build trust and increase purchase intention. Another technology for customer support is machine learning. Machine learning algorithms can analyze customer inquiries, provide automated responses, and even identify the most common issues that customers face (Gnewuch et al., 2018). One of the main benefits of machine learning is that it can automate customer support processes, reducing response times and increasing efficiency (McClelland, 2020).

There are several applications of machine learning in customer support, including natural language processing (NLP), chatbots, and predictive analytics as can be seen from the Table 2. For example, chatbots powered by machine learning algorithms can provide customers with quick and accurate answers to their questions, without the need for human intervention (Dahiya, 2017). These chatbots can be implemented for understanding natural language and can use contextual information to provide personalized responses. In a study by Gartner, it was estimated that by 2020, 85% of all customer interactions will be handled without the need for human intervention ("CRM Strategies and Technologies to Understand, Grow and Manage Customer Experiences," 2011). This highlights the increasing role that machine learning is playing in customer support.

Table 2. Summary of the Technologies Used in Customer Services

Technology	Description	Benefits	Citations
Live Chat	Real-time communication with support person	Quick and simple approach for customers	McLean, Osei-Frimpong, Wilson & Pitardi (2020)
Machine Learning	Analyze customer inquiries, provide automated responses, and identify common issues	Automate customer support processes, reduce response times, and increase efficiency	Gnewuch et al. (2018)
Natural Language Processing	Process and understand human language	Provide personalized responses and improve chatbot accuracy	Dahiya (2017)
Chatbots	AI-powered programs that can engage with customers in conversation	Provide quick and accurate answers without the need for human intervention	Gartner (2011)

Another area where machine learning can be useful is personalizing customer interactions. Customer data such as customer purchase history, browsing behavior, and social media activity can be analyzed by machine learning algorithms to identify patterns and make recommendations for products or services that may be of interest to the customer (Verhoef, Kannan, & Inman, 2015). This can improve the customer experience in addition to increase sales and customer loyalty. Also it can be used to identify the most common issues that customers face. By analyzing customer support data, such as chat logs, emails, and call center transcripts, machine learning algorithms can detect patterns and trends in customer issues (Loebbecke & Picot 2015). Another

example to analyze customer data could be given as the customer purchase history and browsing behavior, can be predicting customer needs: machine learning algorithms can predict what products or services a customer may be interested in, and when they may be likely to make a purchase (Loebbecke et al., 2015).

Customer support is an essential component of any business's operations. Effective customer support can help businesses in retaining existing clients, attracting new ones, and locating and resolving problems with their goods or services (Rust, Lemon & Zeithaml, 2004). Organizations automate customer support procedures, personalize client interactions, pinpoint typical customer problems, and forecast consumer behavior to deliver successful customer care. However, to effectively leverage machine learning in customer support, organizations must address challenges such as the need for high-quality data since if the data used to train the algorithms is incomplete, inaccurate, or biased, then the results may be unreliable (Paleyes, Urma & Lawrence 2022).

2.4 Customer segmentation

Customer segmentation is the process of categorizing customers based on shared traits like preferences, behavior, or demographics. Businesses can then modify their marketing, sales, and customer service methods to suit the needs and preferences of each group. (Hu, Liu, Li, Dai, & Nakao, 2023)

Customer segmentation enables businesses to personalize their marketing campaigns to each group's preferences and needs. For example, a study conducted by Budac (2016) founded that segmenting customers based on their preferences resulted in higher click-through rates and conversions in email marketing campaigns. Another

example, a study conducted by Parasuraman & Zinkhan (2002) founded that segmenting customers based on their preferences and needs resulted in higher customer satisfaction and loyalty.

Machine learning algorithms can provide a more accurate and detailed understanding of each group's preferences and needs, which can inform marketing and product development strategies. For example, a study conducted by Loebbecke et al. (2015) founded that using machine learning algorithms to segment customers based on purchase history resulted in more accurate segmentation and improved marketing effectiveness. Another example, a study conducted by Cuadros and Domínguez (2014) founded that using ML algorithms to segment customers in real-time can lead in improved sales and customer satisfaction. By understanding each group's unique needs, businesses can develop messaging that speaks to each demographic specifically (Kotler, 2001). For example, a study conducted by Pawełoszek (2021) founded that using ML algorithms to segment customers based on their browsing history resulted in personalized recommendations that increased sales.

Businesses can use customer segmentation as a strong tool to customize their marketing, sales, and customer service strategies to meet the needs and preferences of each group. Businesses can personalize marketing, increase client retention, and inform product development by using customer segmentation. Businesses may increase accuracy, carry out real-time segmentation, and customize marketing through customer segmentation using machine learning. For example, in a study by Teichert, Shehu, and von Wartburg (2008), customer segmentation was used in the airline industry to develop effective marketing strategies by understanding the unique needs of each customer segment.

2.5 Sentiment analysis

Sentiment analysis is a process of evaluating text data to determine the emotional tone or attitude reflected in the text (Cambria & White, 2014). The primary purpose of sentiment analysis is to determine whether the sentiment expressed in each part of text is positive, negative, or neutral. The rise of social media and online reviews has made sentiment analysis an essential tool for businesses to make informed decisions on product development, customer service, and marketing strategies (Liu & Zhang, 2012).

Businesses can identify areas of improvement in their products or services and make informed decisions to address customer concerns by analyzing customer feedback. For example, a study conducted by Pang and Lee (2009) founded that sentiment analysis can provide useful insights into customer preferences and help businesses improve customer satisfaction. Businesses can learn how their brand is perceived in relation to their competitors by analyzing sentiment across social media platforms. For example, a study conducted by Cao, Duan, and Gan (2011) founded that sentiment analysis can be used to compare brand sentiment across different social media platforms. Also, businesses can identify negative sentiment in real-time and take action to resolve the problem before it escalates. For example, a study conducted by Kaur and Kumar (2015) founded that sentiment analysis can be used to detect negative sentiment during a crisis and help businesses respond quickly.

There are several algorithms that are commonly used in sentiment analysis, and their popularity can depend on the specific application and context. One of the most common algorithm used in sentiment analysis is Support Vector Machines (SVM) (Pang et al., 2009). SVM is a supervised learning algorithm that can be used to predict customer churn, fraud detection, and text classification (Cortes &Vapnik, 1995). SVMs

are proven to be effective on test categorization (Mohammad, Sobhani, & Kiritchenko, 2017). Another popular algorithm is Random Forest (RF) which is an ensemble learning method that constructs multiple decision trees and combines them can be used for customer segmentation, churn prediction, and recommendation systems (Breiman, 2001). In their study, Amrani, Lazaar, and Kadiri (2018) stated that random forest algorithm is one of the best algorithm compared to other classification algorithms that can classify large amounts of data with accuracy.

Sentiment analysis is a powerful tool that helps organizations understand the attitudes, interests, and demands of their customers. Sentiment analysis provides valuable insights into customer sentiment, competitive analysis, and crisis management.

2.6 Natural language processing

Natural language processing (NLP) is a subfield of artificial intelligence (AI) that studies the interaction between computers and humans using natural language. NLP involves analyzing, understanding, and generating human language, both written and spoken (Acheampong, Wenyu, & Nunoo-Mensah, 2020). In recent years, the field has become increasingly prominent due to the availability of vast datasets that contain user-generated textual data on social media platforms. (Pak & Paroubek, 2010).

NLP techniques have been used in various industries as can be seen in the Table 3. They have been used to identify missing or incomplete documentation in electronic health records (EHRs), enabling clinicians to improve the accuracy and completeness of medical records (Demner-Fushman, Chapman & McDonald, 2009). They have been also used to analyze EHRs and identify patterns that are predictive of conditions. Predictive

analytics can enable clinicians to take proactive measures to prevent adverse events and improve patient outcomes (Ribelles et al., 2021). NLP is also used in finance with sentiment analysis. Sentiment analysis can enable investors to make informed decisions about buying and selling securities (Mishev, Gjorgjevikj, Vodenska, Chitkushev, & Trajanov, 2020). NLP can also enable financial institutions to detect and prevent fraudulent activities, protecting investors and maintaining the integrity of financial markets (Chang, Yen, & Hung, 2022). In addition, NLP algorithms have the capability to empower chatbots to understand natural language queries and provide precise and useful responses (Shum, He, & Li, 2018). NLP has been also used with sentiment analysis in marketing to analyze customer feedback and increase customer satisfaction. For example, sentiment analysis can be used to identify common complaints or issues that customers have with a product or service, enabling companies to address these issues and improve customer satisfaction (Ramaswamy & DeClerck, 2018). Another use of NLP in marketing is chatbots. Chatbots can also be programmed to make recommendations based on customer preferences and previous purchase history (Lee, 2020).

NLP has a wide range of applications in various areas and can be used to analyze large datasets containing textual data and extract meaningful insights that can be used to improve decision-making and improve customer experiences.

Table 3. NLP Techniques and Applications in Various Industries

Industry	Application of NLP technique	Citations
Electronic Health Records	Identify missing or incomplete documentation	Demner-Fushman, Chapman & McDonald, 2009
Electronic Health Records	Analyze patterns predictive of conditions	Ribelles et al., 2021
Finance	Sentiment analysis for informed investment decisions	Mishev, Gjorgjevikj, Vodenska, Chitkushev & Trajanov, 2020
Finance	Fraud detection and prevention	Chang, Yen & Hung, 2022
Marketing	Chatbots for product recommendations	Lee, 2020
Marketing	Sentiment analysis for identifying common complaints or issues	Ramaswamy & DeClerck, 2018

2.7 XGBoost algorithm

XGBoost, or Extreme Gradient Boosting, is An increasingly popular machine learning algorithm due to its performance and flexibility (Tarwidi, Pudjaprasetya, Adytia & Apri, 2023). It is a tree-based ensemble model that combines the strengths of decision trees and gradient tree boosting algorithms (GTBs). XGBoost excels in handling sparse data and addressing instance weights during approximate tree learning and it supports parallel and distributed computing, which enhances the efficiency of decision tree construction. (Chen et al., 2016). XGBoost has demonstrated its utility in addressing complex problems across a range of domains, such as natural language processing, computer vision, and financial forecasting (Chen et al., 2016). One of the key strengths of XGBoost is its ability to handle missing data. When missing data is found, the algorithm automatically learns the optimal direction to take, resulting in more accurate predictions.

When working with huge datasets, where missing data is frequently present, this feature is extremely helpful.

Several studies have compared XGBoost to other popular machine learning algorithms, such as Random Forest and Support Vector Machines (SVM). The study by Chen and Guestrin (2016) founded that XGBoost outperformed these algorithms in terms of accuracy and speed in a variety of classification tasks. Similarly, a study by Zhao et al. (2022) founded that XGBoost outperformed majority of other algorithms in predicting financial distress in companies. Also, according to Feng, Yin, Wang, and Dhamotharan (2022) XGBoost outperformed SVM, artificial natural network (ANN), and other methods with its low cost and better handling large amounts of data.

According to a study by S. Rawat, A.S. Rawat, Kumar, and Sabitha (2021), machine learning algorithms are utilized to analyze datasets and their performance are evaluated using metrics to determine the significant factors that influence claim filing and acceptance in the insurance industry. The future scope of the study stated that XGBoost algorithm can be used to address the high imbalance problem of the data. Also, according to Lian, Gao, and Ye (2022), XGBoost algorithm is suggested in future studies to test for possible nonlinear effects in the models and to improve model prediction accuracy in the field of green credit and bank performance. Table 4 and Table 5 puts the strengths and weaknesses of XGBoost in perspective with its use in different studies.

Table 4. Strengths and Weaknesses of XGBoost

Strengths	Weaknesses	Citations
Efficient handling of sparse data	Limited interpretability of models	Chen and Guestrin (2016)
Ability to handle imbalanced data	Higher complexity compared to simpler models	Zhao et al. (2022)
Low cost	Prone to overfitting with large and noisy datasets	Feng et al. (2022)
Regularization techniques	Requires careful tuning of hyperparameters	Rawat et al. (2021)
Optimized computing performance	May not perform as well with small datasets	Lian et al. (2022)

Table 5. Strengths and Weaknesses of XGBoost in Different Studies

Study	Strengths	Weaknesses
Chen and Guestrin (2016)	Outperformed other algorithms in terms of accuracy and speed	Limited interpretability
Zhao et al. (2022)	Outperformed other algorithms in predicting financial distress	Prone to overfitting
Feng et al. (2022)	Outperformed other algorithms in handling missing data and reducing training overhead	Requires careful tuning of hyperparameters
Rawat et al. (2021)	Suggested for addressing high data imbalance	Limited interpretability
Lian et al. (2022)	Suggested for testing for nonlinear effects and improving prediction accuracy	Prone to overfitting with noisy datasets

2.8 LightGBM algorithm

LightGBM is a machine learning framework that uses gradient boosting and tree-based algorithms to model complex relationships in data and make accurate predictions. It is designed to handle large-scale datasets and provides high accuracy and efficiency in training and prediction. Due to its scalability and capacity for handling large amounts of data, LightGBM has grown in popularity among the machine learning field. (Ke et al., 2017).

One of the main strengths of LightGBM is its ability to handle large datasets with high dimensionality. According to Ke et al. (2017), LightGBM can handle datasets that have billions of records and millions of features and improving efficiency of learning a decision tree while adapting GBT framework. This is due to its ability to

partition data horizontally and vertically, which makes it possible to train models on subsets of the data in parallel. This method enhances the model's accuracy while simultaneously reducing the amount of time to train it. Efficiency is another key strength of LightGBM. The algorithm uses a histogram-based algorithm to bin continuous features, which reduces the memory required to store data and speeds up the training process. According to Ke et al. (2017), LightGBM was up to 20 times faster than other widely used gradient boosting algorithms, such as XGBoost. Given their differences, LightGBM and XGBoost can complement each other in different ways. For example, in their study, Ke et al. (2017), the authors used LightGBM and XGBoost in combination to predict the occurrence of heart disease. They founded that the two algorithms complemented each other, with LightGBM being better suited for handling categorical features and XGBoost being better at handling missing values and optimizing computing performance.

LightGBM has been used in wide range of domains like computer vision, natural language processing, and recommendation systems. According to Sun et al. (2022) LightGBM outperformed other algorithms in terms of accuracy and training time in image classification tasks. In natural language processing, LightGBM has been used for sentiment analysis, text classification, and named entity recognition. According to Bian, Ye, Zhang, and Yan (2022), LightGBM achieved state-of-the-art results in sentiment analysis and text classification tasks. LightGBM has also been used in recommendation systems to forecast user behavior and provide personalized recommendations. According to Han, Liang, Bella, Giunchiglia, and Li (2023) LightGBM outperformed other algorithms in terms of accuracy and training time in a recommendation system for

the data acquired from mobile phones. Table 6 puts the strengths and weaknesses of LightGBM in perspective.

Table 6. Strengths and Weaknesses of LightGBM

Strengths	Weaknesses	Citations
Handles large datasets with high dimensionality	Not suitable for small datasets	Ke et al. (2017)
Efficient training and prediction with histogram-based algorithm	Limited interpretability	Ke et al. (2017)
Can handle categorical features efficiently	Requires parameter tuning	Ke et al. (2017)
Complements XGBoost in handling missing values and optimizing computing performance	May overfit on imbalanced data	Ke et al. (2017)
Achieves advanced results wide range of domains like computer vision, natural language processing, and recommendation systems		Sun et al. (2022)

2.9 CatBoost algorithm

Machine learning has exploded in popularity over the past few years, and algorithms like CatBoost have emerged as powerful resources for predictive modeling. CatBoost is a machine learning algorithm that uses gradient boosting and has gained popularity due to its ability to handle categorical variables effectively (Dorogush, Ershov, & Gulin, 2018). CatBoost is an open-source gradient boosting algorithm developed by Yandex, a Russian search engine company (Prokhorenkova et al., 2018). The algorithm is designed

to deal with categorical variables, which are frequently seen in real-world datasets. The name CatBoost stands for categorical boosting. The algorithm is based on the gradient boosting framework and uses decision trees as weak learners. According to a study by Prokhorenkova et al. (2018), CatBoost outperforms other available boosting models with respect to empirical results on a variety of datasets. The main difference between CatBoost and other gradient boosting algorithms like XGBoost and LightGBM is its ability to handle categorical features without the need for one-hot encoding or label encoding. Another ability of CatBoost is that it can handle imbalanced datasets. The algorithm achieves this by introducing a new hyperparameter called "scale_pos_weight," which enables users to determine the weight of positive samples in the dataset. This is especially helpful in applications like fraud detection or anomaly identification when there are frequently fewer positive samples than negative samples.

