THE UNIVERSITY
of ADELAIDE

Cyber Security Industry Project (Part B) Final Report

Project Title: Mobile Apps Cybersecurity and Privacy Analysis Through User Feedback Reviews

Yuen Yuen (a1836569)

Project Coordinator: Dr. Faheem Ullah

25 April 2023

I. Abstract

This final report outlines the methodology and findings of a project that utilised natural language processing to analyse nearly 20 million user reviews of the top 780 Google apps across 45 categories from six English-speaking countries. The project aimed to address three research questions related to user requirements and cybersecurity/privacy issues. The study identified over 2.7 million user requirements and 241 patterns in the reviews that shed light on how users interact with the apps. The analysis also revealed the top 15 mobile apps with cybersecurity and privacy issues, highlighting the need for developers to prioritise security measures. The sentiment analysis of the reviews showed that 49% of the reviews were positive, while 32% were negative. The use of NLP to analyse mobile app reviews has provided valuable insights into cybersecurity and privacy issues, demonstrating the importance of NLP in aiding app developers in making necessary improvements and identifying fraudulent apps that can harm users' personal data. Overall, this project highlights the significance of NLP technology in ensuring the security and privacy of mobile apps as their use continues to expand.

Keywords: *cybersecurity; privacy; machine learning; natural language processing; NLP; mobile app reviews; unstructured text extraction; sentiment analysis; pattern recognition; software requirement engineering*

II. Introduction

Cybersecurity and privacy have become crucial concerns in the era of digital technology, where millions of users are sharing sensitive information online. Google, being one of the largest technology platforms, provides a range of applications to users, including Productivity, Entertainment, Communication and many other categories. These mobile applications (apps) are widely used by billions of users worldwide, and they collect and store sensitive information about users. Hence, it is essential to ensure the security and privacy of these applications to protect user data from potential cyber attacks and privacy breaches.

User reviews of these applications can provide valuable insights into the cybersecurity and privacy related concerns that users have with these applications. However, manually analysing millions of user reviews is a challenging and time-consuming task. Therefore, natural language processing (NLP) techniques can be utilised to extract cybersecurity and privacy related concerns from these reviews automatically.

This research project aims to utilise NLP techniques to extract cybersecurity and privacy related concerns from Google apps user reviews. The project aims to answer the question: what are the most common cybersecurity and privacy related concerns among Google app users? The research project will focus on three key NLP techniques: sentiment analysis, pattern recognition, and software requirement engineering (SRE) extraction.

Sentiment analysis is an NLP technique that can identify the overall sentiment of the review, whether it is positive, negative, or neutral. By applying sentiment analysis to user reviews, users' general perception of the application's cybersecurity and privacy features can be determined.

Pattern recognition is another NLP technique that can identify patterns in user reviews. In the context of cybersecurity and privacy, pattern recognition can detect recurring complaints about specific security issues or privacy concerns. These recurring complaints can provide valuable insights into the most pressing cybersecurity and privacy issues for users.

Software requirement engineering is an NLP technique that can summarise user requirements from the context of user reviews. By applying this technique to user reviews, the most critical cybersecurity and privacy related features that users expect from apps can be identified.

The objective of this research project is to utilise NLP techniques to identify cybersecurity and privacy related issues from user reviews of Google apps. By applying these techniques, the findings can be applied to the industry to create more secure and privacy conscious applications that cater to the needs of users. This study will provide valuable insights to developers to improve the quality of their apps and enhance user satisfaction by addressing their concerns related to cybersecurity and privacy.

III. Objectives

The main objective of this research project is to use NLP techniques to extract cybersecurity and privacy related concerns from user reviews of Google apps on the Google Play Store. The aim is to identify the most prevalent cybersecurity and privacy related issues among Google app users using three NLP techniques: sentiment analysis, pattern recognition, and software requirement engineering.

The objective is achieved by collecting user reviews of Google apps from the Google Play Store using web scraping techniques. A sufficient number of user reviews are gathered, and the data is preprocessed. The NLP techniques are then employed to extract cybersecurity and privacy related concerns from the reviews.

