Unveiling Smartphone User Perceptions: A Comparative Aspect based Sentiment Analysis of iOS and Android Customer Reviews

by

Chirag Hassan Shashidhara

2023

A Dissertation presented in part consideration for the degree of
"MSc Business Analytics"

ABSTRACT

In the ever-evolving landscape of smartphone technology, the rivalry between iOS and Android platforms has captivated users' attention, sparking curiosity about user sentiments, predictive capabilities, and innovation insights. This thesis embarks on a comprehensive exploration, unearthing the facets that distinguish these platforms in the eyes of their users.

The journey commences with an in-depth methodology that gathers, pre-processes, and analyses customer reviews for iOS and Android smartphones. A notable revelation emerges as Android reviews manifest an inclination towards lengthier feedback, whereas iOS reviews exhibit conciseness. These varying review lengths hint at differing user tendencies, possibly stemming from platform usage patterns, user demographics, or technical issues.

The exploration into sentiments unravels a captivating narrative, as sentiment analysis becomes the lens through which the essence of user opinions is distilled. Distinctive patterns in sentiments surrounding key aspects emerge—Android users tend to emphasize camera, battery life, screen, size, and performance, while iOS users foreground battery life, followed by display, price, and camera. These nuanced sentiment landscapes illuminate the diverse priorities of users across platforms and furnish manufacturers with actionable insights for targeted enhancements.

The journey culminates in the realm of predictive modelling, where machine learning prowess discerns predictive markers within sentiments. The journey of model selection leads to the Random Forest Classifier, which deftly classifies customer reviews as iOS or Android with an accuracy of 67.42%. This predictive prowess underscores the viability of leveraging sentiments as predictive indicators, propelling marketing strategies and product advancements based on real-time user opinions.

In summation, this thesis encapsulates a voyage that unfurls the narratives, nuances, and navigational pathways in the iOS and Android landscape. User sentiments, predictive modelling, and innovation insights converge to form a roadmap for manufacturers and marketers to steer their strategies, showcasing the potential of sentiments as beacons guiding product trajectories toward enhancing the human experience.

Table of Contents

1. INTRODUCTION

In today's rapidly evolving technological landscape, smartphones have transcended their status as mere communication devices and emerged as indispensable tools that shape the way we interact, work, and experience the world around us. With their multifaceted capabilities, smartphones have seamlessly integrated into our daily lives, serving as personal assistants, entertainment hubs, and gateways to the digital world. This ubiquity has led to fierce competition between iOS and Android, the two major platforms that dominate the global smartphone market. Each platform offers a distinct ecosystem, features, and attributes that cater to the diverse preferences and needs of users. As these platforms continue to innovate and evolve, the importance of understanding user sentiments and preferences becomes increasingly evident. Manufacturers, marketers, and researchers alike seek insights that can guide product development, refine marketing strategies, and cultivate user satisfaction. In this context, an in-depth exploration of customer reviews emerges as a valuable resource, offering a treasure trove of unfiltered user opinions that can shape the future trajectory of smartphone design and user experience.

1.1. Motivation for Research

The allure of smartphones lies not only in their technical prowess but also in their capacity to seamlessly integrate into our lives, serving as extensions of our identities and expressions. The symbiotic relationship between users and their smartphones creates a unique avenue for understanding user behavior, preferences, and expectations. With the ever-increasing competition between iOS and Android, both platforms strive to capture the hearts and loyalty of users by offering innovative features, intuitive interfaces, and diverse app ecosystems. However, the battle for dominance is not merely technical; it extends into the realm of user sentiments and experiences. In this dynamic landscape, the significance of comprehending user perceptions and sentiments cannot be overstated.

The motivation behind this research is rooted in the recognition that the voices of users, expressed through their reviews, hold valuable insights that can shape the future of smartphone innovation. Unveiling the underlying sentiments, preferences, and concerns expressed by users in their reviews can provide a nuanced understanding of their experiences and expectations. This knowledge, in turn, can guide manufacturers in optimizing their products to meet user needs more effectively. Additionally, insights from user sentiments can inform marketing strategies, helping manufacturers tailor their messaging to resonate with the emotions and desires of their target audiences.

1.2. Research Questions

This thesis seeks to embark on a comprehensive exploration of user sentiments, perceptions, and predictive capabilities in the realm of iOS and Android smartphones. Through a multifaceted analysis of customer reviews, this research aims to answer two pivotal research questions:

i. *Are there discernible differences in sentiments around the major aspects of customer reviews between iOS and Android smartphones?*

This research question delves into the heart of user reviews, aiming to unearth any disparities in sentiments expressed by users of iOS and Android smartphones. By harnessing advanced sentiment analysis techniques, this research endeavors to uncover whether users' sentiments differ across key aspects such as battery life, performance, camera quality, screen, and other pertinent attributes. The investigation into these nuanced sentiments offers a window into the preferences, priorities, and pain points of users on both platforms. It holds the potential to reveal insights that can inform targeted improvements, strategic decisions, and the prioritization of features.

ii. *Can a predictive machine learning model be developed to accurately classify a customer review as pertaining to an iOS or Android-based smartphone?*

The second research question propels this thesis into the realm of predictive analytics, leveraging the power of machine learning to ascertain the feasibility of discerning the brand of a smartphone from the sentiments expressed in customer reviews. By employing sophisticated machine learning algorithms, this research seeks to build a predictive model capable of accurately classifying a review as belonging to either an iOS or Android smartphone. This attempts to not only showcase the potential of machine learning to detect subtle patterns in user sentiments but also opens avenues for improved customer engagement, targeted marketing strategies, and informed product enhancements.

In navigating the intricate terrain of user sentiments, aspects of innovation, and predictive modelling, this thesis embarks on a transformative journey that holds implications for smartphone manufacturers, marketers, and researchers alike. By deciphering insights from the mosaic of customer reviews, this research aims to enhance the landscape of smartphone design, user experience, and the art of predictive analytics. As the boundaries between technology and human experience continue to blur, the findings of this research offer a compass to navigate the ever-evolving landscape of user preferences and expectations.

2. LITERATURE REVIEW

The technological landscape has been reshaped by the widespread adoption of smartphones, particularly those operating on the iOS and Android platforms. The viewpoints and attitudes of consumers towards these brands wield significant influence over producers, marketers, and researchers. Gaining insights into these sentiments provides a deeper understanding of the factors impacting customer satisfaction, purchasing patterns, and brand loyalty.

This literature review delves into recent investigations exploring consumer perspectives on various smartphone brands. We explore methodologies of sentiment analysis applied to customer reviews, uncovering insights into user preferences, satisfaction levels, and opinions. Additionally, we examine machine learning models that predict brand preferences based on sentiment patterns.

Our objective is to navigate this body of research and unveil the intricate tapestry of user sentiments. This exploration enhances our comprehension of the smartphone market and offers actionable suggestions to enhance user experiences through strategic product design and marketing approaches. As we embark on this journey, our aspiration is to craft a comprehensive narrative encompassing the collective perceptions, inclinations, and critiques of smartphone users at large.

2.1. Shifting Landscapes in Mobile Technology: Unveiling the Dynamics of Smartphone Preferences and User Sentiments

The world of mobile technology is changing rapidly, and smartphones are leading the way. These gadgets aren't just phones anymore – they're like mini computers that can do so much more. A bunch of studies have shown just how important this shift is.

Hakoama and Hakoyama (2011) looked into what people want from their mobile devices. They found that mobile phones are getting all sorts of cool features added to them, making them way more than just devices for chatting. The basic phone we used to know has transformed into something that does way more than just talk.

Singh and Goyal (2009) discovered something interesting about how different people buy phones in India. They noticed that age and gender play a role in what people care about when buying a phone. Young folks between 18 and 30 years old care more about how a phone looks, the brand, and all the extra things it can do. But older folks care more about the price and don't care as much about fancy tech stuff.

Chow et al. (2012) figured out that people are moving from simple phones to smartphones. They looked at what young adults aged 17 to 25 want in a phone. They found that things like what the phone can do, the price, the brand name, and even what their friends like, influence their decision to get a smartphone.

So, all these studies show that smartphones aren't just gadgets we talk on anymore. They've become something that's really personal and important to people. And different folks care about different things when they pick a phone – from how it looks to what it can do.

In our research, we examine the emotions expressed in customer reviews concerning various such elements like price, brand name, features, and design.

2.2. Consumer Preferences in Smartphone Selection: Focus on Operating System

According to studies, when purchasing a smartphone, consumers pay attention to both the operating system (OS) and the hardware features as well.

Cromar (2010) identified that potential smartphone buyers are faced with two pivotal decisions: selecting the preferred smartphone model and choosing the appropriate service provider. Among the determinants, the features embedded in the phone play a crucial role in the decision-making process. While 58% of consumers emphasized factors like size, camera, keyboard style, and price, a remarkable 72% of smartphone buyers directed their attention towards the operating system (OS) that powers the device. This shift in focus highlights the growing significance of the OS in shaping consumer preferences.

Isiklar and Buyukozkan (2007) assessed mobile phone preferences using a multi-criteria decision-making approach. Their analysis encompassed physical characteristics, technical features, functionality, brand choice, and customer excitement. Notably, functionality emerged as the most preferred factor across the various phones under scrutiny, while customer excitement and basic requirements had comparatively less influence.

Further emphasizing the importance of specific features, Ling et al. (2006) investigated mobile phone design attributes such as camera, colour screen, voice-activated dialling, Internet browsing, and wireless connectivity. Their findings underscored that features like colour screen, voice-activated dialling, and Internet browsing significantly influenced user satisfaction.

The surge in Android's popularity also came under scrutiny. Mokhlis and Yaakop (2012) examined Malaysian consumer preferences, unveiling that innovative features related to the OS, image, price, personal recommendation, durability, portable aspects, media influence, and post-sale service were key factors driving smartphone selection.

In the US, Nielsen's survey (2011) revealed a shifting preference landscape among smartphone consumers. Android emerged as the preferred OS for 31% of respondents, surpassing Apple's iOS (30%) and RIM Blackberry (11%). Similarly, Apple's own study (2012) highlighted consumers' differentiators, including the Google brand, larger screens, android market preference, integration with Google services, and quality.

The brand aspect gained prominence in smartphone selection as well. Louie (2012) found that 47% of US smartphone users owned an android device, with Samsung enjoying the highest brand preference, followed by Motorola, HTC, and LG.

The significance of brand personality emerged as a crucial factor in consumer choices. Azzawi and Ezeh (2012) discovered that brand played a pivotal role in choosing between Apple and Samsung smartphones. Apple exhibited strong brand equity and loyalty, while the preference between the two brands was influenced by their alignment with users' personalities.

In conclusion, consumer preferences in smartphone selection are shaped by a blend of factors, including the OS and brand considerations. These studies collectively shed light on the intricate decision-making process that guides consumers as they navigate the expanding landscape of smartphone choices.

This research centres on assessing the strength of sentiments related to various factors that impact smartphone buying decisions, with the ability to predict the operating system used by the reviewer, either iOS or Android.

2.3. Consumer reviews quality on Shopping sites

The prevalence of micro-blog sites, product review platforms, and social media channels has ushered in abundant opportunities for consumers to voice their opinions, feedback, and commentary on various products. This form of electronic word-of-mouth (eWOM) communication has emerged as a pivotal strategy for numerous marketers. An emphasis on sales value prompts firms to diligently monitor eWOM activities (D'Arbelles et al., 2020). To safeguard brand trust, providing timely responses on online platforms has gained significant importance for companies (Bhandari and Rodgers, 2018).

Despite their popularity, social media platforms are susceptible to concerns about source credibility and content quality. In contrast, blogs exhibit the capacity to normalize counterarguments concerning products (Pant et al., 2014). A comparison between social media and shopping sites reveals that eWOM information on shopping platforms enjoys greater trust than information on social media. Shopping sites tend to receive higher ratings in terms of content quality and source credibility (Erkan and Evans, 2018). These online reviews and product insights are instrumental in reducing consumer uncertainty during the purchase journey (Sharma et al., 2011), while also influencing the way consumers patronize retailers and counteract negative narratives about them (Chatterjee, 2001).