One important application of CatBoost is in the field of natural language processing (NLP). In NLP, CatBoost has been used for sentiment analysis, text classification, and named entity recognition. For example, in a study by Zhao et al. (2022), CatBoost was used to classify sentiment tone of comments into positive and negative categories. The study founded that CatBoost outperformed other state-of-the-art algorithms like SVM, logistic regression (LR) and artificial neural network (ANN).

In the field of natural language processing, CatBoost has demonstrated promising results, particularly in sentiment analysis and text classification. According to Zhou, Li, Wang, Ding, and Xia, (2019), XGBoost and LightGBM are used to forecast the default probability of each loan for a Peer-to-Peer (P2P) lending platform and CatBoost can be included as an individual classifier in future research due to its ability to handle categorical features, combine category features leverage the relationship between

features and optimize other algorithms to calculate leaf-values which can prevent model overfitting. Overall, CatBoost is a valuable addition to the family of gradient boosting algorithms and has a promising future in the machine learning industry. Table 7 puts the strengths and weaknesses of CatBoost in perspective.

Table 7. Strengths and Weaknesses of CatBoost

Strengths	Weaknesses	Citations
Effective handling of categorical variables without encoding	Relatively new compared to other gradient boosting algorithms	Prokhorenkova et al., 2018
Ability to handle imbalanced datasets with scale_pos_weight hyperparameter	May require longer training times for larger datasets	Prokhorenkova et al., 2018
Outperforms other boosting models on a variety of datasets	May require tuning of hyperparameters for optimal performance	Prokhorenkova et al., 2018
Demonstrated success in natural language processing tasks		Zhao et al., 2022

2.10 XGBoost, CatBoost and LightGBM algorithms for improving CRM

Combination of machine learning algorithms that use gradient boosting technique like XGBoost, CatBoost, and LightGBM have gained popularity in recent years for addressing issues with customer satisfaction (Hu et al., 2023). In addition, these algorithms have demonstrated superior performance in various data mining competitions and applied research studies compared to traditional models (Imran & Amin, 2020).

With the help of these algorithms, it is possible to predict consumer satisfaction, identify the key factors that affect it, and offer suggestions to improve customer satisfaction.

XGBoost, CatBoost, and LightGBM are all gradient boosting algorithms that have gained popularity in recent years due to their high accuracy and efficiency. XGBoost was developed by Chen and Guestrin in 2016 and has won numerous machine learning competitions. CatBoost, developed by Yandex in 2017, has gained popularity due to its ability to handle categorical variables efficiently. LightGBM, developed by Microsoft in 2017, is known for its high speed and scalability. (Chen et al., 2016; Prokhorenkova et al., 2018; Ke et al., 2017) The reason for the selected algorithms is due to their raising popularity and promising performances in previous studies (Xia, Li, He, Xu, & Meng, 2021). XGBoost is selected because it can resist overfitting which can make a machine learning model more robust, CatBoost improves generalization and captures high-order dependencies which makes it a good algorithm to use among with others and LightGBM is good for its capability to handle large-scale data and its time efficiency (Abbasniya et al., 2022).

Several studies have explored the use of XGBoost, CatBoost, and LightGBM in increasing customer satisfaction. For example, Ma and Fildes (2021) used XGBoost to predict expected sales quantity for increase in customer satisfaction in the online retail industry. They founded that XGBoost outperformed other machine learning algorithms in predicting customer satisfaction. Similarly, Chen, Wang, Zhang, Wang and Peng (2021) used XGBoost and LightGBM to forecast customer purchase prediction to help managers provide better tourism services online. Zhan, Li, Jiang, Sha and Guo (2020) used CatBoost to predict sales in the e-commerce industry. They founded that CatBoost outperformed other machine learning algorithms in predicting customer sales. In

addition, Jabeur, Gharib, Mefteh-Wali and Arfi (2021) used CatBoost and other ML algorithms for corporate failure prediction and founded that CatBoost approach improved the performance. Omar, Klibi, Babai and Ducq (2023) used LightGBM to forecast demand from shopping basket data and showed that proposed methods improved the forecasting accuracy and reduced shortage.

Proposing and testing new combinations and comparing results by assembling algorithms like XGBoost, CatBoost, and LightGBM can be beneficial to future works according to Kuncheva (2014). There is limited research on combining XGBoost, CatBoost, and LightGBM, even though they have all been utilized individually to raise customer satisfaction. However, there are several studies that have explored the use of these algorithms in other applications. For example, Zhou, Fujita, Ding and Ma used XGBoost, CatBoost, and LightGBM (2021) together for credit risk modeling. They founded that using these algorithms together outperformed using them individually. Similarly, Zhou et al. (2019) used these algorithms together to predict P2P lending. Also, the study by Severinsen and Myrland (2022) states that in the competition hold by The American Society of Heating, Refrigerating and Air- Conditioning Engineers (ASHRAE) in 2019 which uses machine learning flows, the top five produced solutions all include LightGBM algorithm, three of the top five used CatBoost, and two used XGBoost. Also, in a study by Abbasniye et al. (2022) combined use of XGBoost, CatBoost and LightGBM algorithms produced the best average accuracy in breast cancer tumor classification. One of the most relevant study that combined these algorithms is about conducting customer segmentation for new product development based on online product reviews by Joung and Kim (2023). In the study, which is a contribution to customer segmentation research, the authors used machine learning classifiers to predict

positive and negative star ratings based on sentiment scores for 15 product features. and XGBoost exhibited the highest f1-score followed by CatBoost and LightGBM. Table 8 puts the strengths and weaknesses of XGBoost, LightGBM and CatBoost in perspective.

2.11 Summary

The literature review reveals how machine learning algorithms have the potential to improve customer support services and increase customer satisfaction and loyalty. Various machine learning algorithms such as XGBoost, CatBoost, and LightGBM can be used for sentiment analysis, topic modeling, personalization, and customer segmentation. XGBoost can resist overfitting which can make a machine learning model more robust, CatBoost improves generalization and captures high-order dependencies and LightGBM can handle large-scale data and its time efficiency (Abbasniya, Sheikholeslamzadeh, Nasiri, & Emami, 2022). Other machine learning algorithms used to improve customer support services and increase customer satisfaction and loyalty include Support Vector Machines (SVM) and Random Forest (RF) as well. SVMs can be computationally expensive when handling large datasets (Fine & Scheinberg, 2001). Random forest has its downside like SVM when handling large datasets and it has the potential to overfit (Breiman, 2001). Another algorithm can be K-Nearest Neighbors (KNN), which is a non-parametric algorithm used for classification and regression tasks that has been used for customer segmentation, recommendation systems, and churn prediction (H. Kim, Kim, & Chang, 2017). KNN can be also computationally expensive when dealing with large datasets (H. Kim et al., 2017). Computationally expensiveness and lack of trust with the large datasets are the reasoning for not using these algorithms in this research since the data used is quite large.

The studies discussed in this chapter highlight the effectiveness of these algorithms in identifying customer complaints related to poor customer support service and in spotting product defects. These algorithms can be used by organizations to analyze customer feedback and identify potential issues with their customer support service, allowing them to make improvements and increase customer satisfaction.

Businesses should consider integrating machine learning algorithms into their customer support services to extract valuable insights from social media data and eventually improve their customer satisfaction and relationships. By real time analyzing massive amounts of social media data, companies can identify customer sentiments, understand customer needs and preferences, and make data-driven decisions. In addition to enhancing client connections and customer satisfaction, this can increase sales and profitability. Furthermore, companies can automate their customer support service and reduce response times by leveraging machine learning algorithms, which are all critical parts for customer satisfaction. By using chatbots or virtual assistants powered by machine learning, companies can provide 24/7 customer support service, which can help customers get the help they need quickly and efficiently.

Table 8. Strengths and Weaknesses of XGBoost, LightGBM and CatBoost

Algorithm	Strengths	Weaknesses	Citations
XGBoost	- High accuracy in prediction and classification tasks - Ability to handle missing values - Ability to handle large datasets with high dimensionality	- High memory consumption - Longer training time for large datasets compared to LightGBM - Inefficient handling of categorical variables	Chen et al., 2016
LightGBM	- Ability to handle large datasets with high dimensionality - High accuracy in prediction and classification tasks - Fast training time due to its histogram-based algorithm - Efficient handling of categorical variables	- Prone to overfitting on smaller datasets - May require tuning of hyperparameters for optimal performance	Prokhorenkova et al., 2018
CatBoost	- Efficient handling of categorical variables without the need for one-hot encoding - High accuracy in prediction and classification tasks - Ability to handle imbalanced datasets	- Longer training time for large datasets compared to XGBoost and LightGBM - May require tuning of hyperparameters for optimal performance	Ke et al., 2017

However, there are challenges that need to be addressed when applying machine learning algorithms for customer support service, including data privacy and security, transparency, interpretability of machine learning models, and the necessary resources and expertise for effective implementation. Despite these challenges, the potential benefits of using machine learning in customer support service are enormous. Through the use of machine learning algorithms, companies can obtain insightful knowledge about customer sentiments, preferences, and needs, which can help them improve their products and services, improve customer satisfaction, and build long-lasting relationships with their customers. Companies that invest in this technology are likely to see significant improvements in their customer satisfaction, customer relationships, and overall business performance. Machine learning algorithms have shown great potential in improving customer support service and enhancing customer satisfaction through the analysis of social media data. Due to the growing significance of social media in customer service, businesses must make use of machine learning algorithms to get insightful knowledge about the wants and preferences of their customers.

CHAPTER 3

MATERIALS AND METHODOLOGY

3.1 Dataset

In this research, the data used were retrieved from the Twitter account of an innovative Customer Management BPO provider with 30 years of experience on a global scale, which will be referred as The Company from this point further. The tweets that were sent to The Company's account, mentioned that contain The Company's tag or retweets that contain The Company's tag were collected from the Twitter between the dates of 01/03/2022 and 31/03/2022 from the Twitter APIs. The Twitter data was collected for the operation causes of The Company and the choice holds no significance other than the operation's beginning dates since it was the most recent data available for that case, and dates can be manipulated for future operations. Then data were cleaned from spam tweets and tweets from fake accounts by The Company. The authenticity of the content of the tweet and account holder were considered and The Company made sure that real customers' tweets were included in the dataset.

Dataset, which consists of 12028 entries, contains the tweets, tweets' dates, users' name, tweets' links, main and sub topics of tweets, retweet ids and numbers; users' tweet count, follower and following numbers, location, language and gender. It also contains a label column which indicates whether the sentiments of the tweets are positive, negative or neutral with respect to their context. The sentiments of the tweets are labeled manually by The Company and the labeling of the entries are assumed to

hold no human error in the context of the study. With respect to this research's perspective, the columns that contain tweets and sentiment labels are used in the machine learning algorithms to train and test the results. As it can be seen from the Figure 1, dataset contains 7847 tweets with neutral sentiment, 89 tweets with positive sentiment and 4092 tweets with negative sentiment. This distribution portrays an imbalanced dataset, which will be affecting the evaluation performance decisions which are discussed in Chapter 4.

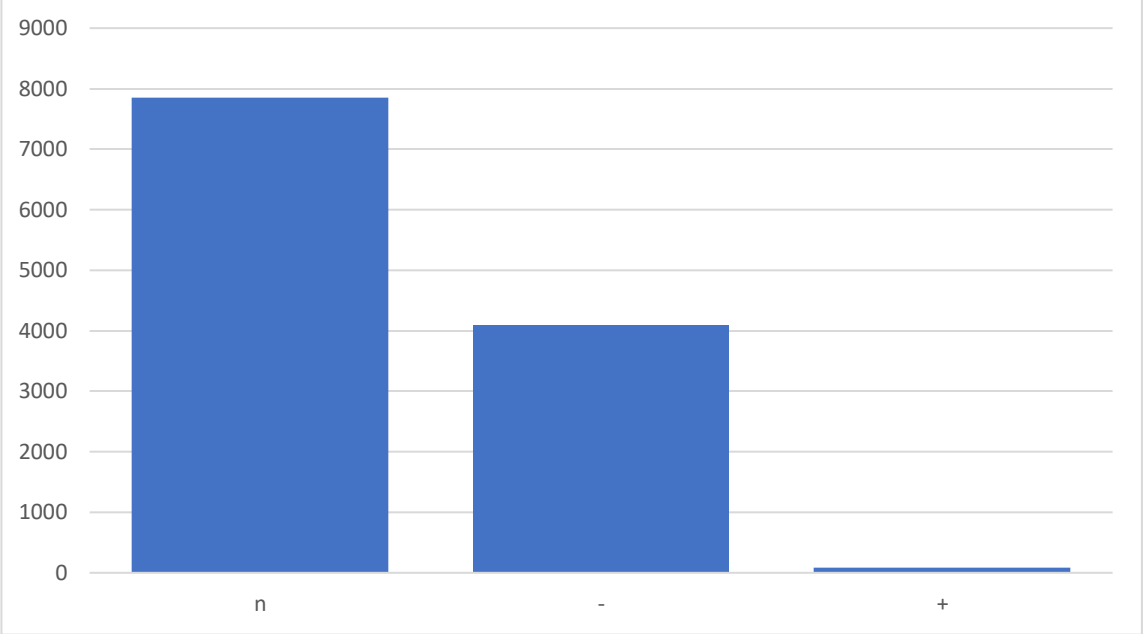

Figure 1. Distribution of negative (-), positive (+) and neutral (n) tweets in the dataset

3.2 Code availability

The source code of the machine learning algorithms that is implemented in Python programming language and used to observe the performances in this research is available in a public repository in GitHub. (see Appendix)

3.3 Methodology

Three gradient boosting techniques are used in this research. Extreme Gradient Boosting (XGBoost), Categorical Boosting (CatBoost), and Light Gradient Boosting Machine (LightGBM) are used to observe the sentiment of the dataset and their classification performances according to the prediction accuracy and F1 score are evaluated. The implementation details and mathematical formulas for XGBoost algorithm are described in 3.3.1, for LightGBM algorithm described in 3.3.2 and for CatBoost algorithm described in 3.3.3. Among with the individual performances, combination of the gradient boosting techniques that are received through the soft voting ensemble classification are evaluated. Ensemble models are created by integrating the predictions of multiple machine learning models and generally shows better results than individual models and has been used for several kinds of tasks (Yousaf et al., 2021). The details of ensemble classification can be found in 3.3.4. Also, SVM and RF algorithms are implemented and their prediction accuracy and F1 score performances are included to evaluate results since they are part of the most commonly used algorithms in sentiment analysis according to previous studies (Pang et al., 2009; Amrani et al., 2018). The details for the implementations of SVM and RF algorithms are in 3.3.5 and 3.3.6. Dataset had been divided in the two sets: one for the training processes of the algorithms and one for the testing process. Training size that had been parted from the dataset was 80% and for testing 20% of the dataset was used, which is a common ratio to prevent overfitting (Hastie, Tibshirani, & Friedman, 2008). The results of the performances show that XGBoost, LightGBM and CatBoost classifier's soft voting ensemble classification gives the best result for prediction accuracy and F1 score. From the induvial performances of XGBoost, LightGBM and CatBoost, LightGBM gives the best

result according to the prediction accuracy and F1 score. The performance evaluation of the gradient boosting techniques and ensemble classifications are further explained in Chapter 4.

3.3.1 XGBoost

The XGBoost algorithm is an optimized version of the gradient boosting algorithm, which is a powerful technique for solving various machine learning problems (Chen et al., 2016). XGBoost has become increasingly popular among data scientists and machine learning practitioners because of its ability to handle large-scale datasets, robustness against overfitting, and high prediction accuracy (Tarwidi et al., 2023). XGBoost has also been used to win several Kaggle competitions, demonstrating its effectiveness in solving complex machine learning problems. For example, in their study, Phan et al. (2018) XGBoost is used to win a Kaggle competition for predicting store sales and it showed that XGBoost was able to handle the high-dimensional and sparse nature of the data and outperformed several other popular algorithms.

The XGBoost algorithm builds an ensemble of decision trees that are trained sequentially, where each subsequent tree corrects the errors of the previous tree. The final prediction is made by aggregating the predictions of all the trees in the ensemble. The XGBoost algorithm minimizes a regularized objective function, which is a combination of a loss function and a regularization term. The loss function measures the discrepancy between the predicted and actual values, while the regularization term controls the complexity of the model and helps prevent overfitting. The objective function can be expressed as follows (Chen et al., 2016):

$$L(\emptyset) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

(1)

where $L(\emptyset)$ is the objective function, i is the model parameters, *l* is the loss function that measures the difference between the predicted values $\hat{y}_i$ and the true labels y_i, K is the number of trees in the ensemble, and $\Omega(f_k)$ is the regularization term that penalizes complex models.