Three research questions going to be answered in this project are:

RQ1: Sentiment Analysis

The first question of the project is to use sentiment analysis to determine the overall sentiment of the user reviews. Through the analysis of the overall sentiment of the reviews, the identification of whether users generally have positive or negative feedback on the security and privacy features of Google apps can be done.

RQ2: Pattern recognition

The second question is to use pattern recognition to identify recurring complaints or patterns in the user reviews that relate to specific security issues or privacy concerns. By identifying these patterns, the most significant cybersecurity and privacy related problems that Google app users face can be determined.

RQ3: Software requirement extraction

The third question is to employ software requirement engineering to summarise the user requirements related to cybersecurity and privacy concerns from the user reviews. By applying this technique, The most critical cybersecurity and privacy related user requirements that users expect from Google apps can be identified through our analysis.

By answering these questions, this research project aims to provide valuable insights into the most pressing cybersecurity and privacy related concerns among Google app users on the Google Play Store. The software requirement engineering (SRE) process for developing more secure and privacy respecting mobile apps can be improved based on the insights gained from this research. This research is particularly relevant given the increasing importance of cybersecurity and privacy in today's digital world. The results of this project will help to identify the most pressing cybersecurity and privacy concerns effectively, and inform the development of applications that better meet user requirements.

IV. Challenges

The challenges for this project are:

- Obtaining a large enough sample of user reviews to achieve statistically significant results may be challenging.
- Filtering out irrelevant reviews that do not pertain to cybersecurity and privacy may be time-consuming.
- Sentiment analysis may not be accurate for certain types of reviews, such as those with mixed sentiments or sarcasm.
- Different users may use different terminology when discussing cybersecurity and privacy issues, which may make it difficult to identify patterns in the reviews.
- Privacy concerns regarding the collection and use of user reviews must be addressed to ensure ethical research practices.
- The availability and quality of NLP tools and libraries may vary for different languages, which may affect the accuracy of the analysis.
- Some user reviews may contain language that is difficult to process using NLP techniques, such as non-standard grammar or misspellings.
- Identifying the most critical cybersecurity and privacy related issues may be challenging since users may have different priorities and expectations.
- Analysing reviews from different Google apps may require different techniques and approaches, which may add complexity to the analysis.
- Keeping up with changes to the Google Play Store and Google apps may require constant updates to the scraping and analysis techniques.
- Dealing with potential biases in the user reviews, such as the demographics of the users or the

type of device they are using, may be challenging.

- Identifying which issues are most important to users may require additional surveys or studies to validate the findings.

- Dealing with the potential for fake or spam reviews may affect the accuracy of the analysis.

- Ensuring the privacy and anonymity of the users whose reviews are being analysed may be challenging, especially if identifiable information is included in the reviews.

- Ensuring the reliability and accuracy of the web scraping techniques used to collect the user reviews may be challenging, as Google may update its security measures to prevent scraping.

- The process of analysing a large amount of user reviews requires high performance computing power to handle the massive amount of data. The computational power is required to run NLP algorithms that can identify patterns, detect sentiments, and summarise user requirements. Without the proper computational resources, the analysis may take an unreasonably long time or may not be possible at all.

- Another challenge is to optimise the algorithms and processing methods to reduce computational requirements while still ensuring accuracy.

V. Literature Review

I. Natural Language Processing (NLP)

Natural Language Processing is a rapidly growing field that focuses on teaching computers to understand human language. Hirschberg and Manning (2015) reviewed the advancements made in NLP research and highlighted the evolution of computational approaches towards automating the analysis, understanding, and production of human language content. Cambria and White (2014) discussed recent developments in NLP that have led to the evolution of natural language understanding. They noted that the analysis of text involves a deep understanding of natural language, which is essential for building intelligent systems that can interpret human language accurately. Khurana et al. (2023) provided an overview of the state of the art, current trends, and challenges in NLP. They discussed the various applications of NLP, such as sentiment analysis and customer reviews, and highlighted the importance of natural language understanding for accurate data analysis.