The data for this research is sourced from e-commerce websites like Amazon and Flipkart, as well as a dedicated online platform focused on providing information and reviews about mobile phones, GSM Arena.

2.4. Text Analysis on Consumer reviews

Given the context of the modern consumer product landscape, characterized by extensive product variety and abundant data, France and Ghose (2016) emphasize the necessity of data-driven tools for analyzing brand competition. A limited number of studies have employed data-driven tools to examine brand positioning through the analysis of online review text.

The expanding landscape of text analytics applications is reshaping our understanding of online reviews through novel metrics, offering fresh insights into consumer attitudes. A notable metric in this realm is sentiment analysis, which examines consumer sentiment - the feelings, consciousness, and thoughts associated with a subject (Bali et al., 2016). Scholars are delving into consumer responses to products and services based on online reviews, feedback, and blogs (Lin et al., 2017). Text analytics tools such as word clouds (Miley and Read, 2011), word storms (Castellà and Sutton, 2014), word-extraction algorithms (Barth et al., 2014), sentiment score analysis methodologies (Lakshmi et al., 2017), and emoji analysis frameworks (Li et al., 2018) provide researchers with avenues to explore the landscape of online consumer reviews.

The length of reviews, a key message attribute, has garnered recent attention in studies. Investigations into review length have established its correlation with content quality (Hu et al., 2008), and subsequently, its association with sentiment scores (Ghasemaghaei et al., 2018). However, the length of a review doesn't emerge as a pivotal factor in predicting the review's helpfulness (Singh J P et al., 2017).

Within our study, text review length and word clouds are employed to contrast fundamental distinctions between comments about iOS and Android, while *sentiment analysis is utilized to conduct a more comprehensive investigation. This analysis also serves as a preliminary step in generating predictive input features, for the machine learning classifier.*

2.5. Sentiment Analysis on Consumer reviews

Sentiment Analysis (SA), also known as Opinion Mining (OM), operates at the crossroads of Information Retrieval (IR), Natural Language Processing (NLP), and Machine Learning (ML) (Bing, L., 2012). This multidisciplinary field encompasses tasks like lexicon generation, subjectivity detection, and polarity detection, collectively aimed at unravelling the sentiments within textual content.

(Mowlaei, M.E. et al., 2020) Lexicon generation involves crafting a lexicon for feature creation in SA. This lexicon constitutes a compilation of words accompanied by their respective sentiment scores. Typically, established sentiment scoring systems utilize a symmetric range of integral or continuous values, such as −5 to +5. In these lexicons, negative scores signify negative sentiments, while positive scores indicate positive sentiments. Scores around zero indicate objectivity, while those significantly deviating in either direction reflect subjectivity. For instance, sentiment scores of −0.7 and −0.93 for words like "bad" and "terrible," respectively, convey that "terrible" holds stronger negative sentiment than "bad."

Subjectivity detection classifies reviews into objective or subjective categories (Keshavarz & Abadeh, 2016). Polarity detection, then, focuses on subjective reviews, determining if they lean towards an overall positive or negative sentiment. These analyses can be applied at varying textual levels, as defined by (Bing, L., 2012), spanning document, sentence, and entity levels, with additional levels identified by Ravi and Ravi (2015), encompassing word, concept, phrase, clause, and more.

Lexicons can be static or dynamic (Keshavarz & Abadeh, 2017). While static lexicons are precompiled with human supervision, dynamic lexicons are generated on-the-fly using ML on text corpora. Static lexicons provide reliable scores but lack the context sensitivity and adaptability of dynamic lexicons. Dynamic lexicons capture evolving slangs from social media, an advantage unavailable to static counterparts. For instance, in the restaurant domain, the term "long" might indicate food preparation time with a negative sentiment, while in the digital asset domain, it could refer to battery life with a positive sentiment. This adaptability becomes particularly crucial in aspect-based sentiment analysis, where the evaluation of individual aspects can be highly domain-specific, requiring lexicons to mirror the nuances of these dynamic linguistic landscapes.

In this study Aspect-based sentiment analysis is used, which demonstrates its superiority in providing a more comprehensive understanding of sentiments. By delving into specific attributes, this approach uncovers a richer spectrum of sentiments associated with different elements of content, enhancing the precision and depth of sentiment interpretation.

2.6. Aspect Based Sentiment Analysis

Sentiment analysis, as well as general opinion assessment, can be conducted across various levels of granularity, ranging from broad to specific. These levels include document-level, sentence-level, aspect-level, and concept-level analyses (referenced from Hemmatian and Sohrabi 2019; Pontiki et al. 2016; Nakov et al. 2019; Rosenthal et al. 2019). The document-level examination (also known as text-level analysis as per Pontiki et al. 2016) focuses on comprehending the sentiment polarity of an entire document. This applies to different types of content such as reviews (Behdenna et al. 2018), news articles (Shirsat et al. 2017), posts, and tweets (Gurini et al. 2013). In this context, the information provided is rather generalized, condensing the sentiment of numerous sentences into a single positive or negative score. This scoring is usually based on a two-point or five-point scale (Nakov et al. 2019).

A more detailed examination is conducted through sentence-level sentiment analysis (mentioned in Appel et al. 2016). The objective here is to determine whether a sentence conveys a positive or negative sentiment. In the contemporary landscape, the distinction between techniques employed for document-level and sentence-level analysis isn't notably pronounced. However, the latter task often presents greater challenges, particularly when dealing with brief sentences or sentences lacking broader contextual cues. This observation is echoed by both (Neviarouskaya et al. 2007) and (Yadollahi et al. 2017), who highlight that the same sentence used in two different contexts might convey diametrically opposing sentiments.

An examination conducted at the document or sentence level yields valuable insights into user opinions regarding specific entities. However, these analyses fall short in revealing the precise elements, attributes, or characteristics of said entities that have influenced users' positive or negative sentiments. To unravel such insights, a more specialized form of analysis known as aspect-level (or aspect-based) sentiment analysis comes into play (referenced from Liu 2012). This level of analysis offers notably more precise outcomes when contrasted with document and sentence-level evaluations.

A comprehensive outline of the primary tasks and subtasks within Aspect-Based Sentiment Analysis (ABSA) has been introduced through the SemEval competition, particularly detailed in the work of Pontiki et al. (2016). According to this framework, a typical aspect-based analysis involves the following tasks:

(1) Identifying sentences containing opinions or subjectivity within the document. (2) Associating each sentence with an entity (also termed as an aspect category). (3) Determining the specific aspect of the entity that the sentence pertains to. (4) Extracting the linguistic expressions employed in the sentence to reference the identified aspect. (5) For each recognized aspect, evaluate its polarity, often represented on a binary scale, five-point scale, or real-number range (Pontiki et al. 2016).

The initial task within Aspect-Based Sentiment Analysis (ABSA) involves aspect extraction, which entails identifying specific elements or attributes within the text. Numerous studies, such as Zhang and Liu (2014), Rana and Cheah (2016), and Hemmatian and Sohrabi (2019), have delved into the analysis and categorization of techniques for aspect extraction. Notably, Rana and Cheah (2016) conducted a comprehensive assessment of recent methods, identifying three primary categories: unsupervised, semi-supervised, and supervised approaches.

In the unsupervised category, techniques are often based on methods such as frequency or statistics (utilized by researchers like Hu and Liu 2004; Bafna and Toshniwal 2013; Rana and Cheah 2018; Wang et al. 2015; Luo et al. 2015). Additionally, heuristic approaches like those proposed by Singh et al. (2013) or the work of Bancken et al. (2014), which utilizes syntactic dependency paths to identify entities, are found in this category. Poria et al. (2014) also adopts a rule-based approach.

Moving to the semi-supervised category, techniques incorporate lexicon-based methods (explored by Yan et al. 2015; D'Aniello et al. 2018; Shah and Swaminarayan 2021; Klyuev and Oleshchuk 2011), as well as approaches using dependency trees (as demonstrated by Yu et al. 2011) and graph-based methods (such as the work of Xu et al. 2013). In contrast, supervised techniques predominantly leverage machine learning models, including random fields, Support Vector Machines (SVM) (as seen in Manek et al. 2017), decision trees, neural networks, and autoencoders (Angelidis and Lapata 2018; Tomasiello 2020).

Following aspect extraction, the subsequent task involves sentiment identification, where each extracted aspect is assigned a sentiment score denoting its polarity or orientation within the text (as outlined in Pontiki et al. 2016). These scores typically fall within a numerical range, indicating the degree of positivity or negativity associated with the opinion regarding the aspect. This range could be expressed using scales like the five-point scale (Nakov et al. 2016) or represented as a decimal value within the range of -1 to +1, where -1 signifies strong negativity and +1 denotes strong positivity. When the objective is solely to ascertain whether the sentiment is positive or negative, this task is also referred to as sentiment classification.

In this study, we conduct aspect-based sentiment analysis utilizing a tool named pyABSA, and the subsequent section provides a brief overview of this tool's functionality.

2.7. A short note on PyABSA

In the realm of Aspect-Based Sentiment Analysis (ABSA), a plethora of models have been proposed. These models often exhibit distinct architectures such as LSTM, GCN, and BERT, as well as unique optimizations encompassing aspects like data pre-processing and evaluation metrics. However, reproducing the reported outcomes of these models, even when their source code is available, can prove challenging due to these variations.

To tackle this challenge and foster equitable comparisons, Yang and Li (2022) have introduced PyABSA, a modularized framework founded upon PyTorch that facilitates reproducible ABSA. This framework is designed to streamline model training, evaluation, and inference across various subtasks within aspect-based sentiment analysis, boasting support for 29 different models and 26 datasets.

One of PyABSA's notable advantages is its user-friendly nature. It allows individuals, including those new to the field, to effortlessly reproduce a model's results on a specific dataset using only a few lines of code. Beyond simplicity, PyABSA follows a modularized organizational structure, comprising five major modules: template class, configuration manager, dataset manager, metric visualizer, and checkpoint manager. This modular architecture enhances flexibility, enabling the extension of provided templates to accommodate different models, datasets, and related tasks with minimal modifications.

This tool serves as the central component of this research, facilitating sentiment comparisons across various aspects of iOS and Android smartphones, while also aiding in the construction of a data frame for the classifier.

2.8. Unveiling Machine Learning Marvels: Navigating Customer Review Analysis

In the landscape of machine learning models tailored for customer reviews, the primary objectives encompass sentiment prediction, user group classification, rating forecasts, and more. Notably, a study by Elli, et al. 2016 delved into this domain, extracting sentiment from reviews and employing the insights to construct a robust business model. Their work showcased the efficacy of the utilized tools, yielding high accuracy and informed decisions through the integration of business analytics. Their research extended to emotion detection from reviews, gender inference based on names, and counterfeit review identification. Python and R emerged as the predominant programming languages, with Multinomial Naïve Bayesian (MNB) and support vector machine (SVM) serving as pivotal classifiers.

In another instance Xu et al. 2015, an author applied existing supervised learning algorithms to forecast review ratings on a numerical scale using solely text input. Employing a holdout cross-validation approach with a 70-30 training-testing data split, the study explored various classifiers to ascertain precision and recall values.

Similarly, in a subsequent study Rain 2013 expanded the scope of natural language processing and sentiment analysis to Amazon review datasets. Naïve Bayesian (NB) and decision list classifiers were employed to tag reviews as positive or negative, focusing on Amazon's books and Kindle section. Data scraping facilitated data acquisition, followed by pre-processing. Bhatt et al. 2015 aimed to create a visualization system that translates review sentiment into charts, utilizing data scraped from Amazon URLs. This paper applied NB, SVM, and maximum entropy, focusing on summarizing product reviews for visualization, with a statistical chart illustrating the results.

Another research endeavour Shaikh, T. and Deshpande, D., 2016 focused on various techniques for feature extraction and selection in sentiment analysis. Collecting an Amazon dataset, pre-processing included the removal of stop words and special characters. This work encompassed phrase-level, single-word, and multiword feature extraction or selection techniques, with Naive Bayes employed as the classifier.