The loss function *l* can be chosen depending on the problem, and common choices include the logistic loss for binary classification, the softmax loss for multiclass classification, and the squared error loss for regression. For the purpose of this research, softmax loss for multiclass function from the XGBoost library is used since the dataset contains multiclass data, which are negative, positive and neutral.

To train the XGBoost model, the algorithm optimizes the objective function by adding decision trees one at a time. At each iteration, the algorithm finds the optimal split point for each leaf node in the tree based on the information gain, which measures the reduction in entropy or impurity achieved by splitting the node. The algorithm then prunes the tree using a greedy approach, keeping only the branches that lead to a reduction in the objective function. The algorithm continues to add trees until the objective function cannot be further optimized or a stopping criterion is met. In the study by Chen et al. (2016) authors introduced XGBoost's level-wise tree growth strategy, which is one of its major strengths and it can improve both accuracy and speed compared to other tree-based algorithms. XGBoost's decision tree that grows by level wise can be found in Figure 1.

3.3.2 LightGBM

LightGBM is a powerful gradient boosting algorithm that employs an innovative technique called Gradient-based One-Side Sampling (GOSS) to accelerate the training process and minimize memory consumption. GOSS works by retaining only the instances with large gradients during the sampling process, which reduces the number of instances that need to be evaluated during the training process (Ke et al., 2017). In addition, LightGBM uses a technique called Exclusive Feature Bundling (EFB), which groups the features into bundles based on their importance, and then performs the split on the bundles rather than on individual features (Islam et al., 2019). The objective function for multi-class classification is defined as (Ke et al., 2017):

$$L(y, f(x)) = -\sum_{i=1}^{n}\sum_{k=1}^{K} y_{ik} log(\hat{p}_{ik}) \quad (2)$$

where y_{ik} is the true label of the i-th sample for the k-th class, $\hat{p}_{ik}$ is the predicted probability of the sample belonging to the k-th class, and f(x) is a matrix of predicted probabilities for each class. Loss function for LightGBM is (Ke et al., 2017):

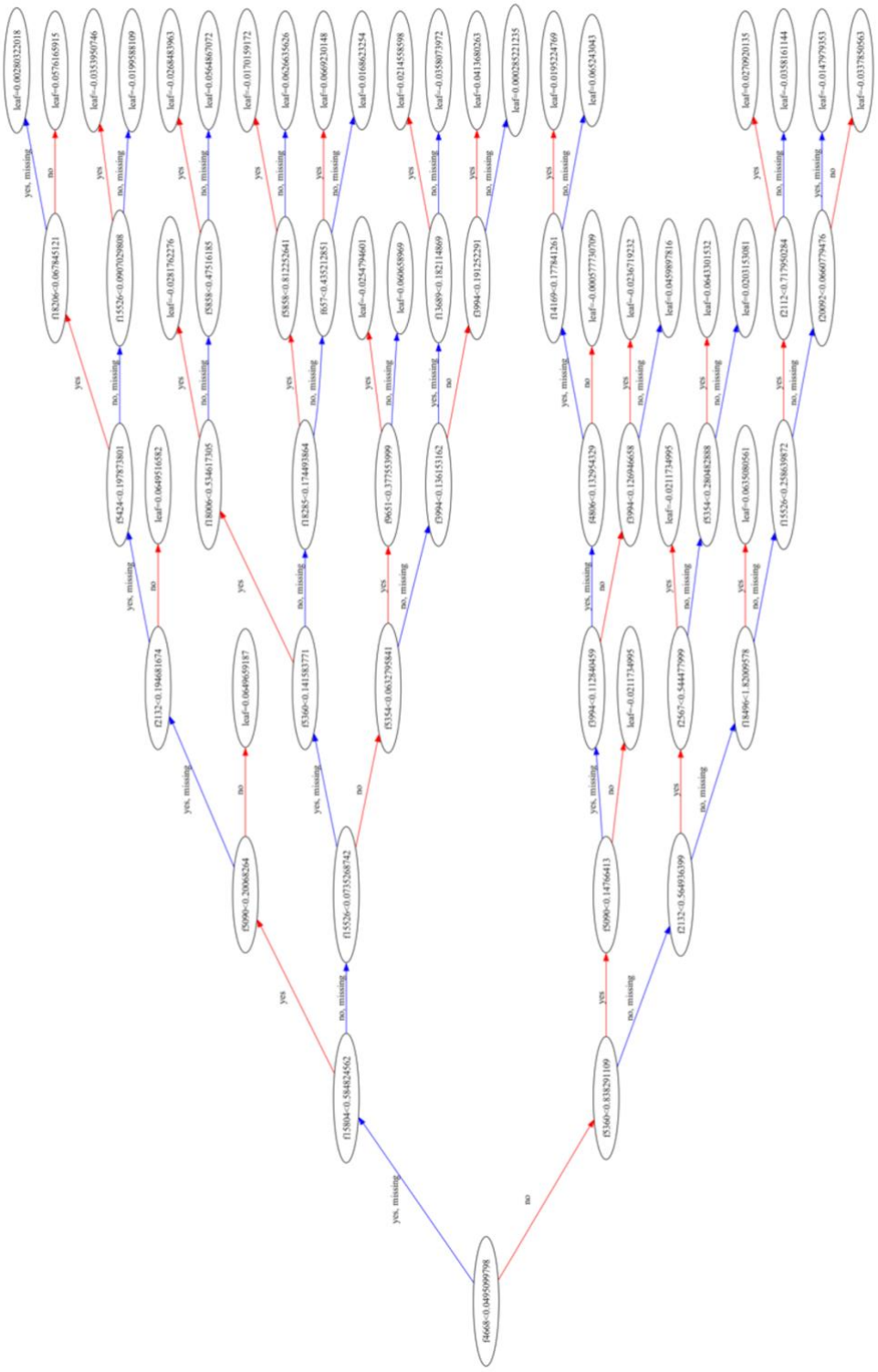

Figure 2. XGBoost's decision tree

$$L(y, f(x)) = \sum_{i=1}^{n} \omega_i l(y_i, f(x_i)) + \sum_{k=1}^{K} \Omega(f_k)$$

(3)

where n is the number of instances, y_i is the true label of the i-th instance, f(x) is the feature vector of the i-th instance, and $f(x_i)$ is the predicted score of the i-th instance. In this research, multi_logloss loss function from the LightGBM's library is used for its compatibility of multiclass objective since the dataset contains of multiclass which are negative, positive and neutral. Also, in their study, Ke at al. (2017) introduced LightGBM and its leaf-wise tree growth strategy which can significantly improve both accuracy and speed compared to other gradient boosting algorithms, especially when dealing with large-scale datasets. Leaf-wise decision tree of LightGBM can be seen in Figure 3.

3.3.3 CatBoost

CatBoost is a gradient boosting algorithm that uses a novel approach to handle categorical features, known as ordered boosting, and incorporates several optimization techniques, such as Newton's method and ordered subsampling, to increase the accuracy and speed of the model. The CatBoost algorithm uses ordered boosting to handle categorical features. This involves sorting the categorical features and splitting them into contiguous intervals that correspond to different classes. The formula for the split gain of a categorical feature is (Prokhorenkova et al., 2018):

$$Split\ Gain = \left|\frac{Gradient\ sum}{Hessian\ sum + \lambda}\right| - |\lambda|$$

(4)

where Gradient sum is the sum of the gradients of the loss function for each instance in the split, Hessian sum is the sum of the Hessians (second derivatives) of the loss function for each instance in the split, lambda is the regularization parameter, and the absolute values ensure that the gain is always positive. The loss function used in CatBoost is a weighted sum of instance-wise losses and a regularization term and it is the same as Equation 3.

The CatBoost algorithm is based on the gradient boosting framework, which trains an ensemble of weak learners, typically decision trees, to make predictions on a given dataset. The goal of gradient boosting is to minimize a loss function, which measures the difference between the true labels and the predicted scores of the ensemble (Prokhorenkova et al., 2018). For this study, MultiClass loss function from the CatBoost library is used for its compatibility of multiclass objective since the dataset contains of multiclass which are negative, positive and neutral. Also in their study, Prokhorenkova et al. (2018) introduced CatBoost and its depth-wise tree growth strategy, which can significantly improve both accuracy and speed compared to other gradient boosting algorithms, especially when dealing with categorical features. Depth-wise decision tree of CatBoost algorithm can be found in Figure 4.

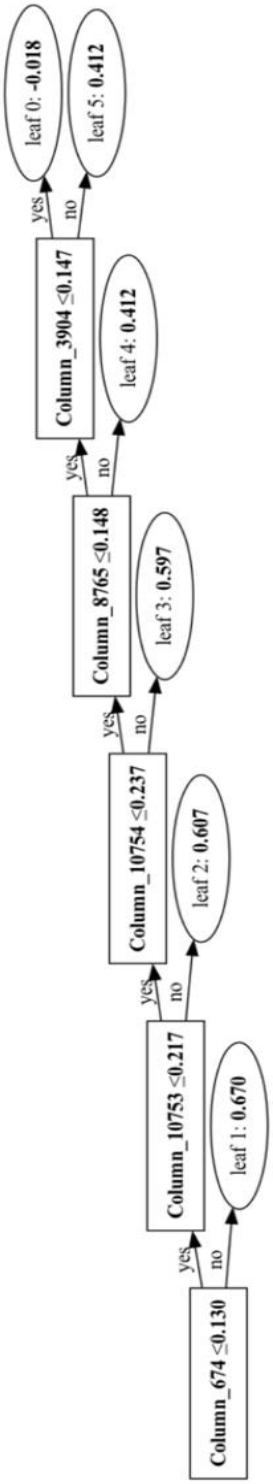

Figure 3. Decision tree of LightGBM

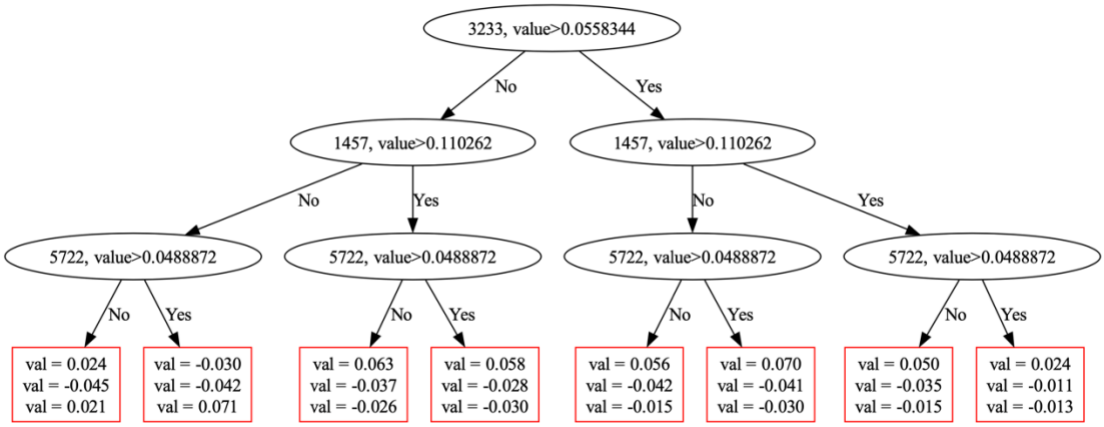

Figure 4. Decision tree of CatBoost

3.3.4 Ensemble classification

Ensemble classification is a machine learning technique that combines multiple models to improve accuracy and generalization (Dietterich, 2000). Polikar (2012) founded that ensemble methods have gained popularity in recent years due to their ability to mitigate overfitting, reduce bias and variance, and improve the performance of machine learning models. The main idea behind ensemble classification is to leverage the strengths of different models and mitigate their weaknesses by combining them (Zhang & Ma, 2012). Ensemble methods can be broadly categorized into two categories: homogeneous and heterogeneous. In homogeneous ensembles, the base models are of the same type, while in heterogeneous ensembles, the base models are of different types (Rokach, 2010).

Voting techniques are used in ensemble classification to make the final prediction by aggregating the predictions of multiple individual models. In hard voting, the final prediction is made by taking the majority vote of all individual models. In soft

voting, the final prediction is made by averaging the predicted probabilities of all individual models. In hard voting, the final prediction y(x) is given by (Bramer., 2016):

$$y(x) = argmax(\sum_{I=1}^{m} f_i(x) == j)$$

(5)

where j is the label being considered.

In soft voting, the final prediction y(x) is given by:

$$y(x) = \sum_{I=1}^{m} p_i j$$

(6)

where $p_i j$ is the predicted probability of the i-th model for the label j.

In this research, ensemble classifications of XGBoost and CatBoost; XGBoost and LightGBM; LightGBM and CatBoost and LightGBM, XGBoost and CatBoost are used to evaluate different combinations of the algorithm. Soft voting is used since soft voting reported to have performed better than hard voting in previous studies. For example, Bramer (2016) compared the performance of hard voting and soft voting in sentiment analysis tasks and founded that soft voting generally outperformed hard voting in terms of accuracy and F1-score.

3.3.5 Support Vector Machines

Support Vector Machines (SVMs) are a powerful and widely used algorithm in the field of machine learning and natural language processing, especially for sentiment analysis tasks (Pang et al., 2009). SVMs are particularly useful in binary classification tasks, where the aim is to classify instances into one of two classes and work by finding the

hyperplane that maximally separates the classes, and then uses this hyperplane to classify new instances (Cortes et al., 1995). The cost function used in SVMs for multiclass classification is the multiclass hinge loss function, which is defined as:

$$L(x) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j \neq y_i} max(0, w_j^T x_i - w_{y_i}^T x_i + \Delta) \quad (7)$$

where w is the weight vector, x is the input feature vector, y is the true class label, n is the number of training examples, and Δ is the margin parameter. The margin parameter determines the degree of separation between the decision boundary and the data points.

For this research, objectivity poly for multiclass implementation is used since the dataset contains more than two classes: negative, positive and neutral.

3.3.5 Random Forest

Random Forest (RF) is a popular ensemble learning algorithm for classification tasks that utilizes multiple decision trees to improve the accuracy and robustness of the model (Breiman, 2001). RF is particularly effective for classification problems, where the goal is to assign a given input to one of multiple possible classes (Amrani et al., 2018). The final prediction of the RF is made by taking the majority vote of the predictions from all the M decision trees (Breiman, 2001):

$$y(x) = argmax(\sum_{m=1}^{M} f_{k_m}(x) == k) \quad (8)$$

where y(x) is the predicted class label of the input feature vector x and $f_{k_m}(x)$ is the predicted class label of the input feature vector x by the m-th decision tree Tk.

For this research, objectivity entropy for multiclass implementation is used since the dataset contains more three classes: negative, positive and neutral.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Evaluation metrics

Machine learning algorithms have a workflow (see Figure 5) where the dataset is divided into training and test sets. Then the machine learning model is trained and finally trained model is being put to test with input data to make predictions (Osman, 2019). Improvements and necessary adjustments can be done in this flow, however, when the result state is reached for a requirement for the evaluation metrics to properly evaluate the results and performances of the algorithms emerges.

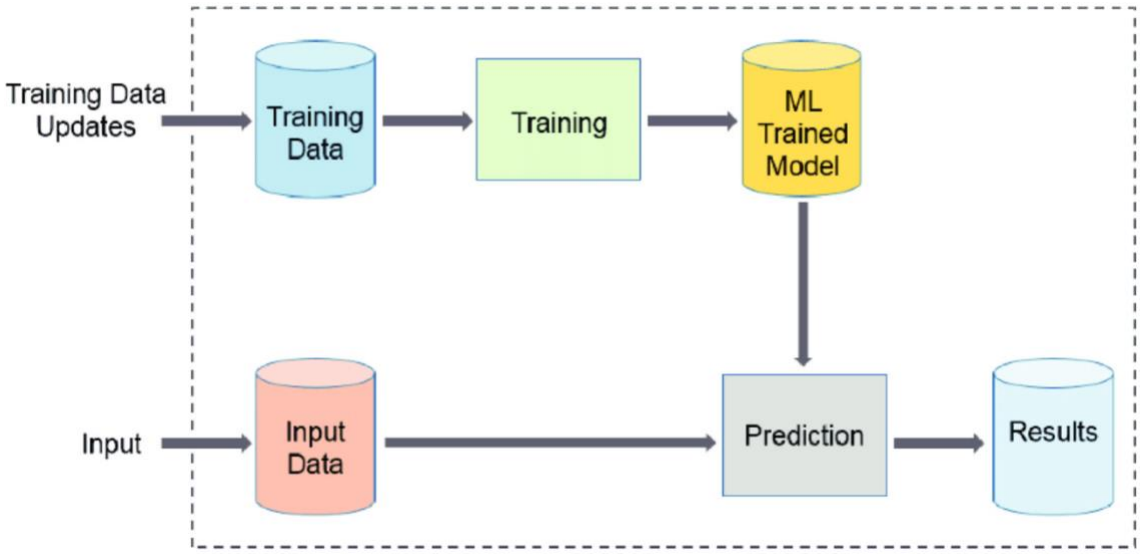

Figure 5. Machine learning flow
Source: Radiation Oncology in the Era of Big Data and Machine Learning for Precision Medicine, 2019

There are various evaluation metrics used in machine learning, each with its strengths and weaknesses depending on the type of problem and dataset. For instance,

accuracy, precision, recall, F1 score, area under the receiver operating characteristic curve (AUC-ROC) and mean squared error (MSE) are some commonly used evaluation metrics in the previous studies (Buda, Maki, & Mazurowski, 2017). Buda et al. (2017) evaluated various metrics such as accuracy, precision, recall, F1 score, and area under the receiver operating characteristic (ROC) curve.