Several studies have explored gender bias in NLP systems, and Sun et al. (2019) reviewed contemporary studies on recognizing and mitigating gender bias in NLP. They analysed four forms of representation bias and discussed methods to mitigate gender bias in NLP systems. Similarly, Kreimeyer et al. (2017) conducted a systematic review of NLP systems for capturing and standardising unstructured clinical information. They highlighted the need for automated solutions utilising NLP to analyse unstructured clinical text and discussed various NLP techniques used for this purpose.

II. Machine learning (ML) and deep learning (DL)

Machine learning and deep learning have

revolutionised the field of NLP by enabling computers to understand human language and generate meaningful responses. NLP has seen tremendous advancements with the application of machine learning and deep learning techniques, which have led to the development of sophisticated models for various NLP tasks such as sentiment analysis, language translation, and speech recognition.

One of the key advancements in NLP is the application of machine learning algorithms for text classification, where documents are classified into different categories based on their content. A study by L. Aristodemou and F. Tietze (2018) provided a comprehensive literature review on machine learning and deep learning methods for analysing intellectual property analytics, which included the use of deep neural network-based corporate performance models and NLP-based hierarchical classification models.

Another recent review by A. Le Glaz et al. (2021) explored the use of machine learning and NLP models for mental health analysis. The study highlighted the potential of machine learning and NLP models in predicting mental health outcomes from social media data. Similarly, a study by O. Sen et al. (2021) provided a comprehensive review of classical, machine learning, and deep learning-based methods for Bangla NLP, which included the use of Convolutional Neural Networks (CNN) for speech processing and recognition.

The application of deep learning techniques in NLP has also seen significant progress. A study by T. Kaluarachchi et al. (2021) reviewed recent deep learning approaches in human-centred machine learning, which included the use of NLP powered user experiences and wizard-of-oz type NLP. Furthermore, A. Yadav and D.K. Vishwakarma (2020) provided a review on sentiment analysis using deep learning architectures, which included the use of Deep Neural Networks (DNNs) for text classification.

III. Methodology for review analysis through NLP

NLP has revolutionised the way data is analysed, making it easier and more efficient for researchers to analyse large volumes of textual data. One area where NLP has shown great potential is in the analysis of literature and reviews. This literature review examines the methodology used for review analysis through NLP, with a focus on recent studies in the field.

Several recent studies have highlighted the potential of NLP in analysing reviews across a variety of fields, including management research, healthcare, and product reviews. One study by Fu et al. (2020) provides a detailed methodology review for clinical concept extraction, while another study by Kang et al. (2020) offers a literature review of NLP in management research. Both studies highlight the importance of employing appropriate toolkits and procedural steps when applying NLP as an analytical technique.

In addition to these studies, several others have focused on specific applications of NLP in review analysis, such as sentiment analysis and opinion mining. Shivaprasad and Shetty (2017) provide a review of sentiment analysis techniques for product reviews, while Sun et al. (2017) present a review of NLP techniques for opinion mining systems.

IV. Mobile app reviews

Mobile app reviews analysis through NLP has gained much attention in recent years due to its potential in extracting valuable insights from user reviews. The application of NLP in analysing mobile app reviews is being explored in various studies. For instance, Tavakoli et al. (2018) discussed the use of NLP, text analysis, and sentiment analysis techniques to extract useful software development information from mobile application reviews. Meanwhile, Rehman et al. (2021) presented a study on the use of sentiment analysis and app ratings to select apps using NLP, which showed the potential of speech-based NLP in analysing user reviews of mobile apps. On the other hand, Srisopha et al. (2020) discussed the challenges of using NLP-based analysis in identifying country-specific feature requests in mobile app reviews. Overall, these studies highlight the importance and potential of NLP-based analysis in extracting valuable insights from mobile app reviews.