Despite significant efforts invested in constructing machine learning classifiers for direct sentiment prediction or other features from reviews, following standard text cleaning procedures such as eliminating stop words and special characters, this study shifts its focus towards utilizing aspect-based sentiment analysis to engineer new features. The objective is to generate sentiment-weighted features aimed at predicting the operating system of the smartphone discussed in the user's comment. This predictive capability can be leveraged for various applications, including foreseeing the likelihood of a reviewer switching smartphones.

3. METHODOLOGY

The primary aim of this research is to conduct a comparative analysis of customer reviews for smartphone brands and develop a machine learning model capable of predicting the brand (iOS or Android smartphones) based on these reviews. To achieve this objective, it was essential to obtain a dataset containing Amazon and GSM arena customer reviews for various smartphones.

This chapter presents a comprehensive overview of the methodology employed at each stage of the research process. It outlines the steps taken to collect and analyze the customer reviews data, as well as the approach used to build and test the predictive machine learning model. By following these methodological insights, a solid foundation is established for the validity and reliability of the research findings.

3.1. Data acquisition

The present investigation seeks to juxtapose the brand perception of iOS and Android smartphones, along with the principal aspects highlighted by customers in their user reviews. To accomplish this, an all-encompassing dataset was procured, consisting of customer reviews pertaining to six distinct brands of smartphones. The datasets were sourced from Kaggle, and a concise overview of the acquired data is presented below.

The table 3.1.1 presents a summary of the data source, brand, model, number of customer reviews, and primary sources for smartphone models collected for the research. The table includes information on several smartphone models from different brands, along with the corresponding number of customer reviews and the primary sources of those reviews. Notably, brands such as Google, OnePlus, Oppo, Samsung, Xiaomi, and Apple are represented in the dataset, encompassing models like Pixel 5, 9 Pro, Find X3 Pro, Galaxy S21 Ultra, Mi11, Pixel 4a, Galaxy A10s, iPhone SE, and iPhone 11. The number of customer reviews varies significantly across the models, ranging from 158 to 9,713 reviews. The primary sources of these reviews include GSM Arena, Amazon, and Flipkart.

Data Source	Brand	Model	Number of customer reviews	Primary Source
Kaggle [1]	Google	Pixel 5	209	GSM Arena
Kaggle [2]	OnePlus	9 Pro	251	GSM Arena
Kaggle [3]	Oppo	Find X3 Pro	158	GSM Arena
Kaggle [4]	Samsung	Galaxy S21 Ultra	346	GSM Arena
Kaggle [5]	Xiaomi	Mi11	264	GSM Arena

Kaggle [6]	Google	Pixel 4a	1759	Amazon
Kaggle [7]	Samsung	Galaxy A10s	1672	Amazon
Kaggle [8]	Apple	iPhone SE	9713	Flipkart
Kaggle [9]	Apple	iPhone 11	5009	Amazon

Table 3.1.1: Sources from which data was acquired

Each dataset comprises multiple features, including review text, rating, username, review title, review URL, and more. However, this research solely focuses on extracting the review text feature from each dataset for analysis and examination.

3.2. Data preprocessing

Data preprocessing was conducted by importing the collected data as a pandas data frame, allowing for efficient data manipulation and analysis. The initial step involved checking the shape of the data frame to understand the size and dimensions of the dataset. Moreover, an examination of the data types was performed to ensure the appropriate data types were assigned to each column, facilitating accurate computations.

During the data inspection, it was observed that there were some missing values; however, the number of missing entries was minimal and did not significantly affect the overall dataset. Subsequently, a data imbalance issue surfaced, as there were over 14,000 reviews for Apple smartphones compared to only 4,500 reviews for Android smartphones. To address this imbalance and maintain equal representation of both classes, a random under-sampling method was employed to reduce the number of Apple smartphone reviews, creating a balanced dataset.

Before balancing the dataset, the text reviews underwent thorough cleaning procedures. Specifically, URLs, stop words, and emojis were removed from the review text. Eliminating URLs ensures that external links and web addresses do not interfere with the sentiment analysis process. Stop words, which are commonly occurring words with little semantic significance (e.g., "and," "the," "is"), were also removed to focus on relevant keywords and phrases. Additionally, emojis were removed as they may not provide direct insights into aspects and sentiments and could lead to inaccurate analysis.

After the data was cleaned, the processed text reviews were fed into the PyABSA model for aspect-based sentiment analysis.

3.3. Aspect based sentiment extraction

PyABSA (Python library for Aspect-based Sentiment Analysis) is a powerful tool that enables the extraction of aspects and sentiments from textual data. The model is designed to identify the main aspects or features being discussed in a review and determine the corresponding sentiment polarity (e.g., positive, negative, neutral) associated with each aspect.

By utilizing PyABSA, this research aimed to gain deeper insights into customer opinions and sentiments regarding different aspects of iOS and Android smartphones. The analysis performed by PyABSA allowed for a comprehensive understanding of how users perceive various features of these smartphones, shedding light on potential strengths and weaknesses of each brand's product offerings.

3.4. Data visualization

After obtaining the JSON output from the PyABSA model, the next step was to convert it into a pandas data frame for further analysis and visualization. The JSON output contained information about each review's aspects and corresponding sentiments.

To create the pandas data frame, the JSON data was processed, and the review text, aspects, and their respective sentiments were extracted and organized. The data frame was structured with three columns: "Review," "Aspects," and "Sentiment." Each row represented a unique aspect-sentiment pair associated with a review.

With the pandas data frame in place, a summary of the top aspects was generated to gather statistics on major aspects and their corresponding sentiments. This summary involved calculating the frequency of each aspect in the data frame and analyzing the distribution of sentiments associated with each aspect.

To visualize the findings, various charts were plotted. A bar chart was used to display the top aspects and their occurrence frequencies, providing a visual representation of which aspects were most commonly mentioned in the reviews. Another bar chart was used to show the sentiment distribution of each aspect, indicating the proportion of positive, negative, and neutral sentiments associated with the major aspects.

Additionally, a pie chart was used to depict the overall sentiment distribution across all aspects, showcasing the percentage of positive, negative, and neutral sentiments present in the reviews collectively.

The data visualization helped in identifying the key aspects that customers frequently discussed in their reviews and revealed the sentiments expressed toward these aspects. By presenting the information in a graphical format, the findings became more accessible and comprehensible, enabling researchers to draw meaningful insights from the aspect-based sentiment analysis of iOS and Android smartphones.

3.5. Feature engineering

To create a dataset suitable for feeding into a classifier, the aspects extracted from the PyABSA analysis, along with their corresponding sentiments, were used to form features for each review. The goal was to identify the top 15 aspects commonly mentioned in iOS and Android reviews and assign sentiment-based weights to these aspects for each review.

The first step involved selecting the top 15 aspects based on their occurrence frequencies in the reviews. These aspects were then used as feature labels for the data frame. For instance, if the top aspects included "Battery life," "Camera quality," "Performance," and so on, these would become the feature labels.

Next, sentiment-based weights were assigned to each aspect for each review in the dataset. The sentiment-based weight values were determined as follows:

- If the aspect was spoken of positively in the review, a weight of +1 was assigned.
- If the aspect was mentioned neutrally, a weight of +0.5 was assigned.
- If the aspect was spoken of negatively, a weight of -1 was assigned.
- If the aspect was not mentioned at all in the review, a weight of 0 was assigned.

For example, if a review contained the aspect "Camera quality" and the sentiment associated with it was positive, the corresponding feature value for that review would be +1. If the review did not mention "Camera quality" at all, the feature value for that aspect would be 0.

3.6. Model selection and prediction

To address the data imbalance issue observed in the dataset, random under-sampling was employed to create a balanced representation of iOS and Android reviews. This involved randomly removing some reviews from the majority class (iOS) to match the number of reviews in the minority class (Android). By balancing the data, the classifiers could make more accurate predictions for both classes without being biased towards the majority class.

After balancing the dataset, it was divided into training and testing sets using a train-test split. The training set was used to train the classifiers, while the testing set was reserved for evaluating their performance.

Three different classifiers were chosen for this study: Logistic Regression, Random Forest Classifier, and Naive Bayes. Each classifier was trained on the balanced training set using the sentiment-weighted features derived from the aspect-based sentiment analysis.

During the training phase, the classifiers learned the patterns and relationships between the features and the target variable (software prediction - iOS or Android). After training, the classifiers were evaluated on the balanced testing set to measure their performance and effectiveness in predicting the software category.

Upon evaluating the classifiers, it was found that the Random Forest Classifier achieved the best performance in terms of accuracy, precision, and recall. Consequently, it was selected as the preferred model for prediction.

To further optimize the Random Forest Classifier, hyperparameter tuning was performed using GridSearchCV. GridSearchCV exhaustively searched through a range of hyperparameters to find the combination that produced the best results. The hyperparameters tuned might include the number of estimators, maximum depth, minimum samples per leaf, and others, depending on the specific implementation of the Random Forest Classifier.

By fine-tuning the hyperparameters using GridSearchCV, the Random Forest Classifier's performance was optimized, leading to improved accuracy and better predictions on the test set.

Research Objective 1: The initial research question, " *Are there discernible differences in sentiments around the major aspects of customer reviews between iOS and Android smartphones?*", will be investigated through the utilization of review text lengths, creation of word clouds through data visualization, conducting aspect-based sentiment analysis and the SHAP outcomes generated by the ultimately trained machine learning model. By employing these techniques, a comparative analysis between the two smartphone platforms will be facilitated.

Research Objective 2: Addressing the second research inquiry, " *Can a predictive machine learning model be developed to accurately classify a customer review as* *pertaining to an iOS or Android-based smartphone?*", will involve employing the outcomes derived from feature engineering and training diverse machine learning models. These findings will be integral in the development of a predictive model geared towards determining the smartphone software mentioned in customer reviews.

4. FINDINGS

The preceding chapters have provided a foundation for this section by elaborating on the research methodology utilized to examine and analyse customer reviews pertaining to smartphone brands, as well as construct a machine-learning model for prediction. In this chapter, we present the results derived from these endeavours, thereby offering a comprehensive understanding of the perceptions, sentiments, and predictive capabilities uncovered over the course of our investigation.

The findings obtained from this study indicate that there exist discernible disparities in customer perceptions and sentiments towards iOS and Android smartphones.

4.1. Descriptive text analysis

The analysis of the dataset uncovered notable disparities in the average text lengths of reviews between the Android and iOS platforms. Specifically, reviews for Android demonstrated a significantly higher average text length of 301 characters, with a median of 163 characters, highlighting users' tendency to provide comprehensive and detailed feedback. In contrast, reviews for iOS displayed a substantially lower average text length of 91 characters, accompanied by a median of 46 characters, indicating a prevailing inclination towards concise expression of opinions.

Table 4.1 depicts the research findings pertaining to the average and median text lengths in terms of character count for reviews on Android and iOS smartphones respectively.

	Average text length	**Median text length**
Android	301 characters	163 characters
iOS	91 characters	46 characters

Table 4.2.1: Review text length averages and medians

Figures 4.1.1 and 4.1.2 exhibit histograms that visually represent the distribution of text lengths in reviews for Android and iOS smartphones, respectively.