Confusion matrix is a widely used tool used to evaluate the performance of machine learning algorithms (Xu, Zhang, & Miao, 2020). The matrix provides a way to visualize the accuracy of a model by comparing its predictions against the actual outcomes of the data. As it can be seen from Figure 6, confusion matrix consists of four elements: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). TP and TN represent the correct predictions, while FP and FN represent the incorrect predictions.

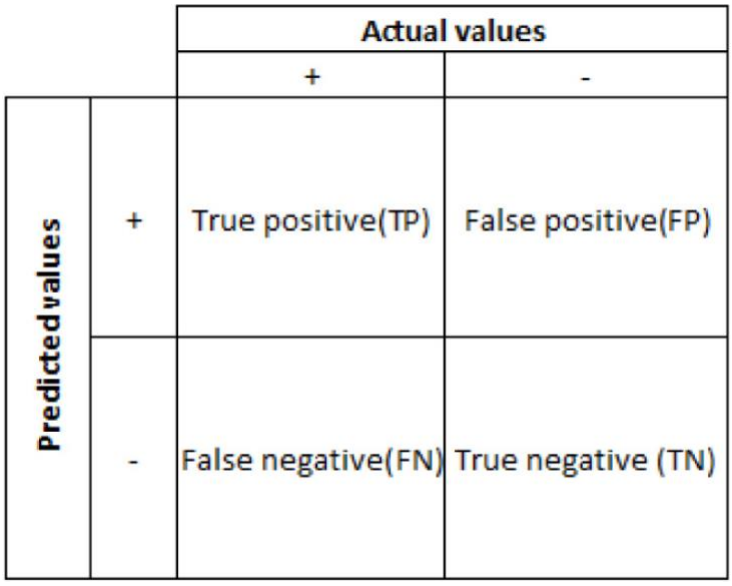

Figure 6. Confusion matrix
Source: Comparing two SVM models through different metrics based on the confusion matrix, 2023

In this study, prediction accuracy and F1 score with the help of the confusion matrix are used for the performance evaluations.

Prediction accuracy is evaluated by:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (7)$$

F1 score is evaluated by:

$$F1\ Score = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad (8)$$

Prediction accuracy is a popular evaluation metric for its ease of implementation and understanding (Kotsiantis, Kanellopoulos, & Pintelas, 2006). When dealing with imbalanced data, prediction accuracy may not be a good choice since it can mislead to a higher score, in fact in the study by Kübler, Liu, and Sayyed (2018), authors suggested that the accuracy can be misleading in the case of imbalanced datasets. Accordingly, other metrics, such as precision, recall, and F1 score, can offer a more robust evaluation of the classifier's performance. For example, if the dataset contains 90% of negative entries and 10% of the positive entries, the model can favor the majority class and label all the entries as negatives. This will give the prediction score of 0.90 which is high, however, it would be misleading since when the dataset contains 60% of negative entries and 40% of the positive entries the score can drop down to 0.6. In order to obtain a more accurate and reliable assessment of the classifier's performance, F1 score is often incorporated into the evaluation metrics of gradient boosting algorithms. F1 score is more preferred when dealing with imbalanced datasets since it takes into account both

precision and recall and can be a better measure of performance than accuracy (Sokolova & Lapalme, 2009). Precision measures the proportion of true positives (TP) to the total number of predicted positives (TP + FP) and measures the ability of a model to accurately identify positive instances, without including any false positives (Japkowicz & Shah, 2011). Recall measures the proportion of true positives to the sum of true positives and false negatives (FN) and measures how well a model can correctly identify all positive instances, without missing any (Galar, Fernández, Barrenechea, Bustince, & Herrera, 2012). In the study by Alakus and Turkoglu (2020), F1 score and accuracy score were compared in detecting cancer cases and they founded that accuracy can be misleading because it penalizes false positives and favors false negatives.

4.2 Results

In this research, the performances of three gradient boosting algorithms, namely XGBoost, CatBoost, and LightGBM, are compared in the task of sentiment analysis on Twitter data with respect to prediction accuracy and F1 score. Then, the performances of their ensemble classifications using soft voting techniques are evaluated with respect to prediction accuracy and F1 score. Additionally, SVM and RF algorithms are included for a compresence with the related studies.

As it can be seen from the Table 9, LightGBM algorithm gave the best result with prediction accuracy 0.8096 and F1 score 0.8096 among the individual performances of the gradient boosting algorithms. Ensemble classification of XGBoost, LightGBM and CatBoost algorithms gave the best result with prediction accuracy 0.8707 and F1 score 0. 8707 among all the algorithms and ensembles. The only

difference between F1 score and prediction accuracy is seen on LightGBM's performance with 0.0000000000000001.

Table 9. Performance Scores of XGBoost, LightGBM, CatBoost and Their Ensemble Classifications along with Performance Scores of SVM and RF

Algorithm	F1 Score (x100)	Accuracy Score (x100)	Average (x100)
XGBoost	79.01	79.01	79.01
LightGBM	80.96	80.96	80.96
CatBoost	77.56	77.56	77.56
XGBoost + CatBoost	86.53	86.53	86.53
CatBoost + LightGBM	87.03	87.03	87.03
XGBoost + LightGBM	86.95	86.95	86.95
XGBoost + LightGBM + CatBoost	87.07	87.07	87.07
Support Vector Machines (SVM)	66.13	66.13	66.13
Random Forest (RF)	68.29	68.29	68.29

In the implementation of the LightGBM algorithm, multiclass objective is used with multi_logloss loss function since the classification of this research requires more than two classes; negative, positive and neutral. The depth of the decision tree is set to 3 and learning rate is set to 0.05. Previous studies show that using shallow trees result in better performance. For example, in their study, Chen et al. (2016) it was founded that using shallow trees (depth 2-4) in their gradient boosting algorithm resulted in better performance than using deep trees (depth 8-16). Keeping the learning rate low provides a smooth learning for the model For example, Buda et al. (2018) showed that keeping the learning rate low in gradient boosting algorithms improved the performance of the model.

The confusion matrix received from the LightGBM algorithm can be seen on the Figure 7. The confusion matrix consists of 611 true negatives, 0 true positive and 1337 true neutrals among the 2406 entries from the test dataset, which is 20% of the original dataset. LightGBM algorithm has the most true predictions among the individual performances of other gradient boosting algorithms.

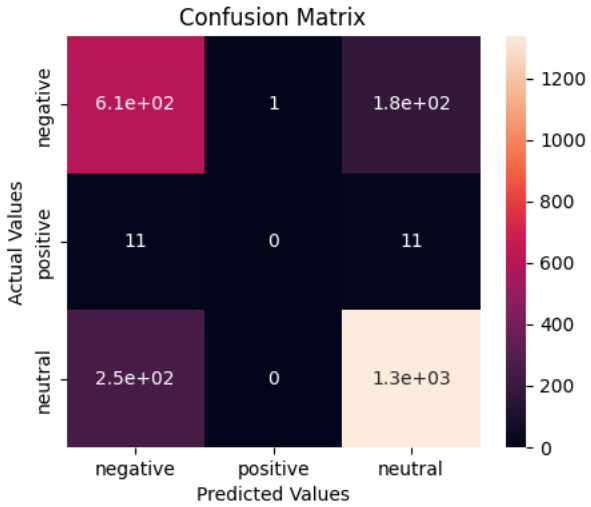

Figure 7. Confusion matrix of LightGBM

In the implementation of the XGBoost algorithm multiclass objective and multi:softmax loss function is used since the classification of this research requires more than two classes; negative, positive and neutral. The depth of the decision tree is set to 3 and learning rate is set to 0.05 for the consistency between gradient boosting algorithms. XGBoost's prediction accuracy and F1 score is the second best among the individual performances of the algorithms with 0.7901.

The confusion matrix received from the XGBoost algorithm can be seen on the Figure 8. The confusion matrix consists of 697 true negatives, 1 true positive and 1203 true neutrals among the 2406 entries.

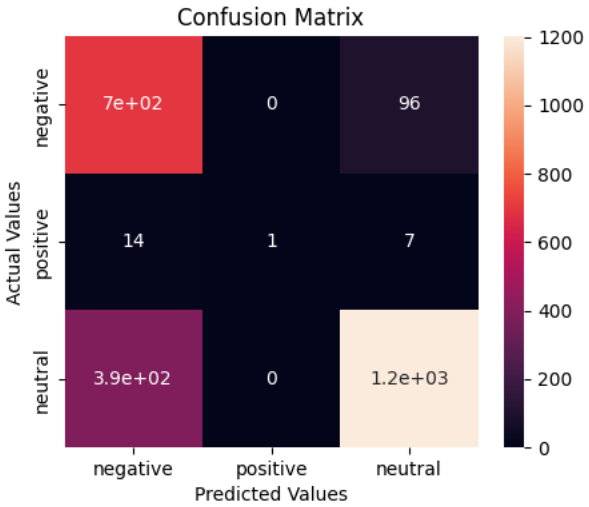

Figure 8. Confusion matrix of XGBoost

In the implementation of the CatBoost algorithm multiclass objective and loss function since in this research there are three classes; negative, positive and neutral. The depth of the decision tree is set to 3 and learning rate is set to 0.05 for the consistency between gradient boosting algorithms. CatBoost's prediction accuracy and F1 score is the worst among the XGBoost, LightGBM and ensembled performances of these three algorithms with 0.7756.

The confusion matrix received from the CatBoost algorithm can be seen on the Figure 9. The confusion matrix consists of 354 true negatives, 0 true positive and 1512 true neutrals among the 2406 entries.

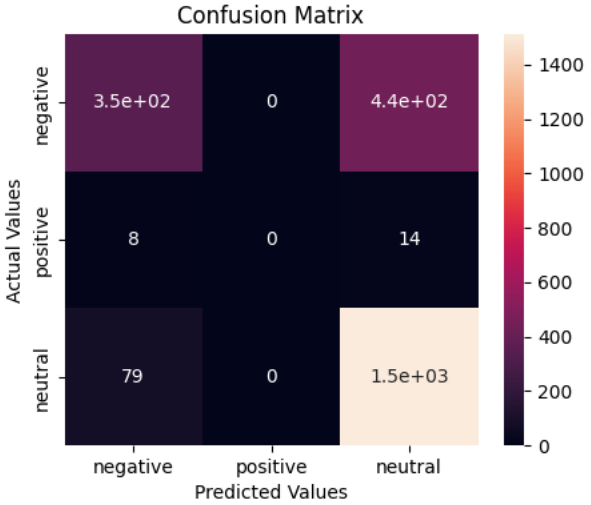

Figure 9. Confusion matrix of CatBoost

For the ensemble classifications, same models and implementations that have been mentioned are used. Ensemble classification of XGBoost, LightGBM and CatBoost gave the best prediction accuracy and F1 score with 87.07 among all the performances of algorithms. Confusion matrix for the classification can be found on Figure 10. The confusion matrix consists of 660 true negatives, 0 true positive and 1435 true neutrals among the 2406 entries.

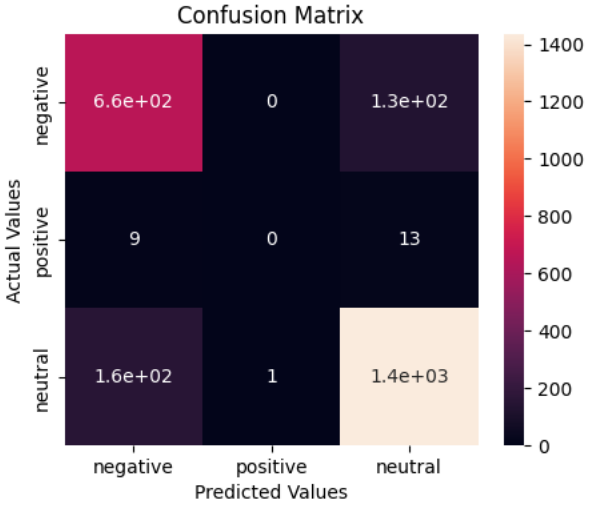

Figure 10. Confusion matrix of ensemble classification of XGBoost, LightGBM and CatBoost

Ensemble classification of LightGBM and CatBoost gave the second best prediction accuracy and F1 score with 87.03 among all the performances of algorithms. Confusion matrix for the classification can be found on Figure 11. The confusion matrix consists of 652 true negatives, 1 true positive and 1441 true neutrals among the 2406 entries.

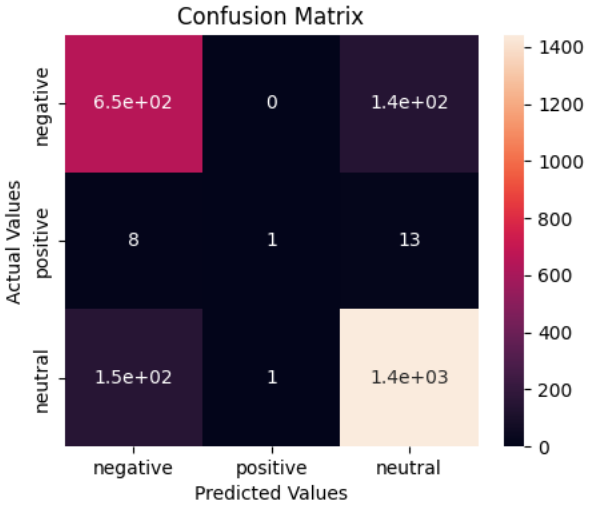

Figure 11. Confusion matrix of ensemble classification of LightGBM and CatBoost

Ensemble classification of LightGBM and XGBoost gave the third best prediction accuracy and F1 score with 86.95 among all the performances of algorithms. Confusion matrix for the classification can be found on Figure 12. The confusion matrix consists of 658 true negatives, 1 true positive and 1433 true neutrals among the 2406 entries.

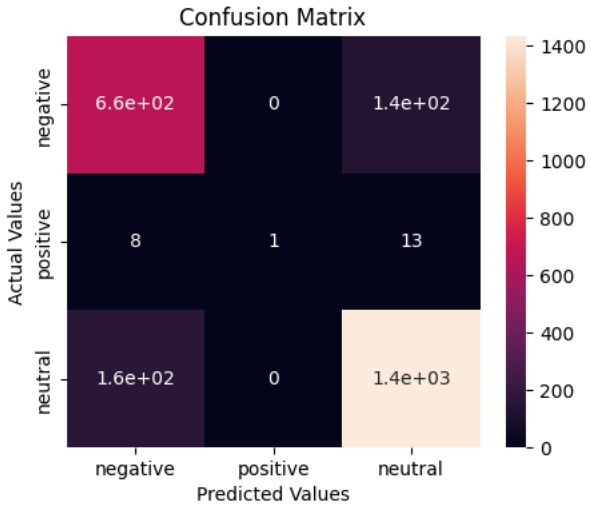

Figure 12. Confusion matrix of ensemble classification of LightGBM and XGBoost

Ensemble classification of CatBoost and XGBoost gave the fourth best prediction accuracy and F1 score with 86.53 among all the performances of algorithms. Confusion matrix for the classification can be found on Figure 13. The confusion matrix consists of 651 true negatives, 1 true positive and 1430 true neutrals among the 2406 entries.