V. Reviews collection

In recent years, mobile applications have gained a significant place in our lives. They offer a wide range of services and have become an essential tool for individuals. One of the essential factors that affect the success of mobile applications is user reviews. User reviews are a rich source of information that provides developers with feedback on their applications' performance and user experience. As such, user reviews have gained attention in NLP research. Several studies have been conducted to collect and analyse user reviews to improve applications' performance.

Sarker et al. (2021) highlighted the importance of NLP in collecting and analysing user reviews for mobile applications. Similarly, Genc-Nayebi and Abran (2017) conducted a systematic literature review on opinion mining studies from mobile app store user reviews. They developed a set of research questions to guide the literature review process and performed an extensive search to find publications that answer the research questions.

VI. Review sampling

Mobile app review sampling has gained popularity in recent times due to the vast number of reviews available on app stores. This literature review focuses on the use of NLP techniques for analysing mobile app reviews. Various studies have utilised NLP techniques to extract and classify user issues raised in mobile app reviews. The majority of datasets have been created through random sampling of available apps on app stores. However, building an automatic identification of country-specific feature requests in mobile app reviews through NLP-based techniques has been explored as well. While NLP techniques have been effective in identifying user issues, challenges like identifying sarcasm and contextual differences between languages remain. Therefore, NLP techniques should be complemented with manual annotation or domain knowledge to improve the accuracy of the analysis. The literature review cites references including Genc-Nayebi and Abran (2017), Srisopha et al. (2020), and Chen et al. (2018) among others.

VII. User feedback analysis

The analysis of user feedback is essential for understanding user preferences and satisfaction. NLP techniques have been adopted to automatically analyse feedback data for better decision making. This literature review aims to summarise the recent trends and challenges in applying NLP techniques to user feedback analysis in education. Shaik et al. (2022) review existing NLP methodologies for feedback analysis and propose a synthesis of these methods to improve student user feedback analysis. Santos et al. (2019) provide an overview of the classification approaches used in user feedback analysis and identify the dominant role of machine learning algorithms. Kastrati et al. (2021) review the application of NLP and deep learning techniques for sentiment analysis in students' feedback.

Overall, NLP and machine learning techniques have shown promising results in analysing user feedback in various domains, including cybersecurity, privacy and software engineering. However, challenges such as the lack of labelled data, domain-specific language, and biases in data must be addressed to improve the accuracy and effectiveness of NLP techniques in user feedback analysis.

VIII. Unstructured data analysis

Unstructured data analysis has become an essential aspect of data analysis in the big data era, providing significant insights into patterns and relationships that are not possible with structured data alone. The article by Eberendu (2016) provides an overview of the significance of unstructured data in predictive analysis, outlining the relevance of this data in decision making and the opportunities to generate relevant information faster through unstructured data analysis. The study by Scherm et al. (2014) highlights the importance of meta-analysis in synthesising structured and unstructured data for plant pathology. The authors emphasise the importance of data-mining approaches to integrate unstructured data into structured data. Balducci and Marinova (2018) explore the application of unstructured data analysis in marketing. The authors provide a comprehensive review of unstructured data sources and the challenges of preparing the data for analysis.

IX. Cybersecurity and privacy for mobile apps

Mobile applications have become an essential tool in everyday life, leading to an increase in concerns about cybersecurity and privacy. Game theory has been applied to cybersecurity and privacy issues in mobile applications, categorising their applications into two classes: security and privacy. Do et al. (2017) demonstrated how game theory is utilised in cyberspace security and privacy. The privacy implications of cyber security systems have also been analysed, and a taxonomy has been provided for a representative set of technologies (Toch et al., 2018). Blockchain technology has been explored as a possible solution for cybersecurity and privacy challenges in various sectors, including mobile applications (Maleh et al., 2020). These studies provide insight into the current state of privacy and cybersecurity technologies and potential solutions to enhance mobile application security and privacy.