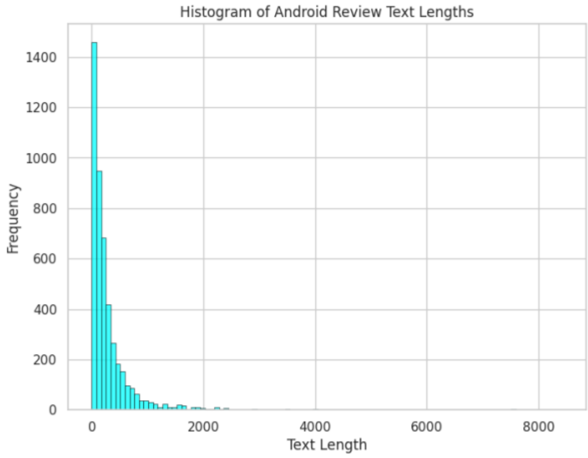


Figure 4.1.1: Android reviews - text length distribution

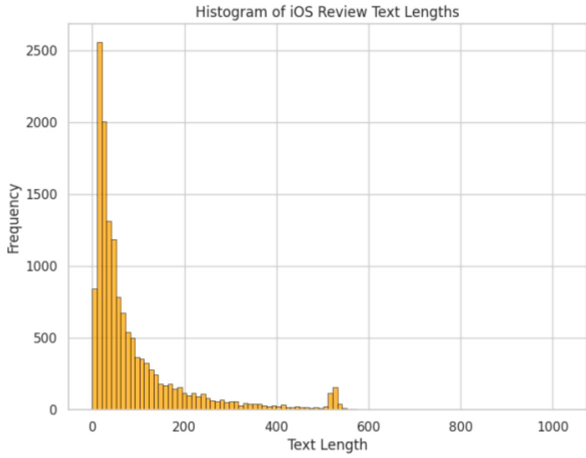


Figure 4.1.2: iOS reviews - text length distribution

Using pyABSA to analyse the facets within each review, a comprehensive examination of the key aspects for both Android and iOS smartphones was possible. The findings indicated that reviews for Android phones placed equal emphasis on camera quality, battery life, screen display, size, and overall performance. On the other hand, iOS reviews mainly focused on battery life as the primary consideration, with secondary factors including display quality, price affordability, and camera features. Additionally, synonymous topics were observed across a significant portion of reviews from both operating systems.

Figures 4.1.3 and 4.1.4 showcase word clouds that visually depict the dominant aspects frequently mentioned in reviews for Android and iOS smartphones, respectively.

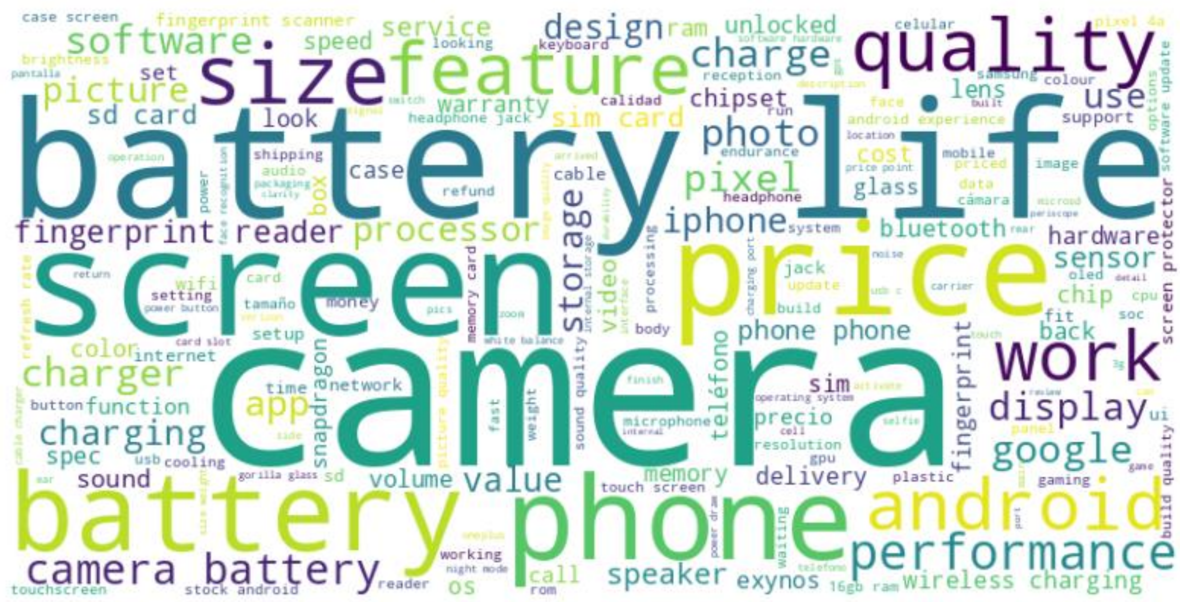


Figure 4.1.3: Android reviews – aspects word cloud

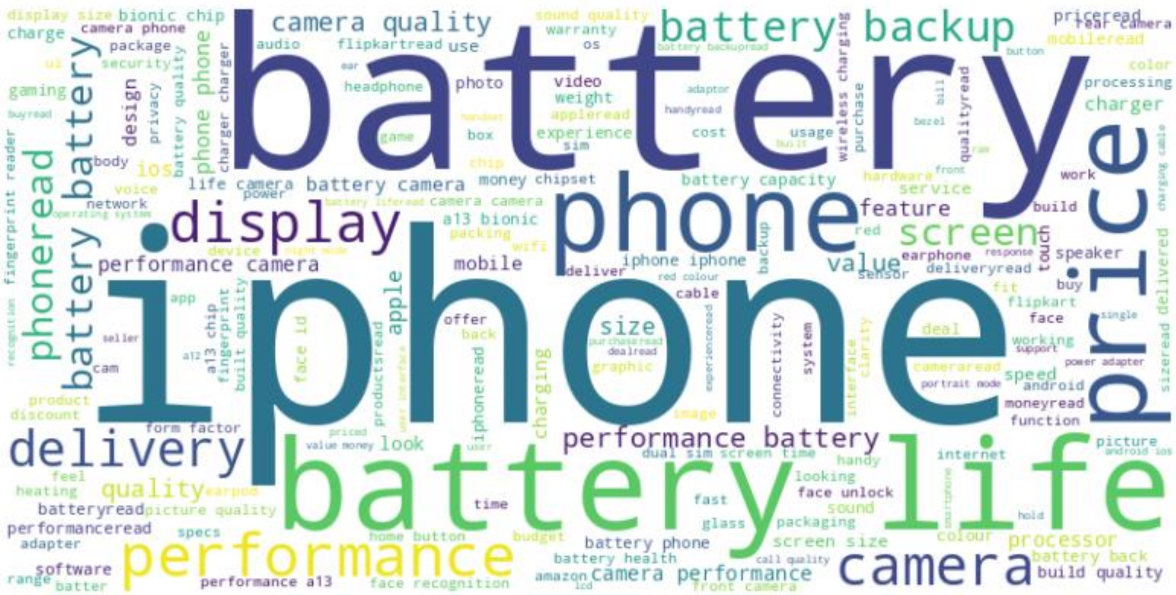


Figure 4.1.4: iOS reviews - aspects word cloud

Presented below is are bar charts (Figure 4.1.5 and 4.1.6) aimed at facilitating a more effective relative comparison among the aspects.

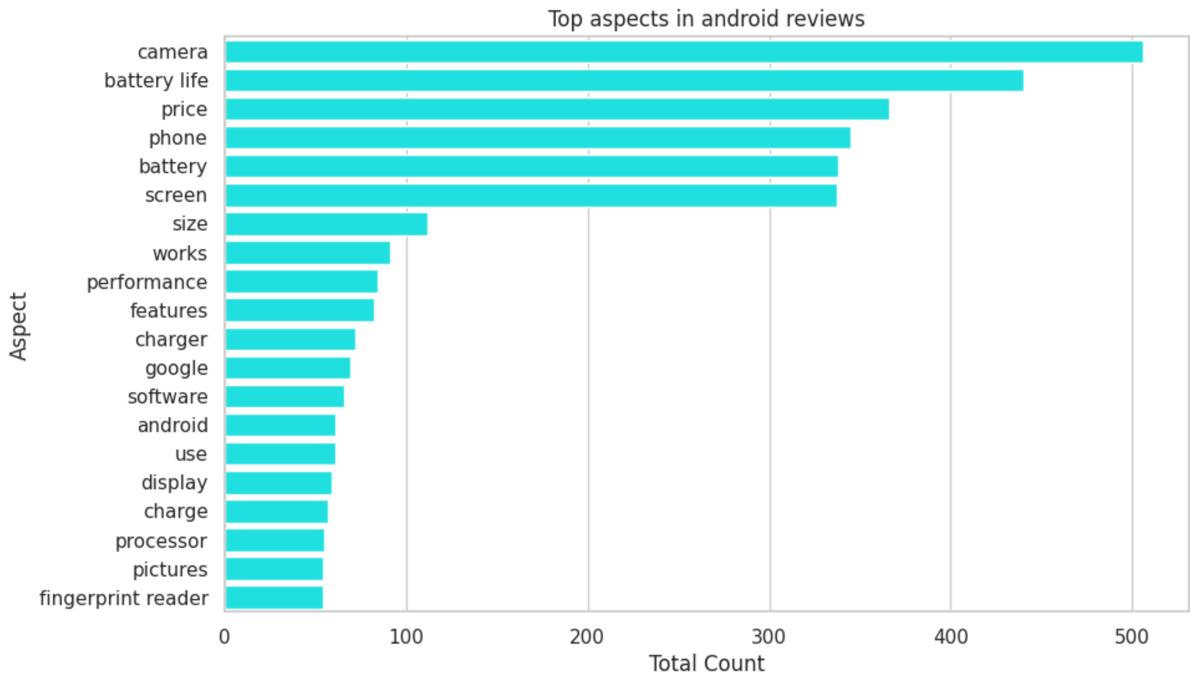


Figure 4.1.5: Android reviews - Top aspects distribution

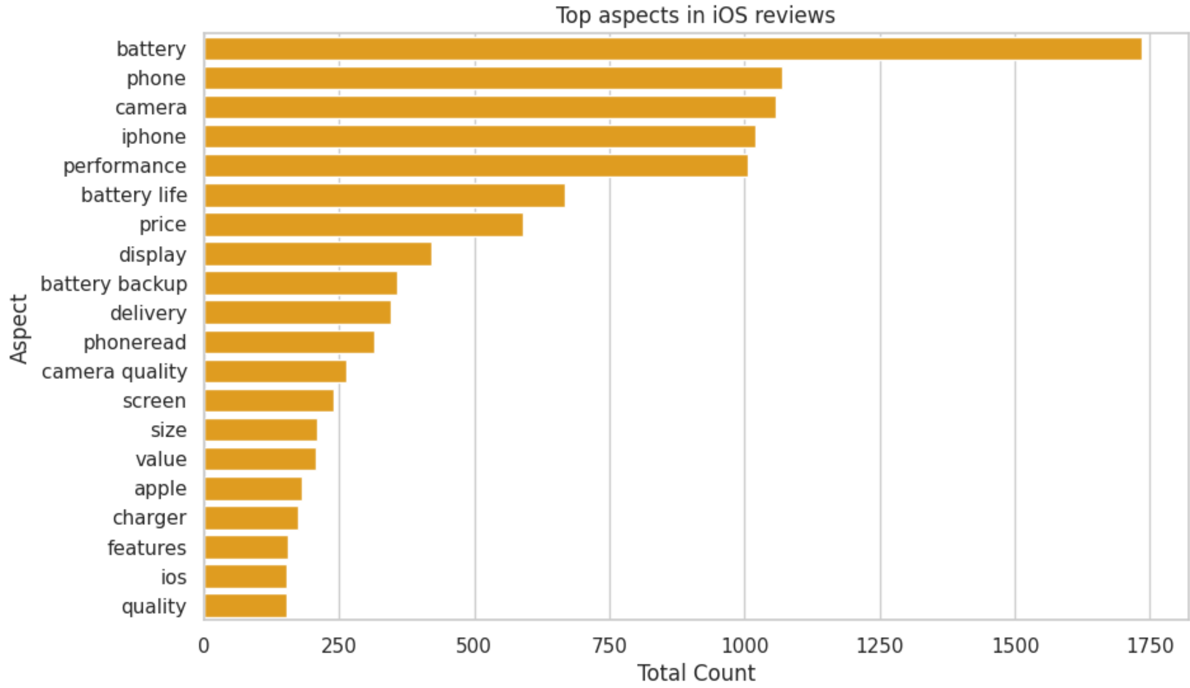


Figure 4.1.6: iOS reviews - Top aspects distribution

4.2. Aspect based sentiment analysis

The aspect-based sentiment analysis conducted in this study offers valuable insights into the sentiments expressed within customer reviews for both Android and iOS smartphones. This analysis focuses on extracting key aspects discussed in these reviews and linking sentiment polarities to each aspect. The findings presented in the following charts provide a concise overview of the sentiments associated with various aspects for both smartphone categories, shedding light on customer perceptions and preferences.