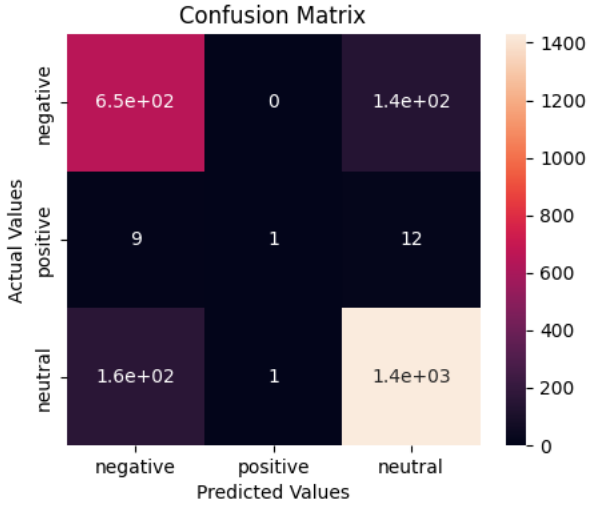

Figure 13. Confusion matrix of ensemble classification of CatBoost and XGBoost

In the implementation of the SVM algorithm poly objective is used since in this research there are mutiple classes; negative, positive and neutral. The depth of the decision tree is set to 3 and learning rate is set to 0.05 for the consistency between gradient boosting algorithms. SVM's prediction accuracy and F1 score is the worst among the algorithms with 0. 6613.

The confusion matrix received from the SVM algorithm can be seen on the Figure 14. The confusion matrix consists of 0 true negatives, 0 true positive and 1591 true neutrals among the 2406 entries.

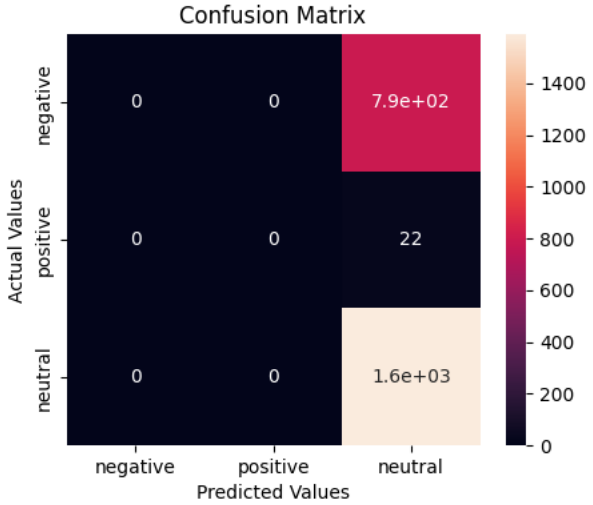

Figure 14. Confusion matrix of SVM

In the implementation of the RF algorithm entropy objective is used since in this research contains three classes; negative, positive and neutral. The depth of the decision tree is set to 3 and learning rate is set to 0.05 for the consistency between gradient boosting algorithms. RF's prediction accuracy and F1 score is the second worst among the algorithms with 0. 6829.

The confusion matrix received from the SVM algorithm can be seen on the Figure 15. The confusion matrix consists of 52 true negatives, 0 true positive and 1591 true neutrals among the 2406 entries.

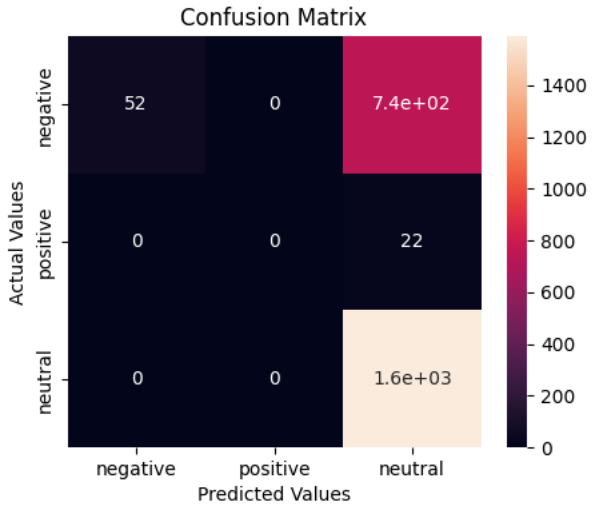

Figure 15. Confusion matrix of RF

After the initial results, K-fold Cross Validation is used as a resampling technique to observe the steadiness and improve the performance of the gradient boosting models (Bui et al., 2020; Chakraborty et al., 2020). Cross-validation is predominantly used in machine learning models to assess the performance of a machine learning model on new, unseen data (Kohavi, 1995). In the K-fold Cross Validation, dataset is shuffled randomly and divided into K subsets, then, (K-1) subsets are used for the training purpose whereas one subset is used for the test purposes. The selection of the K value can affect the performance of the models and it is suggested to choose large values of K like 5, 10 or 20 to discover more combinations in the training process (Anguita, Ghelardoni, Ghio, Oneto, & Ridella, 2012). Also, in their study, Guyon, Saffari, Dror, and Cawley (2010) stated that K values such as 5 or 10 is used by practitioners regardless of the complexity of the problem. In this research, 10-fold Cross Validation is used on the dataset to assess the performances of the XGBoost, LightGBM, CatBoost gradient boosting algorithms and their ensemble classification models since

this K value yields lower bias and variance in the test results (James, Witten, Hastie, & Tibshirani, 2013). The performance results with respect to accuracy and F1 scoring after the 10-fold Cross Validation can be seen in Table 10 and Table 11.

Table 10. Performance Scores of XGBoost, LightGBM, CatBoost and Their Ensemble Classifications with 10-fold Cross Validation (Fold 1 – Fold 5)

Algorithm	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
	Accuracy & F1 Score (x100)	Accuracy & F1 Score (x100)	Accuracy & F1 Score (x100)	Accuracy & F1 Score (x100)	Accuracy & F1 Score (x100)
XGBoost	79.88	81.23	80.05	80.71	81.38
LightGBM	80.13	80.80	79.72	82.96	80.71
CatBoost	76.14	76.64	75.81	78.30	75.31
XGBoost + CatBoost	86.53	87.78	86.53	87.70	85.79
CatBoost + LightGBM	86.62	87.70	85.54	88.28	86.20
XGBoost + LightGBM	87.03	87.28	85.70	87.45	86.12
XGBoost + LightGBM + CatBoost	86.53	87.86	85.95	88.03	86.04

Table 11. Performance Scores of XGBoost, LightGBM, CatBoost and Their Ensemble Classifications with 10-fold Cross Validation (Fold 6 – Fold 10)

Algorithm	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10	Average (x100)
	Accuracy & F1 Score (x100)	Accuracy & F1 Score (x100)	Accuracy & F1 Score (x100)	Accuracy & F1 Score (x100)	Accuracy & F1 Score (x100)	
XGBoost	82.38	81.80	78.80	81.20	83.36	81.07
LightGBM	79.97	82.30	79.05	79.37	80.95	80.60
CatBoost	75.06	77.39	75.31	76.29	75.96	76.22
XGBoost + CatBoost	85.70	87.53	85.87	88.02	89.35	87.08
CatBoost + LightGBM	85.45	88.28	86.78	86.86	88.94	87.06
XGBoost + LightGBM	85.37	88.03	86.12	87.10	89.18	86.94
XGBoost + LightGBM + CatBoost	85.79	87.95	86.28	87.27	89.19	87.09

As it can be seen from Table 10 and Table 11 that XGBoost algorithm gave the best result with prediction accuracy and F1 score 0. 8336 in the Fold 10 with the average prediction score of 0.8107 among the individual performances of the gradient boosting algorithms. Standard deviation, which provides information about the consistency or stability of the model's performance when evaluated on different subsets of the data, of the 10-Fold Cross Validation for the XGBoost model is 1.24% (Kohavi, 1995). The standard deviation of the 10-Fold Cross Validation for the LightGBM model is 1.18% and for the CatBoost model is 0.96%.

Ensemble classification of XGBoost, LightGBM and CatBoost algorithms gave the best average result with respect to prediction accuracy and F1 score with 0.8709 among all the algorithms and ensembles. The standard deviation of the 10-Fold Cross Validation for the ensemble classification of XGBoost, LightGBM and CatBoost models is 1.08%. The highest accuracy and F1 score with the 0.8935 is seen in the Fold 10 with the ensemble classification of XGBoost and CatBoost. The standard deviation of the 10-Fold Cross Validation for the ensemble classification of XGBoost and CatBoost models is 1.13%. The standard deviations of the 10-Fold Cross Validation for the ensemble classification of CatBoost and LightGBM is 1.13%, for the ensemble classification of XGBoost and LightGBM is 1.09%.

4.3 Discussion

In today’s highly competitive market, companies must prioritize establishing enduring bonds with their customers in order to drive loyalty and repeat transactions. Customer relationship management (CRM) is a crucial aspect of establishing and maintaining

these relationships. The focus of CRM is to create a customer-centric business culture that encourages sustained and mutually satisfying engagements with customers (Gummesson, 2017). However, integrating data from different sources and ensuring customer engagement are two of the significant challenges that businesses face while implementing CRM systems (Anshari et al., 2019). To overcome these challenges, companies can employ various strategies, such as utilizing social media to read customer reviews and position themselves as thought leaders in their field, providing special discounts to loyal customers, and using machine learning algorithms to improve customer service and forecast user behavior. Machine learning algorithms can analyze large volumes of customer data, spot patterns and trends, and offer personalized recommendations. For instance, by using ML-based chatbots, businesses can automate customer inquiries, reducing response time, and improving customer satisfaction (Gnewuch et al., 2018). Additionally, businesses can personalize the customer experience by offering customized recommendations, forecasting customer churn, and examining customer feedback. By employing these techniques, companies can establish a customer-centric culture that prioritizes long-term relationships with customers, resulting in increased customer satisfaction, loyalty, and profitability.

Effective customer service is crucial for any organization's operations as it assists and guides customers regarding the goods or services that the business provides. It plays a critical role in customer retention and acquisition, and it has evolved with the advancement of technology (Gupta et al., 2010; Barnes et al., 2009). Machine learning techniques can be used to enhance customer support. Machine learning algorithms can analyze customer inquiries, provide automated responses, and even identify common customer issues. Machine learning can also personalize customer interactions by

analyzing customer data, making recommendations for products or services that may be of interest to the customer, and predicting customer needs. Customer support also helps organizations locate and resolve issues with their goods or services (Rust et al., 2004). By analyzing customer support data such as chat logs, emails, and call center transcripts, machine learning algorithms can identify patterns and trends in customer issues (Loebbecke et al., 2015). This information can help organizations fix problems quickly and build a stronger emotional connection with their customers, driving loyalty and growth. Furthermore, predictive analytics can be used to forecast customer behavior, such as when they are likely to make a purchase.

This research aims to contribute to the growing use of machine learning algorithms in customer support and provide practical insights for organizations seeking to improve their customer support services. In this context, machine learning algorithms XGBoost, CatBoost, LightGBM and ensemble classification of these algorithms are used to analyze data coming from Twitter account of a global service provider as a case and provide insights into customer behavior. SVM and RF algorithms are also used to analyze the sentiment of the Twitter data. XGBoost, CatBoost, and LightGBM gradient boosting algorithms are used due to their high prediction accuracy and their good performances on training (Abbasniya et al., 2022). SVM and RF algorithms are used due to their popularity for sentiment analysis tasks among the previous studies (Pang et al., 2009; Amrani et al., 2018). Also, ensemble classification with soft voting is included in this research to improve the model's predictivity and performance.

The ensemble classification of XGBoost, LightGBM and CatBoost was founded best in performance with prediction accuracy and F1 score 0.8707, which is in accordance with the results if Abbasniya et al.'s study (2022). This could be because of

ensemble classification's ability to mitigate overfitting, reduce bias and variance, and improve the performance of machine learning models (Polikar, 2012). In individual performances, LightGBM gradient boosting algorithm outperformed XGBoost and CatBoost with prediction accuracy and F1 score 0.8096, which is consistent with the Ke et al.'s study (2017). This could be because of LightGBM improves efficiency of learning a decision tree while adapting gradient boosting tree (GBT) framework which increase accuracy performance (Ke et al., 2017). RF algorithm outperformed SVM algorithm which is in contradiction with Amrani et al.'s (2018) study where they found SVM algorithm made higher correct predictions in their study. This could be because RF algorithms improve the performance in case of small texts where SVM algorithm improves the performance for large reviews. Gradient boosting algorithms presented significantly higher prediction accuracy and F1 score than SVM and RF algorithms. This is parallel with Zhao et al.'s (2022) study where they compared XGBoost, CatBoost and SVM algorithms for predicting financial distress by combining sentiment tone features and found that XGBoost and CatBoost gave better results than SVM. This can be because of the ability that gradient boosting algorithms are handle large scale data sets and express nonparametric and nonlinear relationships (Wang, Chen, & Chu, 2018). Also, in their study, Satrya, Aprilliyani, and Yossy (2023) showed that XGBoost outperformed RF algorithm in terms of accuracy for sentiment analysis task. Finally, in their study, Zhou et al. (2019) compared gradient boosting algorithms with benchmark models for default probability loan prediction for a P2P lending platform and founded that they outperform benchmark models like RF and SVM. The problem with the more traditional algorithms like SVM can be is that they look the occurrence frequency rather than word context (Kiran, Kumar, & Bhasker, 2020).

The only difference between F1 score and prediction accuracy becomes apparent when observing LightGBM's performance, where there is a very small distinction (0.0000000000000001). Also, it can be seen from the confusion matrixes of algorithms that the model with highest accuracy, which is ensemble classification of XGBoost, LightGBM and CatBoost, also holds the highest score for predicting the true positives, true negatives and true neutrals which is consistent with the prediction accuracy and F1 scores. In the implementation phases of the algorithms, multiclass objective is used with appropriate loss function since the classification of this research requires more than two classes: negative, positive and neutral. The depth of the decision trees is set to 3 and learning rates are set to 0.05 for all the models. Previous studies show that using shallow trees and low learning rates result in better performance (Chen et al., 2016; Buda et al., 2018). For the ensemble classifications, the same models and implementations that have been mentioned are used and the model is selected with soft voting. Ensemble classification of LightGBM and CatBoost gave the second best prediction accuracy and F1 score with 87.03 and ensemble classification of LightGBM and XGBoost gave the third-best prediction accuracy and F1 score with 86.95 among all the performances of algorithms.

Using 10-Fold Cross Validation increased accuracy and F1 scores of all the gradient boosting algorithms and their ensemble classifications except from the LightGBM algorithm. Increase in the prediction performance after 10-fold Cross Validation is in parallel with the Malakouti's study (2023) where they employed 4-fold Cross Validation, 5-fold Cross Validation and 10-fold Cross Validation to different machine learning models and 10-fold Cross Validation showed the best increase in the performance. After the 10-Fold Cross Validation, XGBoost algorithm gave the best

result with prediction accuracy and F1 score 0. 8336 in the Fold 10 with the average prediction score of 0.8107 among the individual performances of the gradient boosting algorithms. This is contradiction with the study of Liu, Chen, Yang, Li, and Wang, (2023) where CatBoost has performed the best among LightGBM, XGBoost and other models. This can be because that CatBoost is more efficient in categorical feature processing and this feature presented prominent advantages in their study. Ensemble classification of XGBoost, LightGBM and CatBoost algorithms gave the best average result with respect to prediction accuracy and F1 score with 0. 8709 among all the algorithms and ensemble classifications, which is in parallel with the study of Malakouti (2023). This can be because both 10-Fold Cross Validation and ensemble classification improves the prediction performances of the machine learning models.

The minimum standard deviation in the 10-Fold Cross Validations of the gradient boosting models and ensemble classifications is from CatBoost model with 0.96% which means it has the lowest error values among all models which is accordance with the study of Lai, Demartino, and Xiao, (2023) and it can be because of the CatBoost's categorical handling feature reduces variability. It is also in accordance with Omer and Shareef's (2022) study where the authors compared CatBoost, AdaBoost, XGBoost, LightGBM and other models and CatBoost model showed the lowest mean and standard deviation. Standard deviation in the 10-Fold Cross Validation of the XGBoost model is higher than LightGBM model, which is contradiction with Lai et al.'s (2023) work and this can be because LightGBM uses utilized gradient-based optimization techniques to build decision trees.

The evaluation results suggest that gradient boosting algorithms, particularly LightGBM, are suitable for sentiment analysis tasks. Ensemble classifications using

multiple gradient boosting algorithms and soft voting techniques can improve the prediction accuracy and F1 score significantly. Applying 10-fold Cross Validation also can improve the prediction accuracy and F1 score while it gives an insight about the steadiness of the model. The study provides valuable insights for researchers and practitioners who are interested in the sentiment analysis area and selecting appropriate algorithms for their tasks.