X. Data protection

Data protection is a critical issue that has been gaining significant attention due to the proliferation of personal data online and the increased risk of cyber threats. The European Union (EU) legal framework has been exploring challenges and opportunities in this area, including personal data protection and cyber security. Andraško, Mesarčík, and Hamuľák (2021) provide an in -depth review of this topic, highlighting the need for more stringent regulations and guidelines to ensure the safety of personal data. Mulligan, Freeman, and Liu (2019) also emphasise the importance of data protection laws, privacy, and cyber security, as they are intertwined and necessary to safeguard sensitive information. In addition, the authors point out that existing legislation may need to be revised to meet the growing demands of modern technology. Finally, Wylde et al. (2022) examine the relationship between blockchain, data privacy, and cyber security, highlighting the potential of blockchain technology to enhance data protection. The authors identify key challenges and opportunities that need to be addressed to ensure effective implementation of blockchain in data protection, including legal and regulatory challenges.

XI. Personal information and data security

Personal information and data security are critical concerns in the era of digitalization, where data is generated and collected on a massive scale. The unauthorised access to personal information, such as social security numbers and financial accounts, can lead to identity theft and fraud (Thakur et al., 2015). Abouelmehdi et al. (2017) note that organisations need to safeguard personal information and address their legal responsibilities concerning data protection. However, Conger et al. (2013) point out that there is a tension between individuals' privacy and security concerns and government's access to personal data for collective good. This tension is further exacerbated by the emergence of new technologies, such as mobile apps for depression, which collect and use personal information (O'Loughlin et al., 2019). Nevertheless, Achar et al. (2022) emphasise that data breaches of personal information are likely to be the focus of criminal organisations, violating individuals' privacy and causing significant harm.

XII. Cybersecurity and privacy keywords

The use of NLP for cybersecurity and privacy has gained significant attention in recent years. Sousa and Kern (2022) conducted a systematic review of deep learning methods for privacy-preserving NLP, where they discussed privacy issues related to NLP and open challenges in privacy-preserving NLP. Rahman et al. (2020) conducted a literature review on mining cyber threat intelligence from unstructured texts, where they listed NLP techniques related to topic classification and keyword extraction as effective tools for aiding cybersecurity. Sarker et al. (2021) provided an overview of AI-driven cybersecurity and discussed the use of semantic analysis and keyword extraction for intelligent cybersecurity.

XIII. Sentiment analysis

Sentiment analysis is a field of NLP that aims to classify the polarity of text as positive, negative, or neutral. In recent years, sentiment analysis has gained significant attention in various domains, including healthcare, product reviews, and student feedback. Abualigah et al. (2020) provide a brief review of sentiment analysis in healthcare, highlighting the use of NLP, text analysis, and computational linguistics to identify and classify sentiments. Shivaprasad and Shetty (2017) review the application of sentiment analysis in product reviews, emphasising the importance of NLP algorithms in the analysis process. Khan et al. (2016) focus on sentiment analysis techniques and the need to address open challenges specific to NLP. Kastrati et al. (2021) present a systematic mapping study of sentiment analysis in students' feedback, emphasising the application of NLP, deep learning, and machine learning solutions. The study by Mathew and Bindu (2020) reviews NLP techniques for sentiment analysis using pre-trained models.

XIV. Tokenization and lemmatization

Tokenization and lemmatization are two important techniques used in NLP to preprocess text data. Tokenization refers to the process of splitting a text document into smaller units of text, usually words or phrases, while lemmatization refers to the process of reducing words to their base or dictionary form. These techniques are essential for several NLP tasks, such as sentiment analysis, machine translation, and named entity recognition.

Various studies have explored the application of tokenization and lemmatization in different contexts. In radiology, for example, the use of NLP techniques, including tokenization and lemmatization, has been investigated for the evaluation of radiology reports (Kaggal et al., 2020). Additionally, the use of NLP techniques for morphological analysis, including tokenization and lemmatization, has been reviewed in the context of Arabic language processing (Elaziz et al., 2019).