Upon closer examination, as shown in Figure 4.2.1, it becomes evident that iOS reviews demonstrate a convergence of 65.7% positive reviews addressing various aspects, along with 26.9% negative reviews discussing distinct facets. In contrast, Android reviews display a distribution of 56.9% positive feedback and 34.1% negative feedback, highlighting varying sentiments across different aspects.

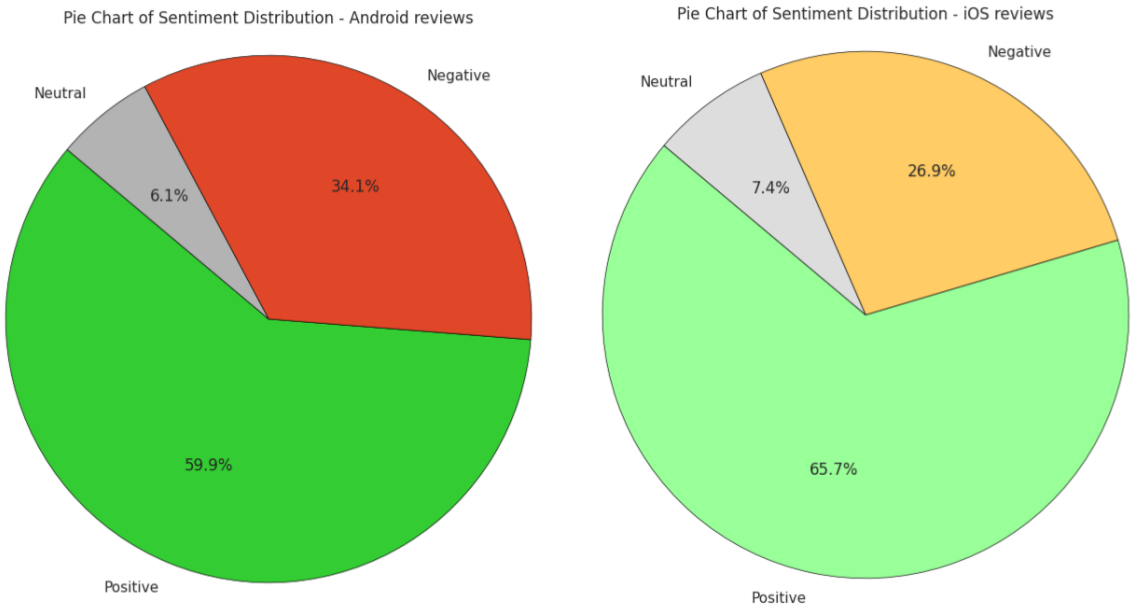


Figure 4.2.1: Distribution of different polarities of sentiments among Android and iOS reviews

Moreover, upon analysing the sentiments associated with each aspect, it was noted that in Android reviews, there were notable negative sentiments towards the screen and battery features. Conversely, positive sentiments were particularly evident for aspects like price and features. On the other hand, in iOS reviews, prevalent negative sentiments revolved around battery and charger features while aspects such as performance received highly positive feedback from users.

The following figures, Figure 4.2.2, and Figure 4.2.3, offer an in-depth presentation of this information.

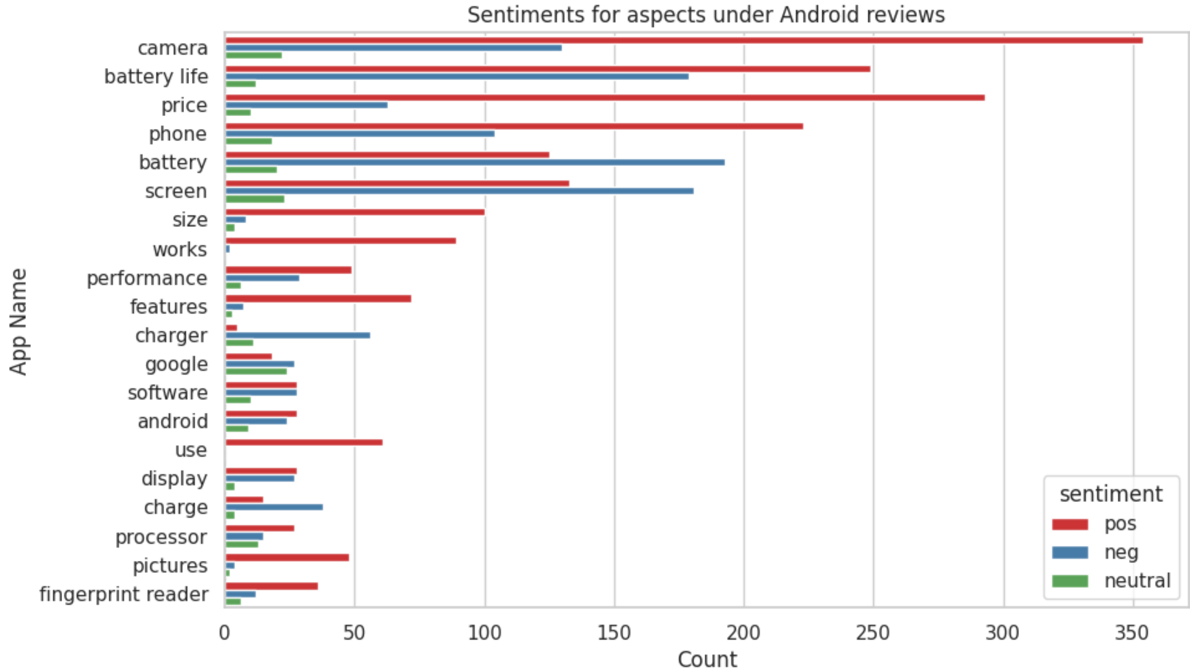


Figure 4.2.2: Android reviews: Sentiments around different aspects

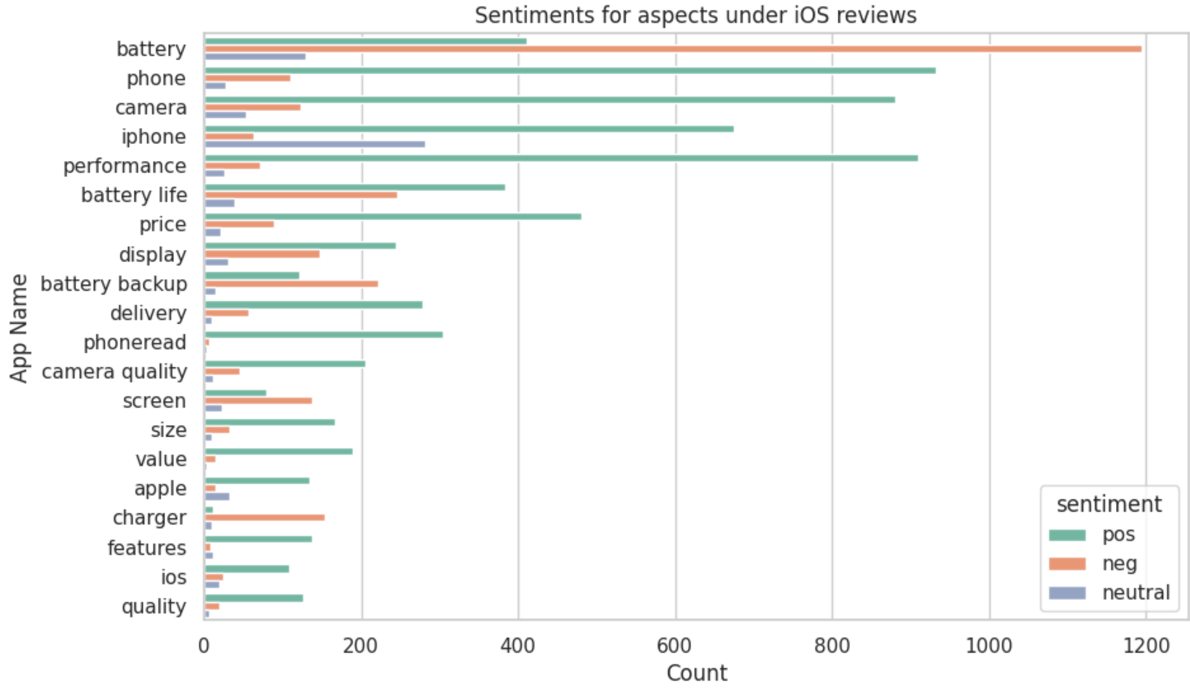


Figure 4.2.3: iOS reviews: Sentiments around different aspects

4.3. Machine learning prediction model performance

The subsequent stage of this research focuses on the evaluation and comparison of different machine learning models used for predicting smartphone brands based on sentiment-weighted features. This section examines the outcomes of diverse models, evaluating their accuracy, precision, and recall in brand prediction. From the examined models, one model stands out as the optimal choice due to its superior performance metrics. The ensuing presentation presents the results of each model's performance and delineates the criteria that led to selecting the most accurate predictive model.

The dataset utilized for brand prediction underwent meticulous pre-processing, which included under sampling reviews to ensure equal representation between iOS and Android. The final dataset consisted of 4174 rows. To effectively evaluate various machine learning models' accuracy in predicting smartphone brands, this balanced dataset was subjected to a stratified train-test split analysis with 80% of the data used for training and 20% used for testing.

The dataset was used to evaluate the predictive performance of three different models: Logistic Regression, Random Forest Classifier, and Naïve Bayes. Accuracy metrics were utilized to quantify their performance. Table 4.3.1 provides detailed information on the accuracy outcomes obtained from these experiments with the three models.

Model	**Accuracy Score**
Logistic Regression	63.13%
Random Forest Classifier	**68.02%**
Naïve Bayes Classifier	63.13%

Table 4.3.1: Accuracies of different classifiers

To achieve balanced predictions for both iOS and Android software, the accuracy score serves as a crucial criterion in this study. Unlike precision or recall, accuracy ensures an equitable representation of both software categories. Through experimentation with Logistic Regression, Random Forest Classifier, and Naïve Bayes algorithms, we identified the model that demonstrated the highest accuracy based on our results.

With the aim of refining the predictive capabilities further, the Random Forest Classifier was selected as the candidate for additional optimization. The subsequent step involved employing the grid search technique to fine-tune the hyperparameters of the Random Forest Classifier. This meticulous optimization process was undertaken to enhance the model's performance by achieving the best possible accuracy and ensuring a robust brand prediction mechanism. The table 4.3.2 elaborates on the efficacy of tis refinement process and the accuracy improvements achieved through the utilization of the grid search technique.

Hyperparameters tuned:

Hyperparameter	Optimal value
n_estimators	200
min_samples_split	10
min_samples_leaf	1
max_features	sqrt
criterion	gini

Table 4.3.2: Hyperparameters tuned using GridSearchCV

The utilization of these hyperparameters yielded an optimal accuracy of 67.42%. While this value is slightly below the previous accuracy score, it holds greater robustness due to the incorporation of cross-validation.

4.4. SHAP values – Feature contribution

Building upon the refined Random Forest Classifier that was fine-tuned using optimal parameters, the investigation now delves into the realm of SHAP values. These values serve as insightful tools to uncover the influence of individual features on the predictive outcomes of the model. By analysing the SHAP values derived from the Random Forest Classifier, this section sheds light on the extent and direction of each feature's contribution to the brand prediction process. The subsequent presentation of SHAP value findings elucidates the crucial role each feature plays in determining the classification outcomes and offers a nuanced understanding of their relative impact.

Upon examining the summary plot of SHAP values, it becomes evident that the most influential feature pertains to sentiments surrounding the brand itself. This is closely trailed by sentiments regarding performance, followed by considerations for screen quality, battery life, and value for money. As discussed in the methodology section, within the sentiment weighting framework, higher feature values indicate positive sentiments, while lower values reflect negative sentiments.