Higher accuracy in machine learning models for sentiment classification can lead to several benefits for organizational performance like better decision making, improved customer satisfaction and increased operational efficiency. There are several studies suggests that increased accuracy of sentiment analysis can lead to increased efficiency and performance in organizations. The study by Verhoef, Kooge, and Walk (2016) investigated the impact of big data analytics, including sentiment analysis, on marketing decisions and its influence on organizational performance. The study found that accurate sentiment analysis can lead to improved customer service efficiency and effectiveness, which in turn increases customer satisfaction and loyalty. The study suggests that sentiment analysis can provide organizations with valuable insights into customer preferences and concerns, enabling them to make informed decisions that can improve their marketing campaigns and customer service processes. Additionally, the study found that organizations that use big data analytics, including sentiment analysis, to inform their marketing decisions are more likely to achieve superior performance compared to those that do not. Overall, the study supports that increased accuracy of sentiment analysis can lead to improved organizational performance, specifically in terms of customer service efficiency and effectiveness.

Higher accuracy in machine learning models holds substantial practical implications in the context of customer support and various business processes. With higher accuracy, ML models can effectively handle tasks such as transferring complaints to the appropriate customer support agent or department. By accurately classifying and routing customer complaints, businesses can streamline their support operations, reduce manual effort, and improve response times and this can lead to enhanced customer satisfaction as issues are addressed promptly and efficiently (Parviainen, Tihinen, Kääriäinen, & Teppola, 2022). Moreover, higher accuracy in ML models can enable businesses to gain valuable insights from customer interactions. Accurate sentiment analysis, for instance, helps in understanding customer sentiments, preferences, and pain points, enabling organizations to tailor their support strategies and improve overall customer experience (Müller, 2016). Additionally, accurate ML models can assist in automating routine customer inquiries, freeing up human support agents to focus on more complex and specialized tasks (Mehr, Ash, & Fellow, 2017). Overall, higher accuracy in ML models empowers businesses to optimize their customer support processes, enhance customer satisfaction, and drive operational efficiency.

The findings of the research indicate that machine learning algorithms, particularly sentiment classification with gradient boosting algorithms, can be effective in improving customer support services. The research recommends that businesses can use machine learning techniques to enhance their customer support by personalizing customer interactions, analyzing customer data for predicting customer needs. Additionally, the research suggests that employing ensemble classification with multiple gradient boosting algorithms can significantly improve the accuracy and F1 score of sentiment analysis tasks. Implementing machine learning algorithms to analyze

customer data can lead to better decision-making and improved organizational performance, resulting in increased customer satisfaction, loyalty, and profitability. Furthermore, this research provides valuable insights for researchers and practitioners who are interested in the sentiment classification area and selecting appropriate algorithms for their tasks. By leveraging machine learning algorithms, businesses can establish a customer-centric culture that prioritizes long-term relationships with customers and builds a stronger emotional connection with their customers, ultimately driving loyalty and growth in today's highly competitive market.

4.4 Limitations and future research

While the classification model with highest performance has good prediction score and F1 score (Table 9), it can perform better and improved performance results can be achieved with more balanced dataset. Expanding the dataset could decrease the imbalance in the dataset and it could also increase the performance results of the study. Also, algorithm choice for sentiment analysis with text data will be varying on the use and focus of the research and application like where will it be used: marketing, product development, customer support etc. Additionally, text data can be gathered from several sources like different companies, blogs, social medias.

Also, this researched can be improved by introducing the model with highest performance to the organizations and making qualitative research with the organizations. This could bring an understanding on impacts of higher accuracy of the sentiment analysis in organizational performance.

CHAPTER 5

CONCLUSION

In today's competitive market, establishing and maintaining long-term relationships with customers is crucial for businesses. Customer relationship management (CRM) is a vital aspect of building these relationships and creating a customer-centric culture. However, implementing CRM poses significant challenges for businesses, such as integrating data from different sources and ensuring customer engagement. To overcome these challenges, companies can employ various strategies, including utilizing social media, providing special discounts to loyal customers, and using machine learning algorithms to improve customer service and forecast user behavior.

Effective customer support is essential for any organization's success, as it helps guide customers regarding the goods or services provided. With the advancement of technology, customer support has evolved, and machine learning techniques can be used to enhance it. Machine learning algorithms can analyze customer inquiries, provide automated responses, and even identify common customer issues. By analyzing customer support data, such as chat logs, emails, and call center transcripts, machine learning algorithms can identify patterns and trends in customer issues, which can help organizations quickly resolve problems and build a stronger emotional connection with their customers.

This research contributes to the growing use of machine learning algorithms in customer support by providing practical insights for organizations seeking to improve their customer support services. The research focused on using XGBoost, CatBoost, LightGBM, and ensemble classification of these algorithms to analyze data coming from

a global service provider's Twitter account. The gradient boosting algorithms were chosen for their high prediction accuracy and performance on training. Ensemble classification with soft voting was included in the research to improve the model's predictivity and performance. The results showed that ensemble classification of XGBoost, LightGBM, and CatBoost was the best-performing model, with a prediction accuracy and F1 score of 0.8707. The individual performances showed that LightGBM outperformed XGBoost and CatBoost, with a prediction accuracy and F1 score of 0.8096. The evaluation results suggest that ensemble classifications using multiple gradient boosting algorithms and soft voting techniques can significantly improve prediction accuracy and F1 scores.

Higher accuracy in machine learning models for sentiment analysis can lead to several benefits for organizational performance, such as better decision-making and improved customer satisfaction. With higher accuracy, ML models can effectively handle tasks such as transferring complaints to the appropriate customer support agent or department. Additionally, accurate ML models can assist in automating routine customer inquiries, freeing up human support agents to focus on more complex and specialized tasks. By analyzing customer data, businesses can personalize the customer experience, forecast customer behavior, and improve customer support services. The implementation of machine learning algorithms in CRM and customer support can help businesses establish a customer-centric culture and build long-term relationships with their customers, resulting in increased customer satisfaction, loyalty, and profitability.

In conclusion, the use of machine learning algorithms in CRM and customer support is a valuable tool for businesses seeking to improve customer satisfaction and loyalty. This research provides insights into the effectiveness of gradient boosting

algorithms for sentiment analysis tasks and highlights the benefits of using ensemble classifications with soft voting techniques. By leveraging these technologies, businesses can establish a customer-centric culture and build lasting relationships with their customers, leading to increased profitability and success in today's competitive market.

APPENDIX

CODE

```
# Load the data from the Excel file
data = pd.read_excel('/Users/ibm/BOUN/datacopy.xlsx')
# Define the target column and convert it to numerical labels
target_map = {"+": 1, "-": 0, "n": 2}
data["label"] = data["label"].map(target_map)
# Define the feature and target data
features = data["ileti"]
target = data["label"]
# Split the data into training and testing sets
train_features, test_features, train_target, test_target =
train_test_split(features, target, test_size=0.2, random_state=42
# Transform documents to document-term matrix
vectorizer = TfidfVectorizer()
train_features = vectorizer.fit_transform(train_features)
test_features = vectorizer.transform(test_features)


# Confusion matrix for multiclass
def conf_matrix(predicted_y):
    cm = confusion_matrix(test_target, predicted_y)
    cm_df = pd.DataFrame(cm,
                         index=['negative', 'positive', 'neutral'],
                         columns=['negative', 'positive', 'neutral'])
    plt.figure(figsize=(5, 4))
    sns.heatmap(cm_df, annot=True)
```

```
    plt.title('Confusion Matrix')
    plt.ylabel('Actual Values')
    plt.xlabel('Predicted Values')
    plt.show()


# XGBoost Model
def xgb_boost():
    global train_features, test_features
    # Define the XGBoost model and train it on the training data
    dtrain = xgb.DMatrix(train_features, label=train_target)
    params = {"objective": "multi:softmax", "num_class": 3,
"num_leaves": 3, "learning_rate": 0.05}
    model = xgb.train(params, dtrain)
    # Predict the labels for the test data and calculate the accuracy
    dtest = xgb.DMatrix(test_features)
    preds = model.predict(dtest)
    # Calculate accuracy score and f1 score
    accuracy = accuracy_score(test_target, preds)
    print("Accuracy: %.2f%%" % (accuracy))
    f1 = sklearn.metrics.f1_score(test_target, preds, average='micro')
    print("f1:", f1)
    # decision tree
    plot_tree(model, num_trees=3, rankdir='LR')
    plt.gcf().set_size_inches(18.5, 10.5)
    plt.show()
    # Confusion matrix
    conf_matrix(preds)
```

```
# XGBoost Model with cross-validation
def xgb_boost_k_fold():
    global features, target
    params = {"objective": "multi:softmax", "num_class": 3,
"num_leaves": 3, "learning_rate": 0.05}
    # Perform 10-fold cross-validation
    cv = StratifiedKFold(n_splits=10, shuffle=True, random_state=42)
    a_scores = []
    f_scores = []
    for train_idx, test_idx in cv.split(features, target):
        train_features, test_features = features[train_idx],
features[test_idx]
        train_target, test_target = target[train_idx], target[test_idx]
        # Transform documents to document-term matrix
        vectorizer = TfidfVectorizer()
        train_features = vectorizer.fit_transform(train_features)
        test_features = vectorizer.transform(test_features)
        dtrain = xgb.DMatrix(train_features, label=train_target)
        model = xgb.train(params, dtrain)
        # Predict the labels for the test data and calculate the
accuracy
        dtest = xgb.DMatrix(test_features)
        preds = model.predict(dtest)
        # Calculate accuracy score and f1 score
        accuracy = accuracy_score(test_target, preds)
        print("Accuracy: %.2f%%" % (accuracy))
        a_scores.append(accuracy)
        f1 = sklearn.metrics.f1_score(test_target, preds,
average='micro')
```

```
        print("f1:", f1)
        f_scores.append(f1)
    # Print the cross-validation scores
    a_scores = np.array(a_scores)
    print("Accuracy Cross-validation scores:", a_scores)
    print("Average accuracy: %.2f%%" % (a_scores.mean() * 100))
    print("Standard deviation: %.2f%%" % (a_scores.std() * 100))
    # Print the cross-validation scores
    f_scores = np.array(f_scores)
    print("f1 Cross-validation scores:", f_scores)
    print("Average accuracy: %.2f%%" % (f_scores.mean() * 100))
    print("Standard deviation: %.2f%%" % (f_scores.std() * 100))

# LightGBM Model
def lightgbm():
    global train_features, test_features

    # Define the training dataset
    train_data = lgb.Dataset(train_features, label=train_target)
    # Set the hyperparameters for the LightGBM model
    params = {
        "objective": "multiclass",
        "metric": "multi_logloss",
        "num_classes": 3,
        "num_leaves": 3,
        "learning_rate": 0.05

    # Train the LightGBM model
    model = lgb.train(params, train_data, num_boost_round=100)
```

```
    # Use the model to make predictions on the testing data
    y_pred = model.predict(test_features)
    # Convert predictions from one-hot encoding to class indices
    y_pred = np.argmax(y_pred, axis=1)
    # Calculate accuracy score and f1 score
    accuracy = accuracy_score(test_target, y_pred)
    f1 = sklearn.metrics.f1_score(test_target, y_pred, average='micro')
    print("Accuracy:", accuracy)
    print("f1:", f1)
    # decision tree
    lgbm_tree(model, tree_index=4)
    plt.gcf().set_size_inches(18.5, 10.5)
    plt.show()
    # Confusion matrix
    conf_matrix(y_pred)

# LightGBM Model with cross-validation
def lightgbm_k_fold():
    global features, target
    # Set the hyperparameters for the LightGBM model
    params = {
        "objective": "multiclass",
        "metric": "multi_logloss",
        "num_classes": 3,
        "num_leaves": 3,
        "learning_rate": 0.05
    }
    # Perform 10-fold cross-validation
    cv = StratifiedKFold(n_splits=10, shuffle=True, random_state=42)
```

```
    a_scores = []
    f_scores = []
    for train_idx, test_idx in cv.split(features, target):
        train_features, test_features = features[train_idx],
features[test_idx]
        train_target, test_target = target[train_idx], target[test_idx]
        # Transform documents to document-term matrix
        vectorizer = TfidfVectorizer()
        train_features = vectorizer.fit_transform(train_features)
        test_features = vectorizer.transform(test_features)
        # Define the training dataset
        train_data = lgb.Dataset(train_features, label=train_target)
        # Train the LightGBM model
        model = lgb.train(params, train_data, num_boost_round=100)
        # Use the model to make predictions on the testing data
        preds = model.predict(test_features)
        # Convert predictions from one-hot encoding to class indices
        preds = np.argmax(preds, axis=1)
        # Calculate accuracy score and f1 score
        accuracy = accuracy_score(test_target, preds)
        print("Accuracy: %.2f%%" % (accuracy))
        a_scores.append(accuracy)
        f1 = sklearn.metrics.f1_score(test_target, preds,
average='micro')
        print("f1:", f1)
        f_scores.append(f1)
    # Print the cross-validation scores
    a_scores = np.array(a_scores)
    print("Accuracy Cross-validation scores:", a_scores)
```

```
    print("Average accuracy: %.2f%%" % (a_scores.mean() * 100))
    print("Standard deviation: %.2f%%" % (a_scores.std() * 100))
    # Print the cross-validation scores
    f_scores = np.array(f_scores)
    print("f1 Cross-validation scores:", f_scores)
    print("Average accuracy: %.2f%%" % (f_scores.mean() * 100))
    print("Standard deviation: %.2f%%" % (f_scores.std() * 100))


# CatBoost Model
def catboost():
    global train_features, test_features
    # Define the CatBoost model and train it on the training data
    pool = ctb.Pool(train_features, train_target)
    cf = ctb.CatBoostClassifier(iterations=10,
learning_rate=0.05,depth=3,loss_function='MultiClass', verbose=False)
    model_CBC = cf.fit(pool)
    # decision tree
    decision_tree = model_CBC.plot_tree(tree_idx=3)
    decision_tree.render()
    # Use the model to make predictions on the testing data
    expected_y = test_target
    predicted_y = model_CBC.predict(test_features)
    # Calculate accuracy score and f1 score
    accuracy = accuracy_score(expected_y, predicted_y)
    print("Accuracy: %.2f%%" % (accuracy * 100.0))
    f1 = sklearn.metrics.f1_score(test_target, predicted_y,
average='micro')
    print("f1:", f1)
    # Confusion matrix
```

```
    conf_matrix(predicted_y)


# CatBoost Model with cross-validation
def catboost_k_fold():
    global features, target
    # Perform 10-fold cross-validation
    cv = StratifiedKFold(n_splits=10, shuffle=True, random_state=42)
    a_scores = []
    f_scores = []
    for train_idx, test_idx in cv.split(features, target):
        train_features, test_features = features[train_idx],
features[test_idx]
        train_target, test_target = target[train_idx], target[test_idx]
        # Transform documents to document-term matrix
        vectorizer = TfidfVectorizer()
        train_features = vectorizer.fit_transform(train_features)
        test_features = vectorizer.transform(test_features)
        # Define the CatBoost model and train it on the training data
        pool = ctb.Pool(train_features, train_target)
        cf = ctb.CatBoostClassifier(iterations=10, learning_rate=0.05,
depth=3, loss_function='MultiClass',
                                    verbose=False)
        model_CBC = cf.fit(pool)
        # Use the model to make predictions on the testing data
        preds = model_CBC.predict(test_features)
        # Calculate accuracy score and f1 score
        accuracy = accuracy_score(test_target, preds)
        print("Accuracy: %.2f%%" % (accuracy))
        a_scores.append(accuracy)
```

```
        f1 = sklearn.metrics.f1_score(test_target, preds,
average='micro')
        print("f1:", f1)
        f_scores.append(f1)
    # Print the cross-validation scores
    a_scores = np.array(a_scores)
    print("Accuracy Cross-validation scores:", a_scores)
    print("Average accuracy: %.2f%%" % (a_scores.mean() * 100))
    print("Standard deviation: %.2f%%" % (a_scores.std() * 100))
    # Print the cross-validation scores
    f_scores = np.array(f_scores)
    print("f1 Cross-validation scores:", f_scores)
    print("Average accuracy: %.2f%%" % (f_scores.mean() * 100))
    print("Standard deviation: %.2f%%" % (f_scores.std() * 100))

# Ensemble Classification Model
def ensemble():
    global train_features, test_features
    # CatBoost Model
    gb1 = ctb.CatBoostClassifier(verbose=False)
    # LightGBM Model
    gb2 = lgb.LGBMClassifier()
    # XGBoost Model
    gb3 = xgb.XGBClassifier()
    # Soft voting classification
    egb = VotingClassifier(estimators=[('cb', gb1), ('lb', gb2), ('xb',
gb3)], voting='soft')
    # Use the selected model to make predictions on the testing data
    egb.fit(train_features, train_target)
```

```
    y_pred = egb.predict(test_features)
    # Calculate accuracy score and f1 score
    score = accuracy_score(test_target, y_pred)
    print("Accuracy: %.2f%%" % (score * 100.0))
    f1 = sklearn.metrics.f1_score(test_target, y_pred, average='micro')
    print("f1:", f1)
    # Confusion matrix
    conf_matrix(y_pred)

# Ensemble Model with cross-validation
def ensemble_k_fold():
    global features, target
    # CatBoost Model
    gb1 = ctb.CatBoostClassifier(verbose=False)
    # LightGBM Model
    gb2 = lgb.LGBMClassifier()
    # XGBoost Model
    gb3 = xgb.XGBClassifier()
    # Soft voting classification
    egb = VotingClassifier(estimators=[('xb', gb3), ('lb', gb2)],
voting='soft')
    # Perform 10-fold cross-validation
    cv = StratifiedKFold(n_splits=10, shuffle=True, random_state=42)
    a_scores = []
    f_scores = []
    for train_idx, test_idx in cv.split(features, target):
        train_features, test_features = features[train_idx],
features[test_idx]
        train_target, test_target = target[train_idx], target[test_idx]
```

```
        # Transform documents to document-term matrix
        vectorizer = TfidfVectorizer()
        train_features = vectorizer.fit_transform(train_features)
        test_features = vectorizer.transform(test_features)
        # Use the selected model to make predictions on the testing
data
        egb.fit(train_features, train_target)
        preds = egb.predict(test_features)
        # Calculate accuracy score and f1 score
        accuracy = accuracy_score(test_target, preds)
        print("Accuracy: %.2f%%" % (accuracy))
        a_scores.append(accuracy)
        f1 = sklearn.metrics.f1_score(test_target, preds,
average='micro')
        print("f1:", f1)
        f_scores.append(f1)
    # Print the cross-validation scores
    a_scores = np.array(a_scores)
    print("Accuracy Cross-validation scores:", a_scores)
    print("Average accuracy: %.2f%%" % (a_scores.mean() * 100))
    print("Standard deviation: %.2f%%" % (a_scores.std() * 100))
    # Print the cross-validation scores
    f_scores = np.array(f_scores)
    print("f1 Cross-validation scores:", f_scores)
    print("Average accuracy: %.2f%%" % (f_scores.mean() * 100))
    print("Standard deviation: %.2f%%" % (f_scores.std() * 100))
```

```
# SVM Model
def svm():
    global train_features, test_features
    clf = SVC(kernel='poly', degree=3, gamma=0.05)
    clf.fit(train_features, train_target)
    prediction = clf.predict(test_features)
    # Calculate accuracy score and f1 score
    accuracy = accuracy_score(test_target, prediction)
    print("Accuracy:", accuracy)
    f1 = sklearn.metrics.f1_score(test_target, prediction,
average='micro')
    print("f1:", f1)
    # Confusion matrix
    conf_matrix(prediction)
# RF Model
def rf():
    global train_features, test_features
    classifier = RandomForestClassifier(max_depth = 3, criterion =
'entropy', random_state = 42)
    classifier.fit(train_features, train_target)
    prediction = classifier.predict(test_features)
    # Calculate accuracy score and f1 score
    accuracy = accuracy_score(test_target, prediction)
    print("Accuracy:", accuracy)
    f1 = sklearn.metrics.f1_score(test_target, prediction,
average='micro')
    print("f1:", f1)
    # Confusion matrix
    conf_matrix(prediction)
```