In the medical education domain, NLP techniques, including tokenization and lemmatization, have been explored for various applications, such as question-answering systems and medical text summarization (Karami et al., 2019). In the same way, a systematic review of named entity recognition techniques, which involve tokenization and lemmatization, has been conducted to identify the best practices in this area (Sharma, Chakraborty & Kumar, 2022).

XV. NLP models

NLP models have revolutionised NLP tasks such as language translation, sentiment analysis, and named entity recognition, achieving state-of-the-art performance in many benchmark datasets (Devlin et al., 2018; Brown et al., 2020). However, training these models is computationally expensive and requires a significant amount of resources, such as high-performance computing clusters and large datasets (Sharir et al., 2020). Moreover, deploying large NLP models in resource-constrained environments is challenging due to their high memory requirements (Qiu et al., 2020). To address these issues, researchers have developed efficient NLP models, such as lightweight convolutional neural networks (CNN) and recurrent neural networks (RNN), which have shown promising results in various NLP tasks (Vaswani et al., 2017; Lample & Conneau, 2019). Additionally, pre-trained language models, such as BERT (Devlin et al., 2018) and GPT (Brown et al., 2020), have become a popular approach to reduce the cost of training NLP models while improving their performance. These models are pre-trained on massive text corpora and fine-tuned for specific NLP tasks, allowing them to achieve state-of-the-art performance with less training data (Qiu et al., 2020).

XVI. Pattern recognition

Pattern recognition is a fundamental technique in NLP used for analysing language data and identifying meaningful patterns. Kocaleva et al. (2016) describe pattern recognition as a method that models regularities found in data. Supervised learning is a common approach in NLP, which uses training data to identify matching patterns. Gu et al. (2018) discuss the recent advancements in Convolutional Neural Networks (CNN) for visual recognition, speech recognition, and NLP. They emphasise the use of CNN in computer vision, speech, and NLP. Lavanya and Sasikala (2021) explain that pattern recognition models are useful in social healthcare for uncovering masked patterns and finding meaningful insights. In addition, they discuss the integration of NLP techniques with pattern recognition models. Paolanti and Frontoni (2020) review multidisciplinary pattern recognition applications and emphasise the incorporation of NLP output in addition to traditional approaches.

XVII. Text classification

Text classification is a crucial task in NLP, which involves categorising text data into predefined classes or categories such as cybersecurity and privacy. The use of deep learning techniques in text classification has become increasingly popular due to their high performance in various NLP applications. Lavanya and Sasikala (2021) conducted a comprehensive survey on the application of deep learning techniques in text classification within social healthcare networks. Bhavani and Kumar (2021) provided a review of the state-of-the-art text classification algorithms, including machine learning-based and deep learning-based models. Additionally, Dreisbach et al. (2019) conducted a systematic review of NLP and text mining techniques for symptom information extraction from electronic patient-authored text data. Finally, Dogra et al. (2022) presented a detailed review of the complete process of text classification using state-of-the-art NLP models.

XVIII. Cybersecurity threat detection

The field of cybersecurity has been increasingly adopting NLP techniques to detect and prevent various threats. Several studies have conducted systematic reviews and surveys to evaluate the application of NLP in cybersecurity. Sworna et al. (2022) reviewed NLP methods in host-based intrusion detection systems and highlighted the need for future research in this area. Sarker et al. (2021) provided an overview of AI-driven cybersecurity and discussed the potential of NLP for intelligent cybersecurity modelling. Salloum et al. (2022) conducted a literature survey on phishing email detection using NLP techniques and emphasised the need for effective detection methods to prevent monetary and identity loss for the user. Mathews (2019) evaluated the application of explainable artificial intelligence, including NLP, in malware classification in the context of cybersecurity.

VI. Methodology

As businesses are constantly searching for ways to improve their products and services, understanding user feedback becomes essential. Conducting a project to identify the sentiment of the reviews, detect patterns, and summarise user requirements can help businesses to address their user needs.