The sentiment dynamics play a pivotal role in the model's predictions. Higher positive sentiment directed towards the brand tends to lead to predictions of iOS reviews, mirroring a similar trend for performance. Conversely, screen-related sentiments lean towards Android phones, signifying positive sentiment in this context. Battery aspects emerge as a significant influence, with iOS garnering more negative reviews compared to the positive sentiments associated with Android. Furthermore, Android phones exhibit positive sentiments in relation to their value for money.

The comprehensive Figure 4.4.1 below provides a visual representation that encapsulates these insights, unravelling the intricate interplay between different aspects and sentiments.

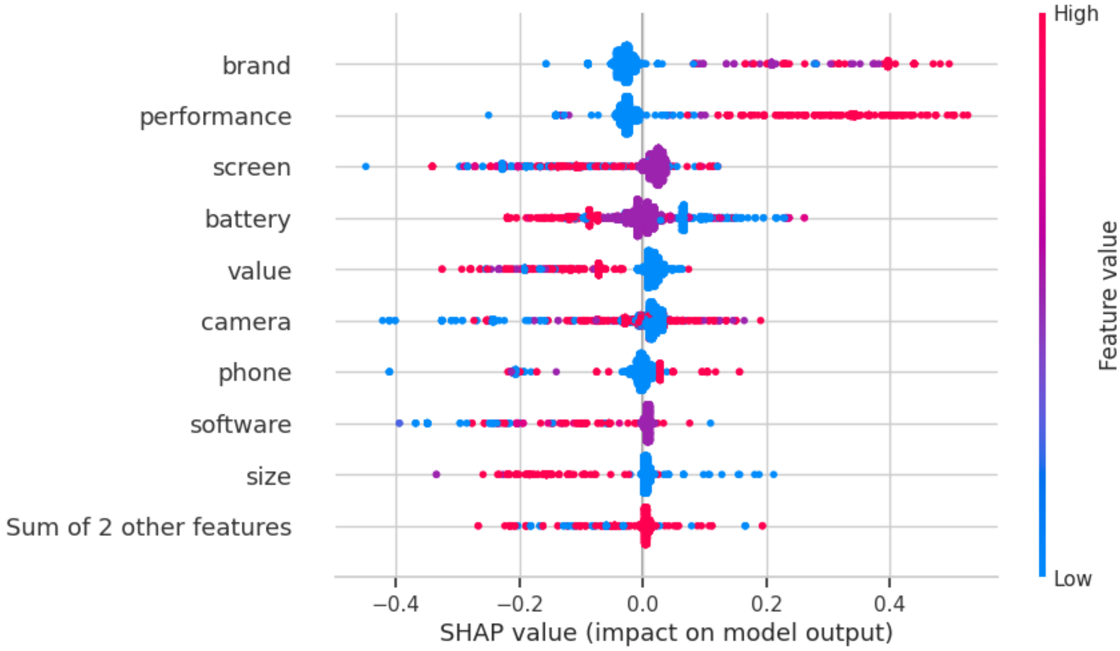


Figure 4.4.1: SHAP summary plot

5. DISCUSSION

In today's world, smartphones have become an integral part of our lives. They have revolutionized the way we communicate, work, and entertain ourselves. With the growing popularity of smartphones, the market has been dominated by two major platforms - iOS and Android. These platforms have their own unique features and characteristics that attract different types of users.

In our current research, we are exploring the potential differences in customer perceptions and sentiments of innovation between iOS and Android smartphones. We aim to analyse user reviews and investigate various aspects such as text length, expressed sentiments, and the ability to predict the brand based on these sentiments. By conducting this research, we seek to gain insights that can help smartphone manufacturers optimize their products, marketing strategies, and customer experiences.

5.1. Text Length of Reviews: More Than Meets the Eye

One intriguing finding of this research is the significant difference in text length between reviews of Android and iOS smartphones. The analysis revealed that Android reviews tend to have a notably higher average text length compared to iOS reviews. This discovery implies that Android users are more inclined to provide detailed and comprehensive feedback about their smartphones.

There could be various reasons behind this variation in review length. One potential explanation is the different ways users interact with the two platforms. Android offers a higher degree of customizability, enabling users to have more control over their overall experience. This increased control may empower Android users to express their opinions on specific aspects of the device, resulting in lengthier reviews.

Another factor that might contribute to the longer text length in Android is that the user demographics between the two platforms vary. Android tends to attract power users and developers who often provide detailed feedback about their devices due to their expertise and higher engagement with the platform. In contrast, iOS is more popular among casual users who may be less inclined to provide detailed feedback.

Bugs and software issues can significantly impact user experiences and may be a contributing factor to the text length discrepancy. Users on a platform experiencing more bugs or technical issues may feel compelled to provide more detailed feedback in order to document and report these issues. This can be particularly true for Android users who are known for their technical expertise and active involvement in the platform. They may have a higher expectation for bug-free software and may be more willing to share detailed information about their encounters with bugs, including steps to reproduce them or potential workarounds. In contrast, iOS users, who generally have a smoother and more controlled user experience, may encounter fewer bugs and therefore have less reason to provide extensive details in their feedback.

5.2. Sentiments Expressed in Reviews: A Tale of Differences

Apart from the difference in text length, the research delved into the sentiments expressed in reviews of Android and iOS smartphones. Through aspect-based sentiment analysis, significant distinctions in the sentiments expressed for various features of the two platforms was discovered.

One interesting finding is that the most pronounced negative sentiments in Android reviews were directed towards the screen aspect. This could be due to the fact that Android smartphones generally have larger screens than iOS devices, which can lead to problems with screen quality or battery life. Additionally, Android users may have higher expectations for their smartphones' screen quality due to the platform's reputation in this area.

Another interesting finding is that the most prevalent negative sentiments in iOS reviews were centered around the battery and charger aspects. This could be due to the fact that iOS devices typically have smaller batteries than Android devices, which can lead to shorter battery life.

Yet another important difference between the two platforms is performance. iOS devices are generally considered to have better performance than Android devices. This is due to a number of factors, including the use of a custom-designed processor and operating system.

However, there are also some similarities between the two platforms. For example, both platforms have good cameras, with iOS devices generally being considered to have better image quality.

5.3. Predicting the Brand of a Smartphone: Sentiments as a Valuable Indicator

The findings of this research suggest that sentiment analysis can be used as an effective indicator to predict the brand of a smartphone in customer reviews. The Random Forest Classifier model, which was fine-tuned using the grid search technique, achieved an accuracy of 67.42% in predicting the brand of a smartphone based on the sentiment-weighted features of the reviews.

This suggests that the sentiment of the reviews can be used to identify the different aspects of the smartphone that are important to users, and that these aspects can be used to distinguish between iOS and Android devices. For example, the reviews of iOS devices were more likely to express positive sentiments about the performance of the devices while the reviews of Android devices were more likely to express positive sentiments about the price and features of the devices.

Also the SHAP values provide valuable insights into the topmost sentiment weighted aspects that the model relied on. Sentiments around the brand played the most crucial feature in predicting followed by performance, screen, battery and value for money.

The reasons for the differences in these aspects could be

i. **Brand**: iOS is a more premium brand than Android, and users who buy iOS devices are generally more likely to have higher expectations for the quality and features of the product.
ii. **Performance**: iOS devices are generally considered to be faster and more responsive than Android devices. This is due to a number of factors, including the use of a custom-designed processor and operating system.
iii. **Screen**: iOS devices typically have higher-quality screens than Android devices. This is due to the use of higher-resolution displays and better colour reproduction.
iv. **Battery life**: iOS devices typically have shorter battery life than Android devices. This is due to the use of lower capacity batteries.
v. **Value for money**: There is a wider range of Android devices available at different price points, so users can find an Android device that fits their budget.

These findings have a number of implications for smartphone manufacturers and marketers. For example, manufacturers can use sentiment analysis to identify the features that are most important to users and to prioritize the development of these features. Marketers can also use sentiment analysis to identify the aspects of their products that are most likely to be praised or criticized by users, and to tailor their marketing messages accordingly.

5.4. Implications for Design and Marketing: Insights for Smartphone Manufacturers

The findings of this research carry several implications for smartphone design and marketing. Manufacturers of Android smartphones may consider focusing on improving screen quality and battery life, as these aspects were identified as significant points of concern for users. By addressing these areas, manufacturers can enhance the overall user experience and reinforce the positive sentiments expressed in reviews.

On the other hand, manufacturers of iOS smartphones may want to prioritize improving performance in response to users' negative sentiments. By enhancing performance, Apple can not only meet the expectations of iOS users but also create a competitive advantage in the market.

Moreover, these research findings can be employed to develop more targeted marketing campaigns for smartphones. By understanding the primary concerns and preferences of users, manufacturers can tailor their marketing messages to address these specific areas, resulting in more effective and persuasive campaigns. For example, if users are dissatisfied with battery life, manufacturers can focus on promoting the long-lasting battery life of their products and highlight features that help users extend battery life.

5.5. Implications for Machine Learning: Fine-tuning for Better Predictions

The research findings have significant implications for the development of machine learning models used for brand prediction. Future studies might focus on improving the accuracy of these models by incorporating additional features such as smartphone price or user location. These additional variables could contribute to a more comprehensive understanding of the factors influencing brand prediction.

For example, if researchers find evidence suggesting that smartphone price significantly affects brand determination, manufacturers can incorporate this information into their machine learning models. By doing so, they can enhance the accuracy of their brand predictions and gain valuable insights for marketing and product development strategies.

Furthermore, the insights from sentiment analysis can extend beyond brand prediction to predicting customer churn. By monitoring the sentiment patterns associated with different aspects of smartphones, manufacturers can gain early signals of dissatisfaction and predict potential customer churn. This proactive approach empowers companies to address issues and concerns before they lead to customer attrition, thereby enhancing user retention and loyalty.

5.6. Implications for Customer Experience: Enhancing User Satisfaction

Ultimately, these research findings can be leveraged to improve the overall customer experience for smartphone users. Manufacturers can utilize the insights gained to identify areas where their products are falling short and make necessary improvements. Addressing the areas of concern highlighted in user reviews can lead to enhanced satisfaction and loyalty among customers.

For instance, if users express dissatisfaction with a particular smartphone manufacturer's customer service, the manufacturer can reflect on these reviews and make improvements to their customer service department accordingly. By doing so, they can demonstrate their commitment to providing excellent customer experiences and foster stronger relationships with their user base.

In conclusion, the findings of this research provide valuable insights into the differences in customer perceptions and sentiments of innovation between iOS and Android smartphones. The variations observed in text length, sentiments expressed, and brand predictions offer opportunities for smartphone manufacturers to optimize their products, marketing efforts, machine learning models, and customer experiences. By understanding these differences, manufacturers can better meet the needs and preferences of their users, leading to improved customer satisfaction and brand loyalty.

6. CONCLUSION

In the dynamic realm of smartphone technology, where iOS and Android platforms vie for user allegiance, understanding user sentiments, preferences, and predictive potential holds the key to innovation and success. This research embarked on a journey to explore the multifaceted landscape of customer reviews, unravelling insights that not only illuminate user sentiments but also inform predictive models and shape business strategies.

The journey began by dissecting the data landscape, where a meticulous methodology was employed to collect, pre-process, and analyse customer reviews for iOS and Android smartphones. Through this groundwork, the research unveiled stark differences in review lengths between the two platforms. Android reviews emerged as longer and more comprehensive, possibly reflecting Android users' inclination to provide detailed feedback on various aspects of their smartphones. In contrast, iOS reviews were succinct, indicating users' preference for concise expression.

The exploration then delved into the sentiments embedded within these reviews, showcasing a panorama of perceptions around key aspects of iOS and Android smartphones. The sentiment analysis unveiled nuanced differences in user sentiments, with Android users expressing positive sentiments towards aspects like price and features, while also highlighting aspects such as camera, battery life, screen, size, and performance. Conversely, iOS users predominantly conveyed negative sentiments towards battery life, followed by positive sentiments primarily directed at camera and performance, and to a lesser extent, display and price. These findings underscore the distinct priorities and preferences of users across the platforms, offering manufacturers valuable insights for targeted product improvements.