REFERENCES

A, S., & Myrland, Ø. (2022). Statistical learning to estimate energy savings from retrofitting in the Norwegian food retail market. *Renewable & Sustainable Energy Reviews, 167,* 112691. https://doi.org/10.1016/j.rser.2022.112691

Abbasniya, M. R., Sheikholeslamzadeh, S. A., Nasiri, H. R., & Emami, S. (2022). Classification of breast tumors based on histopathology images using deep features and ensemble of gradient boosting methods. *Computers & Electrical Engineering, 103*, 108382. https://doi.org/10.1016/j.compeleceng.2022.108382

Acheampong, F. A., Wenyu, C., & Nunoo-Mensah, H. (2020). Text-based emotion detection: Advances, challenges, and opportunities. *Engineering Reports,* 2(7). https://doi.org/10.1002/eng2.12189

Akter, S., Dwivedi, Y. K., Sajib, S., Biswas, K., Bandara, R., & Michael, K. (2022). Algorithmic bias in machine learning-based marketing models. *Journal of Business Research, 144*, 201–216. https://doi.org/10.1016/j.jbusres.2022.01.083

Alakus, T. B., & Turkoglu, I. (2020). Comparison of deep learning approaches to predict COVID-19 infection. *Chaos Solitons & Fractals, 140*, 110120. https://doi.org/10.1016/j.chaos.2020.110120

Amrani, Y. A., Lazaar, M., & Kadiri, K. E. E. (2018). Random forest and support vector machine based hybrid approach to sentiment analysis. *Procedia Computer Science, 127*, 511–520. https://doi.org/10.1016/j.procs.2018.01.150

Anguita, D. Ghelardoni, L. Ghio, A. Oneto, L & Ridella, S. The 'K' in k-fold cross validation in European Symposium on artificial neural networks. *Computational Intelligence and Machine Learning,* 441–446 (2012).

Anshari, M., Almunawar, M. N., Lim, S. A., & Al-Mudimigh, A. S. (2019). Customer relationship management and big data enabled: Personalization & customization of services. *Applied Computing and Informatics, 15*(2), 94–101. https://doi.org/10.1016/j.aci.2018.05.004

Barnes, N. G., & Mattson, E. (2009). Social media in the 2009 Inc. 500: New tools and new trends. *Journal of New Communication Research, 4*(2), 70-79.

Bell, B., & Shankar, R. (2019). Big data analytics: Opportunities and challenges. *Handbook of Research on Big Data and the IoT* (pp. 1-19).

Bian, Y., Ye, R., Zhang, J., & Xintian, Y. (2022). Customer preference identification from hotel online reviews: A neural network based fine-grained sentiment analysis. *Computers & Industrial Engineering, 172,* 108648. https://doi.org/10.1016/j.cie.2022.108648

Bramer, M. (2016). Ensemble classification. *Principles of Data Mining* (pp. 209–220). https://doi.org/10.1007/978-1-4471-7307-6_14

Breiman, L. (2001). Random forests. *Machine learning, 45*, 5-32.

Bui, X., Nguyen, H., Choi, Y., Nguyen-Thoi, T., Zhou, J., & Pradhan, B. (2020). Prediction of slope failure in open-pit mines using a novel hybrid artificial intelligence model based on decision tree and evolution algorithm. *Scientific Reports, 10*(1). https://doi.org/10.1038/s41598-020-66904-y

Buda, M., Maki, A., & Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks, 106*, 249–259. https://doi.org/10.1016/j.neunet.2018.07.011

Budac, C. (2016). Theoretical approaches on successful email marketing campaigns. *Ovidius University Annals: Economic Sciences Series, 16(*2), 306-311.

Buyya, R., Calheiros, R. N., & Dastjerdi, A. V. (Eds.). (2016). *Big data: principles and paradigms.* Morgan Kaufmann.

Cambria, E., & White, B. (2014). Jumping NLP curves: A review of natural language processing research. *IEEE Computational Intelligence Magazine, 9*(2), 48–57. https://doi.org/10.1109/mci.2014.2307227

Cao, Q., Duan, W., & Gan, Q. (2011). Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach. *Decision Support Systems, 50*(2), 511–521. https://doi.org/10.1016/j.dss.2010.11.009

Chakraborty, K., Bhatia, S., Bhattacharyya, S., Platos, J., Bag, R., & Hassanien, A. E. (2020). Sentiment analysis of COVID-19 tweets by deep learning classifiers—A study to show how popularity is affecting accuracy in social media. *Applied Soft Computing, 97*, 106754. https://doi.org/10.1016/j.asoc.2020.106754

Chang, J. W., Yen, N. Y., & Hung, J. C. (2022). Design of a NLP-empowered finance fraud awareness model: the anti-fraud chatbot for fraud detection and fraud classification as an instance. *Journal of Ambient Intelligence and Humanized Computing, 13*(10), 4663–4679. https://doi.org/10.1007/s12652-021-03512-2

Chaudhary, A., & Tomar, P. (2019). Big data and IoT applications in real life environment. In *Advances in data mining and database management book series*. IGI Global. https://doi.org/10.4018/978-1-5225-7432-3.ch001

Chen, D., & Zhao, H. (2012). *Data Security and Privacy Protection Issues in Cloud Computing*. https://doi.org/10.1109/iccsee.2012.193

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *Management Information Systems Quarterly, 36*(4), 1165. https://doi.org/10.2307/41703503

Chen, I. J., & Popovich, K. (2003). Understanding customer relationship management (CRM). *Business Process Management Journal, 9*(5), 672–688. https://doi.org/10.1108/14637150310496758

Chen, S., Wang, L., Zhang, H., Wang, J., & Peng, J. (2021). Customer purchase forecasting for online tourism: A data-driven method with multiplex behavior data. *Tourism Management, 87*, 104357. https://doi.org/10.1016/j.tourman.2021.104357

Chen, T., & Guestrin, C. (2016). *XGBoost*. https://doi.org/10.1145/2939672.2939785

Chen, Y., Fay, S., & Wang, Q. (2011). The role of marketing in social media: How online consumer reviews evolve. *Journal of Interactive Marketing, 25*(2), 85–94. https://doi.org/10.1016/j.intmar.2011.01.003

Clarke, D. (2018). *Experience is everything: Here's how to get it right.* Retrieved May 7, 2023, from https://www.pwc.de/de/consulting/pwc-consumer-intelligence-series-customer-experience.pdf

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*(3), 273–297. https://doi.org/10.1007/bf00994018

Cortez, R. M., Clarke, A. E., & Freytag, P. V. (2021). B2B market segmentation: A systematic review and research agenda. *Journal of Business Research, 126*, 415–428. https://doi.org/10.1016/j.jbusres.2020.12.070

Cuadros, A. J., & Domínguez, V. (2014). Customer segmentation model based on value generation for marketing strategies formulation. *Estudios Gerenciales, 30*(130), 25–30. https://doi.org/10.1016/j.estger.2014.02.005

CRM Strategies and Technologies to Understand, Grow and Manage Customer Experiences. (2011). Retrieved May 7, 2023, from https://www.gartner.com/imagesrv/summits/docs/na/customer-360/C360_2011_brochure_FINAL.pdf

Dahiya, M. (2017). A tool of conversation: Chatbot. *International Journal of Computer Sciences and Engineering, 5*(5), 158-161.

Davenport, T. (2014). Big data at work: Dispelling the myths, uncovering the opportunities. *Harvard Business Review Press.*

Davenport, T. H., & Harris, J. G. (2007). Competing on analytics: The new science of winning. *Harvard business review press, Language, 15*(217), 24.

Demner-Fushman, D., Chapman, W. W., & McDonald, C. J. (2009). What can natural language processing do for clinical decision support? *Journal of Biomedical Informatics, 42*(5), 760–772. https://doi.org/10.1016/j.jbi.2009.08.007

Dietterich, T. G. (2000). Ensemble methods in machine learning. In *Lecture Notes in Computer Science* (pp. 1–15). Springer Science+Business Media. https://doi.org/10.1007/3-540-45014-9_1

Dorogush, A. V., Ershov, V., & Gulin, A. (2018). CatBoost: Gradient boosting with categorical features support. *arXiv preprint arXiv:1810.11363.* https://doi.org/10.48550/arXiv.1810.11363

Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996, August). A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd* (Vol. 96, No. 34, pp. 226-231).

Feng, Y., Yin, Y., Wang, D., & Dhamotharan, L. (2022). A dynamic ensemble selection method for bank telemarketing sales prediction. *Journal of Business Research, 139,* 368–382. https://doi.org/10.1016/j.jbusres.2021.09.067

Fine, S., & Scheinberg, K. (2001). Efficient SVM training using low-rank kernel representations. *Journal of Machine Learning Research, 2*(Dec), 243-264.

Galar, M., Fernández, A. Á., Barrenechea, E., Bustince, H., & Herrera, F. (2012). A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches. *IEEE Transactions on Systems, Man and Cybernetics, 42*(4), 463–484. https://doi.org/10.1109/tsmcc.2011.2161285

Gandomi, A. H., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*(2), 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

Gnewuch, U., Morana, S., Adam, M., & Maedche, A. (2018). Faster is not always better: Understanding the effect of dynamic response delays in human-chatbot interaction.

Grönroos, C. (2007). *Service management and marketing: customer management in service competition*. John Wiley & Sons.

Gulfraz, M., Sufyan, M., Mustak, M., Salminen, J., & Srivastava, D. (2022). Understanding the impact of online customers' shopping experience on online impulsive buying: A study on two leading E-commerce platforms. *Journal of Retailing and Consumer Services, 68,* 103000. https://doi.org/10.1016/j.jretconser.2022.103000

Gummesson, E. (2017). *Case theory in business and management: Reinventing case study research.* Sage.

Gupta, R., & Kabadayi, S. (2010). The relationship between trusting beliefs and web site loyalty: The moderating role of consumer motives and flow. *Psychology & Marketing, 27*(2), 166–185. https://doi.org/10.1002/mar.20325

Guyon, I., Saffari, A., Dror, G., & Cawley, G. (2010). Model selection: Beyond the bayesian/frequentist divide. *Journal of Machine Learning Research, 11*(1).

Han, H., Liang, Y., Bella, G., Giunchiglia, F., & Li, D. (2023). LFDNN: A novel hybrid recommendation model based on DeepFM and LightGBM. *Entropy, 25*(4), 638. https://doi.org/10.3390/e25040638

Hastie, T., Tibshirani, R., & Friedman, J. H. (2008). Unsupervised learning. In *Springer series in statistics* (pp. 485–585). Springer Science+Business Media. https://doi.org/10.1007/978-0-387-84858-7_14

Hennig-Thurau, T., Gwinner, K. P., & Gremler, D. D. (2002). Understanding relationship marketing outcomes. *Journal of Service Research, 4*(3), 230–247. https://doi.org/10.1177/1094670502004003006

Homburg, C., Jozić, D., & Kuehnl, C. (2017). Customer experience management: Toward implementing an evolving marketing concept. *Journal of the Academy of Marketing Science, 45*(3), 377–401. https://doi.org/10.1007/s11747-015-0460-7

Hu, M., Stephen, B., Browell, J., Haben, S., & Wallom, D. (2023). Impacts of building load dispersion level on its load forecasting accuracy: Data or algorithms? Importance of reliability and interpretability in machine learning. *Energy and Buildings, 285,* 112896. https://doi.org/10.1016/j.enbuild.2023.112896

Hu, X., Liu, A., Li, X., Dai, Y., & Nakao, M. (2023). Explainable AI for customer segmentation in product development. *CIRP Annals.* https://doi.org/10.1016/j.cirp.2023.03.004

Imran, A. A., & Amin, N. (2020). Predicting the return of orders in the E-tail industry accompanying with model interpretation. *Procedia Computer Science, 176*, 1170–1179. https://doi.org/10.1016/j.procs.2020.09.113

Jabeur, S. B., Gharib, C., Mefteh-Wali, S., & Arfi, W. B. (2021). CatBoost model and artificial intelligence techniques for corporate failure prediction. *Technological Forecasting and Social Change, 166*, 120658. https://doi.org/10.1016/j.techfore.2021.120658

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: with applications in R*, ser. Springer texts in statistics.

Jane, J. B., & Ganesh, E. N. (2019). A review on big data with machine learning and fuzzy logic for better decision making. *International Journal of Scientific & Technology Research, 8*(10), 1121-1125.

Japkowicz, N., & Shah, M. (2011). *Evaluating learning algorithms: a classification perspective.* Cambridge University Press.

Jordan, M. I., & Mitchell, T. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*(6245), 255–260. https://doi.org/10.1126/science.aaa8415

Joung, J., & Kim, H. M. (2023). Interpretable machine learning-based approach for customer segmentation for new product development from online product reviews. *International Journal of Information Management, 70*, 102641. https://doi.org/10.1016/j.ijinfomgt.2023.102641

Kaur, H., & Kumar, R. (2015). *Sentiment analysis from social media in crisis situations.* https://doi.org/10.1109/ccaa.2015.7148383

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T. Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems, 30.*

Kim, A. J., & Ko, E. (2010). Impacts of luxury fashion brand's social media marketing on customer relationship and purchase intention. *Journal of Global Fashion Marketing, 1*(3), 164–171. https://doi.org/10.1080/20932685.2010.10593068

Kim, H., Kim, H., & Chang, J. (2017). A secure kNN query processing algorithm using homomorphic encryption on outsourced database. *Data and Knowledge Engineering, 123,* 101602. https://doi.org/10.1016/j.datak.2017.07.005

Kiran, R., Kumar, P., & Bhasker, B. (2020). Oslcfit (organic simultaneous LSTM and CNN Fit): A novel deep learning based solution for sentiment polarity classification of reviews. *Expert Systems With Applications, 157,* 113488. https://doi.org/10.1016/j.eswa.2020.113488

Klibi, W., Babai, M. Z., Ducq, Y., & Akher, H. O. a. E. (2021). Basket data-driven approach for omnichannel demand forecasting. *International Journal of Production Economics, 257,* 108748. https://doi.org/10.1016/j.ijpe.2022.108748

Kohavi, R. (1995, August). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai* (Vol. 14, No. 2, pp. 1137-1145).

Kotler, P. (2001). *Marketing management, millenium edition*. Prentice-Hall, Inc..

Kotsiantis, S., Kanellopoulos, D., & Pintelas, P. (2006). Handling imbalanced datasets: A review. *GESTS international transactions on computer science and engineering, 30*(1), 25-36.

Kuncheva, L. I. (2014). *Combining pattern classifiers: methods and algorithms.* John Wiley & Sons.

Kübler, S., Liu, C., & Sayyed, Z. A. (2018). To use or not to use: Feature selection for sentiment analysis of highly imbalanced data. *Natural Language Engineering, 24*(1), 3–37. https://doi.org/10.1017/s1351324917000298

Lai, D., Demartino, C., & Xiao, Y. (2023). Interpretable machine-learning models for maximum displacements of RC beams under impact loading predictions. *Engineering Structures, 281,* 115723. https://doi.org/10.1016/j.engstruct.2023.115723

Lee, S. B. (2020). Chatbots and communication: The growing role of artificial intelligence in addressing and shaping customer needs. *Business Communication Research and Practice, 3*(2), 103–111. https://doi.org/10.22682/bcrp.2020.3.2.103

Lian, Y., Gao, J., & Ye, T. (2022). How does green credit affect the financial performance of commercial banks? ——Evidence from China. *Journal of Cleaner Production, 344*, 131069. https://doi.org/10.1016/j.jclepro.2022.131069

Linden, G., Smith, B. R., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing, 7*(1), 76–80. https://doi.org/10.1109/mic.2003.1167344

Liu, B., & Zhang, L. (2012). A Survey of Opinion Mining and Sentiment Analysis. In *Springer eBooks* (pp. 415–463). https://doi.org/10.1007/978-1-4614-3223-4_13

Liu, Y., Chen, X., Yang, J., Li, L., & Wang, T. (2023). Snow avalanche susceptibility mapping from tree-based machine learning approaches in ungauged or poorly-gauged regions. *Catena, 224*, 106997. https://doi.org/10.1016/j.catena.2023.106997

Liu, Y., Hu, B., Yan, W., & Lin, Z. (2023). Can chatbots satisfy me? A mixed-method comparative study of satisfaction with task-oriented chatbots in mainland China and Hong Kong. *Computers in Human Behavior, 143*, 107716. https://doi.org/10.1016/j.chb.2023.107716

Loebbecke, C., & Picot, A. (2015). Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *Journal of Strategic Information Systems, 24*(3), 149–157. https://doi.org/10.1016/j.jsis.2015.08.002

Ma, S., & Fildes, R. (2021). Retail sales forecasting with meta-learning. *European Journal of Operational Research, 288*(1), 111–128. https://doi.org/10.1016/j.ejor.2020.05.038

Malakouti, S. M. (2023). Babysitting hyperparameter optimization and 10-fold-cross-validation to enhance the performance of ML methods in predicting wind speed and energy generation. *Intelligent Systems with Applications*, 200248. https://doi.org/10.1016/j.iswa.2023.200248

Marino, V., & Lo Presti, L. (2018). Engagement, satisfaction and customer behavior-based CRM performance: An empirical study of mobile instant messaging. *Journal of Service Theory and Practice, 28*(5), 682-707.

Marr, B. (2015). *Big Data: Using SMART big data, analytics and metrics to make better decisions and improve performance.* John Wiley & Sons.

McClelland, C. (2020). The impact of artificial intelligence–widespread job losses. *IT for All, 1*.

McLean, G., Osei-Frimpong, K., Wilson, A. A., & Pitardi, V. (2020). How live chat assistants drive travel consumers' attitudes, trust and purchase intentions. *International Journal of Contemporary Hospitality Management, 32*(5), 1795–1812. https://doi.org/10.1108/ijchm-07-2019-0605

Mehr, H., Ash, H., & Fellow, D. (2017). Artificial intelligence for citizen services and government. *Ash Cent. Democr. Gov. Innov. Harvard Kennedy Sch., no. August,* 1-12.

Mishev, K., Gjorgjevikj, A., Vodenska, I., Chitkushev, L. T., & Trajanov, D. (2020). Evaluation of sentiment analysis in finance: From lexicons to transformers. *IEEE Access, 8,* 131662–131682. https://doi.org/10.1109/access.2020.3009626

Misischia, C. V., Poecze, F., & Strauss, C. (2022). Chatbots in customer service: Their relevance and impact on service quality. *Procedia Computer Science, 201*, 421–428. https://doi.org/10.1016/j.procs.2022.03.055

Müller, O. (2016, December 31). *Using Text Analytics to Derive Customer Service Management Benefits from Unstructured Data.* Retrieved from https://pure.itu.dk/en/publications/using-text-analytics-to-derive-customer-service-management-benefi

Narayanan, A., & Shmatikov, V. (2010). Myths and fallacies of "Personally Identifiable Information." *Communications of the ACM, 53*(6), 24–26. https://doi.org/10.1145/1743546.1743558

Ngai, E. W., Xiu, L., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems With Applications, 36*(2), 2592–2602. https://doi.org/10.1016/j.eswa.2008.02.021

Odusami, M., Abayomi-Alli, O., Misra, S., Abayomi-Alli, A., & Sharma, M. M. (2021). A hybrid machine learning model for predicting customer churn in the telecommunication industry. In *Innovations in Bio-Inspired Computing and Applications: Proceedings of the 11th International Conference on Innovations in Bio-Inspired Computing and Applications (IBICA 2020) held during December 16-18, 2020 11* (pp. 458-468). Springer International Publishing.

Omar, H., Klibi, W., Babai, M. Z., & Ducq, Y. (2023). Basket data-driven approach for omnichannel demand forecasting. *International Journal of Production Economics, 257*, 108748. https://doi.org/10.1016/j.ijpe.2022.108748

Omer, Z. M., & Shareef, H. (2022). Comparison of decision tree based ensemble methods for prediction of photovoltaic maximum current. *Energy Conversion and Management: X, 16*, 100333. https://doi.org/10.1016/j.ecmx.2022.100333

Osman, A. F. (2019). Radiation oncology in the era of big data and machine learning for precision medicine. In *Artificial Intelligence-Applications in Medicine and Biology*. IntechOpen.

Pak, A., & Paroubek, P. (2010, May). Twitter as a corpus for sentiment analysis and opinion mining. In *LREc* (Vol. 10, No. 2010, pp. 1320-1326).

Paleyes, A., Urma, R., & Lawrence, N. D. (2022). Challenges in deploying machine learning: A survey of case studies. *ACM Computing Surveys, 55*(6), 1–29. https://doi.org/10.1145/3533378

Pang, B., & Lee, L. (2009). Opinion mining and sentiment analysis. *Comput. Linguist, 35*(2), 311-312.

Parasuraman, A., & Zinkhan, G. M. (2002). Marketing to and serving customers through the Internet: An overview and research agenda. *Journal of the academy of marketing science, 30*(4), 286-295. https://doi.org/10.1177/009207002236906

Parviainen, P., Tihinen, M., Kääriäinen, J., & Teppola, S. (2022). Tackling the digitalization challenge: How to benefit from digitalization in practice. *International Journal of Information Systems and Project Management, 5*(1), 63–77. https://doi.org/10.12821/ijispm050104

Pawełoszek, I. (2021). Customer segmentation based on activity monitoring applications for the recommendation system. *Procedia Computer Science, 192,* 4751–4761. https://doi.org/10.1016/j.procs.2021.09.253

Polikar, R. (2012). Ensemble learning. In *Springer eBooks* (pp. 1–34). https://doi.org/10.1007/978-1-4419-9326-7_1

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. *Advances in neural information processing systems, 31*.

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data, 1*(1), 51–59. https://doi.org/10.1089/big.2013.1508

Ramaswamy, S., & DeClerck, N. (2018). Customer perception analysis using deep learning and NLP. *Procedia Computer Science, 140,* 170–178. https://doi.org/10.1016/j.procs.2018.10.326

Rawat, S., Rawat, A. S., Kumar, D., & Sabitha, A. S. (2021). Application of machine learning and data visualization techniques for decision support in the insurance sector. *International Journal of Information Management Data Insights, 1*(2), 100012. https://doi.org/10.1016/j.jjimei.2021.100012

Reinartz, W., Krafft, M., & Hoyer, W. D. (2004). The customer relationship management process: Its measurement and impact on performance. *Journal of Marketing Research, 41*(3), 293–305. https://doi.org/10.1509/jmkr.41.3.293.35991

Ribelles, N., Jerez, J. M., Rodriguez-Brazzarola, P., Jiménez, B., Diaz-Redondo, T., Mesa, H., Márquez, A., Sánchez-Muñoz, A., Pajares, B., Carabantes, F., Bermejo, M. a. V., Villar, E., Dominguez-Recio, M. E., Saez, E., Gálvez, L., Godoy, A. L. P. C., Franco, L., Ruiz-Medina, S., Lopez, I., & Alba, E. (2021). Machine learning and natural language processing (NLP) approach to predict early progression to first-line treatment in real-world hormone receptor-positive (HR+)/HER2-negative advanced breast cancer patients. *European Journal of Cancer, 144,* 224–231. https://doi.org/10.1016/j.ejca.2020.11.030

RightNow Technologies (Ed.). (2010). *Customer Experience Impact Report.* Retrieved May 7, 2023, from https://www.slideshare.net/RightNow/2010-customer-experience-impact

Rokach, L. (2010). Ensemble-based classifiers. *Artificial intelligence review, 33*, 1-39.

Rtayli, N., & Enneya, N. (2020). Enhanced credit card fraud detection based on SVM-recursive feature elimination and hyper-parameters optimization. *Journal of Information Security and Applications, 55*, 102596. https://doi.org/10.1016/j.jisa.2020.102596

Rust, R. T., Lemon, K. N., & Zeithaml, V. A. (2004). Return on marketing: Using customer equity to focus marketing strategy. *Journal of marketing, 68*(1), 109-127. https://doi.org/10.1509/jmkg.68.1.109.2403

Sagi, O., & Rokach, L. (2018). Ensemble learning: A survey. *Wiley Interdisciplinary Reviews-Data Mining and Knowledge Discovery, 8*(4). https://doi.org/10.1002/widm.1249

Satrya, W. F., Aprilliyani, R., & Yossy, E. H. (2023). Sentiment analysis of Indonesian police chief using multi-level ensemble model. *Procedia Computer Science, 216*, 620–629. https://doi.org/10.1016/j.procs.2022.12.177

Severinsen, A., & Myrland, Ø. (2022). Statistical learning to estimate energy savings from retrofitting in the Norwegian food retail market. *Renewable and Sustainable Energy Reviews, 167,* 112691. https://doi.org/10.1016/j.rser.2022.112691

Siegel, E. (2013). *Predictive analytics: The power to predict who will click, buy, lie, or die*. John Wiley & Sons.

Shum, H., He, X., & Li, D. M. (2018). From Eliza to XiaoIce: Challenges and opportunities with social chatbots. *Frontiers of Informaion Technology & Electronic Engineering, 19*(1), 10–26. https://doi.org/10.1631/fitee.1700826

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing and Management, 45*(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Sun, K., He, M., Xu, Y., Wu, Q., He, Z., Li, W., Liu, H., & Pi, X. (2022). Multi-label classification of fundus images with graph convolutional network and LightGBM. *Computers in Biology and Medicine, 149*, 105909. https://doi.org/10.1016/j.compbiomed.2022.105909

Supriya, M., & Deepa, A. J. (2020). Machine learning approach on healthcare big data: A review. *Big Data & Information Analytics, 5*(1), 58–75. https://doi.org/10.3934/bdia.2020005

Tarwidi, D., Pudjaprasetya, S. R., Adytia, D., & Apri, M. (2023). An optimized XGBoost-based machine learning method for predicting wave run-up on a sloping beach. *MethodsX*, 102119. https://doi.org/10.1016/j.mex.2023.102119

Teichert, T., Shehu, E., & Von Wartburg, I. (2008). Customer segmentation revisited: The case of the airline industry. *Transportation Research Part A-policy and Practice, 42*(1), 227–242. https://doi.org/10.1016/j.tra.2007.08.003

Verhoef, P. C., Kannan, P., & Inman, J. J. (2015). From multi-channel retailing to omni-channel retailing. *Journal of Retailing, 91*(2), 174–181. https://doi.org/10.1016/j.jretai.2015.02.005

Verhoef, P. C., Kooge, E., & Walk, N. (2016). Creating value with big data analytics: Making smarter marketing decisions. *Routledge*.

Verhoef, P. C., Lemon, K. N., Parasuraman, A., Roggeveen, A. L., Tsiros, M., & Schlesinger, L. A. (2009). Customer experience creation: Determinants, dynamics and management strategies. *Journal of Retailing, 85*(1), 31–41. https://doi.org/10.1016/j.jretai.2008.11.001

Verhoef, P. C., Reinartz, W., & Krafft, M. (2010). Customer engagement as a new perspective in customer management. *Journal of Service Research, 13*(3), 247–252. https://doi.org/10.1177/1094670510375461

Wang, G., Chen, G., & Chu, Y. (2018). A new random subspace method incorporating sentiment and textual information for financial distress prediction. *Electronic Commerce Research and Applications, 29*, 30–49. https://doi.org/10.1016/j.elerap.2018.03.004

Watson, H. J. (2014). Tutorial: Big data analytics: Concepts, technologies, and applications. *Communications of the Association for Information Systems, 34*(1), 65.

Weerasinghe, K., Pauleen, D. J., Scahill, S., & Taskin, N. (2018). Development of a theoretical framework to investigate alignment of big data in healthcare through a social representation lens. *Australasian Journal of Information Systems, 22*. https://doi.org/10.3127/ajis.v22i0.1617

Xia, Y., Li, Y., He, L., Xu, Y., & Meng, Y. (2021). Incorporating multilevel macroeconomic variables into credit scoring for online consumer lending. *Electronic Commerce Research and Applications, 49*, 101095. https://doi.org/10.1016/j.elerap.2021.101095

Xu, J., Zhang, Y., & Miao, D. (2020). Three-way confusion matrix for classification: A measure driven view. *Information Sciences, 507*, 772–794. https://doi.org/10.1016/j.ins.2019.06.064

Yousaf, A., Umer, M., Sadiq, S., Ullah, S., Mirjalili, S., Rupapara, V., & Nappi, M. (2021). Emotion recognition by textual tweets classification using voting classifier (LR-SGD). *IEEE Access, 9*, 6286–6295. https://doi.org/10.1109/access.2020.3047831

Zhan, C., Li, J., Jiang, W., Sha, W., & Guo, Y. (2020). E-commerce sales forecast based on ensemble learning. https://doi.org/10.1109/ispce-cn51288.2020.9321858

Zhang, C., & Ma, Y. (Eds.). (2012). Ensemble machine learning: Methods and applications. *Springer Science & Business Media*.

Zhang, Z., Srivastava, P. R., Sharma, D., & Eachempati, P. (2021). Big data analytics and machine learning: A retrospective overview and bibliometric analysis. *Expert Systems With Applications, 184*, 115561. https://doi.org/10.1016/j.eswa.2021.115561

Zhao, S., Xu, K., Yang, X., Liang, C., Lu, W., & Chen, B. (2022). Financial distress prediction by combining sentiment tone features. *Economic Modelling, 106*, 105709. https://doi.org/10.1016/j.econmod.2021.105709

Zhou, J., Li, W., Wang, J., Ding, S., & Xia, C. (2019). Default prediction in P2P lending from high-dimensional data based on machine learning. *Physica D: Nonlinear Phenomena, 534*, 122370. https://doi.org/10.1016/j.physa.2019.122370

Zhou, L., Fujita, H., Ding, H., & Ma, R. (2021). Credit risk modeling on data with two timestamps in peer-to-peer lending by gradient boosting. *Applied Soft Computing, 110*, 107672. https://doi.org/10.1016/j.asoc.2021.107672