The research journey culminated in the realm of predictive modelling, where machine learning's prowess was harnessed to classify customer reviews as belonging to either iOS or Android smartphones. Through rigorous experimentation, the Random Forest Classifier emerged as the optimal model, achieving an accuracy of 67.42%. This predictive power highlights the feasibility of using sentiments as indicative markers to distinguish between the two platforms, facilitating refined marketing strategies and data-driven product enhancements.

6.1. Limitations

While this research sheds valuable light on the perceptions and sentiments surrounding iOS and Android smartphones, it is important to acknowledge certain limitations that warrant consideration in interpreting the findings and their implications. These limitations provide insights into the scope and direction for future research endeavours:

i. **Demographic Factors Exclusion**: This research focused exclusively on textual analysis of customer reviews, a limitation that excludes key demographic factors such as age, gender, and location. These variables are known to exert significant influence on user preferences and sentiments. Future studies could explore the intricate interplay between these demographic variables and user sentiments to unravel deeper insights into how different groups perceive and interact with smartphone platforms. Understanding how distinct demographics perceive various aspects could lead to more tailored products and marketing strategies that cater to diverse user preferences.

ii. **Text Length Context**: The investigation into text length disparities between Android and iOS reviews highlights intriguing trends, but it also leaves open questions about the underlying causes. While this research offers a glimpse into the potential reasons behind these variations, further exploration is warranted. Future work could employ qualitative research methods, such as surveys or interviews, to gain a deeper understanding of user motivations behind the length of their reviews. This could unveil whether the differences are driven by user preferences, technical issues, or other factors that could be leveraged for improved user experiences.

iii. **Platform Evolution**: The research is anchored in a specific time frame, and the technological landscape of smartphones is continuously evolving. The sentiments expressed in customer reviews are influenced by the current state of the platforms, including software updates, hardware advancements, and market trends. Consequently, the findings may not fully capture potential changes that occur in the platforms over time. A longitudinal study spanning multiple years could provide a more comprehensive understanding of how sentiments evolve alongside the platforms' development.

iv. **Sample Size Variability**: The dataset used for sentiment analysis was drawn from various sources and contained reviews for different smartphone models. While efforts were made to balance the dataset and ensure representation, variations in the number of reviews for each model and brand could introduce biases. Future studies could aim for a more standardized approach to data collection, potentially involving a broader range of models across different time periods to achieve a more balanced and representative dataset.

v. **Feature Engineering Complexity**: The sentiment-based feature engineering approach for predictive modelling is effective in highlighting the importance of sentiment aspects. However, the simplicity of the sentiment-weighted features used in this research may not encapsulate the full complexity of user opinions. Future research could explore more sophisticated methods of feature engineering that capture nuances in sentiment expressions, potentially utilizing more advanced natural language processing techniques to better capture the intricacies of user sentiments.

vi. **Generalizability**: While the findings of this research provide valuable insights into the sentiments and predictive modelling for iOS and Android smartphones, caution should be exercised when generalizing these findings to other smartphone platforms or contexts. Different platforms may have unique user dynamics, software ecosystems, and market positioning that could result in varying sentiment patterns. Future research could extend the analysis to other smartphone platforms to ascertain the generalizability of the identified sentiment patterns.

6.2. *Business Recommendations*

The insights garnered from this research harbour profound implications for smartphone manufacturers and marketers alike. Manufacturers of Android smartphones could leverage the findings by prioritizing improvements in screen quality and battery life. Addressing these key concerns can enhance user satisfaction and align products with user expectations. Conversely, iOS manufacturers can capitalize on their performance advantage by further enhancing this aspect to secure a competitive edge.

Marketers can tailor their strategies to resonate with user sentiments. For example, if Android users prioritize camera quality, manufacturers can highlight camera advancements in their marketing materials. Similarly, if iOS users emphasize battery life, marketing campaigns can underscore extended battery performance.

6.3. *Further Suggested Research*

The exploration into user sentiments, though comprehensive, unveils avenues for further research. Investigating the impact of user demographics on sentiments could provide deeper insights into the differing priorities of various user segments. Additionally, delving into the role of bugs and technical issues in shaping review lengths and sentiments could offer a more nuanced understanding of user experiences.

Further research could also extend the predictive modelling to incorporate additional variables such as smartphone price or user location, enhancing the accuracy of brand predictions. Moreover, exploring the use of sentiment analysis for predicting customer churn based on reviews could empower manufacturers to take proactive measures to retain customers and enhance loyalty.

In conclusion, this research voyage traversed the landscapes of customer sentiments, predictive modelling, and innovation insights in the world of iOS and Android smartphones. The findings resonate far beyond the digital realm, casting a transformative light on the convergence of technology and human experience. In an era where user preferences sculpt product trajectories, this research equips manufacturers and marketers with navigational tools to navigate the dynamic currents of user expectations, enabling them to create products that resonate with the heartbeats of their users.

REFERENCES

- Angelidis, S. and Lapata, M., 2018. Summarizing opinions: Aspect extraction meets sentiment prediction and they are both weakly supervised. *arXiv preprint arXiv:1808.08858*.
- Appel, O., Chiclana, F., Carter, J. and Fujita, H., 2016. A hybrid approach to the sentiment analysis problem at the sentence level. *Knowledge-Based Systems*, *108*, pp.110-124.
- Azzawi, M.A. and Ezeh, M.A., 2012. Students brand preferences between Apple and Samsung Smartphone. *Malardalen University Eskilstuna-Västerås, School of Sustainable Development of Society andTechnology*.
- Baccianella, S., Esuli, A. and Sebastiani, F., 2010, May. Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In *Lrec* (Vol. 10, No. 2010, pp. 2200-2204).
- Bafna, K. and Toshniwal, D., 2013. Feature based summarization of customers' reviews of online products. *Procedia Computer Science*, *22*, pp.142-151.
- Bali, A., Agarwal, P., Poddar, G., Harsole, D. and Zaman, N.M., 2016. Consumer's sentiment analysis of popular phone brands and operating system preference. *International Journal of Computer Applications*, *155*(4), pp.15-19.
- Bancken, W., Alfarone, D. and Davis, J., 2014. Automatically detecting and rating product aspects from textual customer reviews. In *Proceedings of the 1st international workshop on interactions between data mining and natural language processing at ECML/PKDD* (Vol. 1202, pp. 1-16). CEUR-WS. org.
- Barth, L., Kobourov, S.G. and Pupyrev, S., 2014. Experimental comparison of semantic word clouds. In *Experimental Algorithms: 13th International Symposium, SEA 2014, Copenhagen, Denmark, June 29–July 1, 2014. Proceedings 13* (pp. 247-258). Springer International Publishing.
- Behdenna, S., Barigou, F. and Belalem, G., 2018. Document Level Sentiment Analysis: A survey. EAI Endorsed Transactions on Context-Aware Systems and Applications, 4 (13), 154339.
- Bhandari, M. and Rodgers, S., 2020. What does the brand say? Effects of brand feedback to negative eWOM on brand trust and purchase intentions. In *Electronic Word of Mouth as a Promotional Technique* (pp. 125-141). Routledge.
- Bhatt, A., Patel, A., Chheda, H. and Gawande, K., 2015. Amazon review classification and sentiment analysis. *International Journal of Computer Science and Information Technologies*, *6*(6), pp.5107-5110.
- Bing, L., 2012. Sentiment analysis and opinion mining (synthesis lectures on human language technologies). *University of Illinois: Chicago, IL, USA*.
- Bradley, M.M. and Lang, P.J., 1999. *Affective norms for English words (ANEW): Instruction manual and affective ratings* (Vol. 30, No. 1, pp. 25-36). Technical report C-1, the center for research in psychophysiology, University of Florida.
- Castella, Q. and Sutton, C., 2014, April. Word storms: Multiples of word clouds for visual comparison of documents. In *Proceedings of the 23rd international conference on World wide web* (pp. 665-676).
- Chatterjee, P., 2001. Online reviews: do consumers use them?.
- Chow M M, Chen L H, Yeow J A *et al*. 2012, Factors Affecting the Demand of Smartphone Among Young Adults, *International Journal on Social Science Economics and Arts*, Vol. 2, No. 2, pp. 44-49.
- Cromar, S.A., 2010. Smartphones in the US: Market report. *Accessed*, *5*(7), p.2016.
- D'Aniello, G., Gaeta, M., Loia, F., Reformat, M. and Toti, D., 2018. An environment for collective perception based on fuzzy and semantic approaches. *Journal of Artificial Intelligence and Soft Computing Research*, *8*(3), pp.191-210.
- D'Arbelles, K., Berry, P. and Theyyil, A., 2020. Electronic word-of-mouth marketing on amazon: exploring how and to what extent Amazon reviews affect sales. *The McMaster Journal of Communication*, *12*(1), pp.50-79.

- Deng, S., Sinha, A.P. and Zhao, H., 2017. Adapting sentiment lexicons to domain-specific social media texts. *Decision Support Systems*, *94*, pp.65-76.
- Elli, M.S. and Wang, Y.F., 2016. Amazon reviews, business analytics with sentiment analysis. *Elwalda, Abdulaziz, et al."Perceived Derived Attributes of Online Customer Reviews*.
- Erkan, I. and Evans, C., 2018. Social media or shopping websites? The influence of eWOM on consumers' online purchase intentions. *Journal of Marketing Communications*, *24*(6), pp.617-632.
- France, S.L. and Ghose, S., 2016. An analysis and visualization methodology for identifying and testing market structure. *Marketing Science*, *35*(1), pp.182-197.
- Ghasemaghaei, M., Eslami, S.P., Deal, K. and Hassanein, K., 2018. Reviews' length and sentiment as correlates of online reviews' ratings. *Internet Research*, *28*(3), pp.544-563.
- Gurini, D.F., Gasparetti, F., Micarelli, A. and Sansonetti, G., 2013. A Sentiment-Based Approach to Twitter User Recommendation. *RSWeb@ RecSys*, *1066*, pp.1-4.
- Hakoama, M. and Hakoyama, S., 2011. The Impact of Cell Phone use on Social Networking and Development among college students. AABSS Journal, 15.
- Hemmatian, F. and Sohrabi, M.K., 2019. A survey on classification techniques for opinion mining and sentiment analysis. *Artificial intelligence review*, *52*(3), pp.1495-1545.
- Hu, M. and Liu, B., 2004, August. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 168-177).
- Hu, N., Liu, L. and Zhang, J.J., 2008. Do online reviews affect product sales? The role of reviewer characteristics and temporal effects. *Information Technology and management*, *9*, pp.201-214.
- Işıklar, G. and Büyüközkan, G., 2007. Using a multi-criteria decision making approach to evaluate mobile phone alternatives. *Computer Standards & Interfaces*, *29*(2), pp.265-274.
- Karjaluoto, H., Karvonen, J., Kesti, M., Koivumäki, T., Manninen, M., Pakola, J., Ristola, A. and Salo, J., 2005. Factors affecting consumer choice of mobile phones: Two studies from Finland. *Journal of Euromarketing*, *14*(3), pp.59-82.
- Keshavarz, H. and Abadeh, M.S., 2016, March. SubLex: Generating subjectivity lexicons using genetic algorithm for subjectivity classification of big social data. In *2016 1st Conference on Swarm Intelligence and Evolutionary Computation (CSIEC)* (pp. 136-141). IEEE.
- Keshavarz, H. and Abadeh, M.S., 2017. ALGA: Adaptive lexicon learning using genetic algorithm for sentiment analysis of microblogs. *Knowledge-Based Systems*, *122*, pp.1-16.
- Kilgarriff, A., 2000. Wordnet: An electronic lexical database.
- Klyuev, V. and Oleshchuk, V., 2011. Semantic retrieval: an approach to representing, searching and summarising text documents. *International Journal of Information Technology, Communications and Convergence*, *1*(2), pp.221-234.
- Lakshmi, V., Harika, K., Bavishya, H. and Harsha, C.S., 2017. Sentiment analysis of twitter data. *Int. Res. J. Eng. Technol*, *4*, pp.2224-2227.
- Li, M., Ch'ng, E., Chong, A.Y.L. and See, S., 2018. Multi-class Twitter sentiment classification with emojis. *Industrial Management & Data Systems*, *118*(9), pp.1804-1820.
- Lin, C.Y., Liaw, S.Y., Chen, C.C., Pai, M.Y. and Chen, Y.M., 2017. A computer-based approach for analyzing consumer demands in electronic word-of-mouth. *Electronic Markets*, *27*, pp.225-242.
- Ling, C., Hwang, W. and Salvendy, G., 2006. Diversified users' satisfaction with advanced mobile phone features. *Universal Access in the Information Society*, *5*, pp.239-249.
- Liu, B., 2012. Sentiment analysis: A fascinating problem. In *Sentiment Analysis and Opinion Mining* (pp. 1-8). Cham: Springer International Publishing.

- Louie A (2012), “Android Brand Preferences: US Consumers”, available at https://igr-inc.com/Advisory_And_Subscription_Services/Research_Catalog/ ?year=2012. Accessed on December 12, 2012.
- Luo, W., Zhuang, F., Zhao, W., He, Q. and Shi, Z., 2015. QPLSA: Utilizing quad-tuples for aspect identification and rating. *Information processing & management*, *51*(1), pp.25-41.
- Manek, A.S., Shenoy, P.D. and Mohan, M.C., 2017. Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier. *World wide web*, *20*, pp.135-154.
- Miley, F. and Read, A., 2011. Using word clouds to develop proactive learners. *Journal of the Scholarship of Teaching and Learning*, pp.91-110.
- Mokhlis, S. and Yaakop, A.Y., 2012. Consumer choice criteria in mobile phone selection: An investigation of malaysian university students. *International Review of Social Sciences and Humanities*, *2*(2), pp.203-212.
- Mowlaei, M.E., Abadeh, M.S. and Keshavarz, H., 2020. Aspect-based sentiment analysis using adaptive aspect-based lexicons. *Expert Systems with Applications*, *148*, p.113234.
- Nakov, P., Ritter, A., Rosenthal, S., Sebastiani, F. and Stoyanov, V., 2019. SemEval-2016 task 4: Sentiment analysis in Twitter. *arXiv preprint arXiv:1912.01973*.
- Neviarouskaya, A., Prendinger, H. and Ishizuka, M., 2007. Textual affect sensing for sociable and expressive online communication. In *Affective Computing and Intelligent Interaction: Second International Conference, ACII 2007 Lisbon, Portugal, September 12-14, 2007 Proceedings 2* (pp. 218-229). Springer Berlin Heidelberg.
- Nielsen (2011), “Nielsen Data Indicates Consumer Preference for Android”, The Nielsen Company, available at http://www.michaelsinsight.com/2011/04/nielsen-data-indicates consumer-preference-for-android.html. Accessed on September 15, 2012.
- Nielsen, F.Å., 2011. A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. *arXiv preprint arXiv:1103.2903*.
- Pant, A., Hsieh, Y.C., Lee, M. and Shen, C.H., 2014. Pointing out differences: Consumer evaluations of comparative information across product types. *Journal of Promotion Management*, *20*(3), pp.390-409.
- Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., AL-Smadi, M., Al-Ayyoub, M., Zhao, Y., Qin, B., De Clercq, O. and Hoste, V., 2016. Semeval-2016 task 5: Aspect based sentiment analysis. In *ProWorkshop on Semantic Evaluation (SemEval-2016)* (pp. 19-30). Association for Computational Linguistics.
- Poria, S., Cambria, E., Ku, L.W., Gui, C. and Gelbukh, A., 2014, August. A rule-based approach to aspect extraction from product reviews. In *Proceedings of the second workshop on natural language processing for social media (SocialNLP)* (pp. 28-37).
- Rain, C., 2013. Sentiment analysis in amazon reviews using probabilistic machine learning. *Swarthmore College*, *42*.
- Rana, T.A. and Cheah, Y.N., 2016. Aspect extraction in sentiment analysis: comparative analysis and survey. *Artificial Intelligence Review*, *46*, pp.459-483.
- Rana, T.A. and Cheah, Y.N., 2018. Improving aspect extraction using aspect frequency and semantic similarity-based approach for aspect-based sentiment analysis. In *Recent Advances in Information and Communication Technology 2017: Proceedings of the 13th International Conference on Computing and Information Technology (IC2IT)*(pp. 317-326). Springer International Publishing.
- Ravi, K. and Ravi, V., 2015. A survey on opinion mining and sentiment analysis: tasks, approaches and applications. *Knowledge-based systems*, *89*, pp.14-46.
- Rosenthal, S., Farra, N. and Nakov, P., 2019. SemEval-2017 task 4: Sentiment analysis in Twitter. *arXiv preprint arXiv:1912.00741*.
- Shah, P. and Swaminarayan, P., 2021. Lexicon-based sentiment analysis on movie review in the Gujarati language. *International Journal of Information Technology, Communications and Convergence*, *4*(1), pp.63-72.

- Shaikh, T. and Deshpande, D., 2016. Feature selection methods in sentiment analysis and sentiment classification of amazon product reviews. *International Journal of Computer Trends and Technology (IJCTT)*, *36*(4), pp.225-230.
- Sharma, R., Morales-Arroyo, M. and Pandey, T., 2012. The emergence of electronic word-of-mouth as a marketing channel for the digital marketplace. *Ravi S. Sharma, Miguel Morales-Arroyo, and Tushar Pandey."The Emergence of Electronic Word-of-Mouth as a Marketing Channel for the Digital Marketplace", Journal of Information, Information Technology, and Organizations*, *6*, pp.41-61.
- Shirsat, V.S., Jagdale, R.S. and Deshmukh, S.N., 2017, August. Document level sentiment analysis from news articles. In *2017 international conference on computing, Communication, Control and Automation (ICCUBEA)* (pp. 1-4). IEEE.
- Singh, J. and Goyal, B.B., 2009. Mobile handset buying behavior of different age and gender groups. *International Journal of Business and Management*, *4*(5), pp.179-187.
- Singh, J.P., Irani, S., Rana, N.P., Dwivedi, Y.K., Saumya, S. and Roy, P.K., 2017. Predicting the "helpfulness" of online consumer reviews. *Journal of Business Research*, *70*, pp.346-355.
- Singh, V.K., Piryani, R., Uddin, A. and Waila, P., 2013, March. Sentiment analysis of movie reviews: A new feature-based heuristic for aspect-level sentiment classification. In *2013 International mutli-conference on automation, computing, communication, control and compressed sensing (imac4s)*(pp. 712-717). IEEE.
- Tomasiello, S., 2020. Least-squares fuzzy transforms and autoencoders: some remarks and application. *IEEE Transactions on Fuzzy Systems*, *29*(1), pp.129-136.
- Tsai, C.Y. and Shih, C., 2011. An analysis of preferences in smartphone interface design. *Scientific Research and Essays*, *6*(20), pp.4195-4206.
- Wang, G., Zhang, Z., Sun, J., Yang, S. and Larson, C.A., 2015. POS-RS: A Random Subspace method for sentiment classification based on part-of-speech analysis. *Information Processing & Management*, *51*(4), pp.458-479.
- Wiebe, J., Wilson, T. and Cardie, C., 2005. Annotating expressions of opinions and emotions in language. *Language resources and evaluation*, *39*, pp.165-210.
- Xu, L., Liu, K., Lai, S., Chen, Y. and Zhao, J., 2013, May. Walk and learn: a two-stage approach for opinion words and opinion targets co-extraction. In *Proceedings of the 22nd International Conference on World Wide Web* (pp. 95-96).
- Xu, Y., Wu, X. and Wang, Q., 2015, December. Sentiment analysis of yelp's ratings based on text reviews. In *2015 17th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)* (Vol. 17, No. 1, pp. 117-120).
- Yadollahi, A., Shahraki, A.G. and Zaiane, O.R., 2017. Current state of text sentiment analysis from opinion to emotion mining. *ACM Computing Surveys (CSUR)*, *50*(2), pp.1-33.
- Yan, Z., Xing, M., Zhang, D. and Ma, B., 2015. EXPRS: An extended pagerank method for product feature extraction from online consumer reviews. *Information & Management*, *52*(7), pp.850-858.
- Yang, H. and Li, K., 2022. PyABSA: open framework for aspect-based sentiment analysis. *arXiv preprint arXiv:2208.01368*.
- Yu, J., Zha, Z.J., Wang, M. and Chua, T.S., 2011, June. Aspect ranking: identifying important product aspects from online consumer reviews. In *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies* (pp. 1496-1505).
- Zhang, L. and Liu, B., 2014. Aspect and entity extraction for opinion mining. In *Data mining and knowledge discovery for big data: Methodologies, challenge and opportunities* (pp. 1-40). Berlin, Heidelberg: Springer Berlin Heidelberg.

APPENDIX

Figures

Figure 1: Confusion matrix of Random Forest Classifier

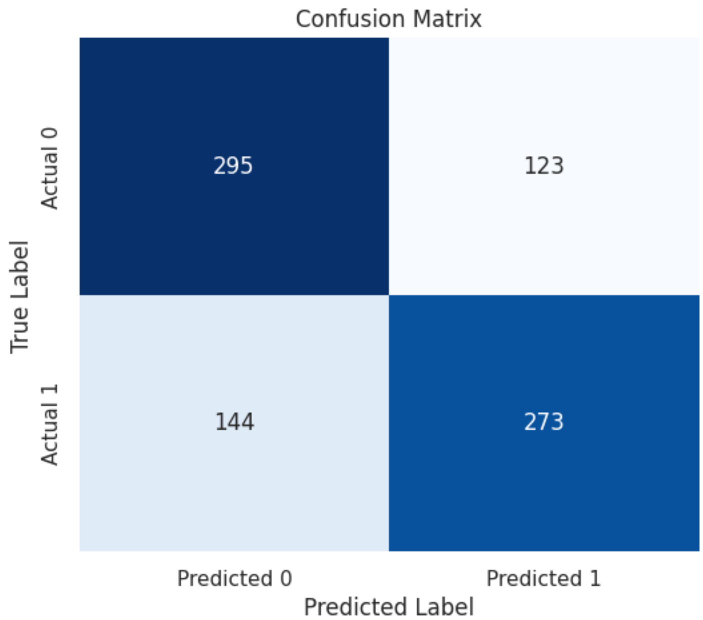
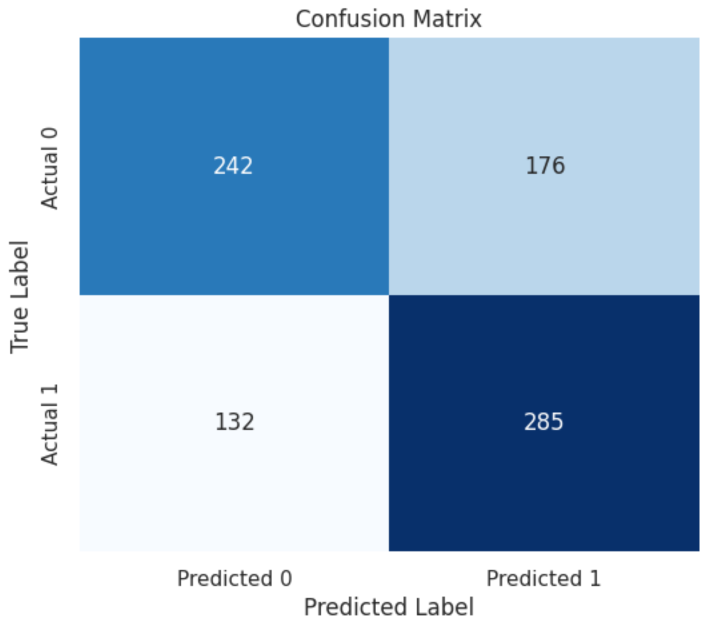


Figure 2: Confusion matrix of Naïve Bayes Classifier



Tables

Table 1: Python notebook details