Sentiment analysis is a process of identifying the overall sentiment of the user reviews. It can help businesses to understand whether the feedback is positive, negative or neutral. For example, it can be used to identify customer satisfaction levels, detect brand reputation or monitor customer service. Sentiment analysis involves NLP techniques such as text preprocessing, tokenization, and machine learning algorithms. Several studies have shown that sentiment analysis can improve customer satisfaction, sales, and user engagement. (Reyes, Saura & Alvarez, 2018; Jabangwe, Edison & Duc,, 2018)

Another important aspect of user feedback is pattern recognition. This process involves detecting recurring patterns or issues that users frequently mention. Pattern recognition can help businesses to understand the areas of their products or services that need improvement. For example, if users frequently complain about cybersecurity issues, businesses can prioritise improving their cybersecurity measures. NLP techniques such as part-of-speech tagging, named entity recognition, and clustering can be used to identify patterns. Studies have shown that pattern recognition can help businesses to improve user satisfaction, product quality, and customer loyalty. (Mansour and Abbasi, 2017; Malavolta et al., 2015)

Software requirement engineering is the process of gathering user requirements and translating them into software specifications. User feedback is a valuable source of information that can be used to identify user requirements. However, analysing user feedback can be challenging due to the large amount of data and the diversity of user feedback. NLP techniques such as topic modelling, sentiment analysis, and text classification can help to identify user requirements from user feedback. Studies have shown that software requirement engineering can improve product quality, user satisfaction, and software development time. (Zhang and Zhang, 2019; Chowdhury et al., 2018)

A practical industry project can be

conducted to address these three aspects of user feedback. The project can involve collecting user feedback from various sources such as mobile apps, social media, or customer service. The collected data can be preprocessed and analysed using NLP techniques to identify the sentiment of the reviews, detect patterns, and summarise user requirements. The findings can be presented to the business stakeholders to help them make data-driven decisions that improve their products or services.

In conclusion, conducting a project to identify the sentiment of the reviews, detect patterns, and summarise user requirements is essential for businesses to understand their user needs. Sentiment analysis, pattern recognition, and software requirement engineering are all NLP techniques that can be used to analyse user feedback. A practical industry project can be conducted to implement these techniques and present the findings to business stakeholders.

General review analysis steps

1. Data collection: Collect user feedback data from mobile app stores or other sources such as social media platforms or online forums. The data should be in a structured format like a CSV file, and it should contain features such as review ID, date, rating, and content.

2. Data pre-processing: This involves data cleaning and preparation before the analysis. The process includes removing irrelevant data like duplicates or incomplete reviews, converting text to lowercase, and removing stop words and punctuation. The pre-processing also includes tokenization, stemming, and lemmatization of the text to normalise the words and reduce the dimensionality of the dataset.

3. Sentiment analysis: Perform sentiment analysis on the pre-processed text data. There are various techniques to perform sentiment analysis, including rule-based methods, machine learning-based methods, and deep learning-based methods. In this case, you can use pre-trained models like VADER (Valence Aware Dictionary and sEntiment Reasoner) or TextBlob, or you can train your own model based on the data. The output of the sentiment analysis should be a score that indicates the polarity of the review, which can be positive, negative, or neutral.

4. Labelling: Assign a sentiment label to each review based on the polarity score obtained from the sentiment analysis. Reviews with a score greater than 0.5 can be labelled as positive, reviews with a score less than -0.5 can be labelled as negative, and reviews with a score between -0.5 and 0.5 can be labelled as neutral.

5. Visualisation: Visualise the results using graphs or charts to provide insights into the overall sentiment of the reviews. This can help to identify trends and patterns in the data and to detect the most common issues or areas of concern.

6. Evaluation: Evaluate the accuracy of the sentiment analysis model by comparing the labelled sentiment with the actual sentiment of the reviews. This can be done using metrics like precision, recall, and F1 score.

7. Iteration: Iterate and improve the model based on the evaluation results. This involves adjusting the parameters of the model, changing the pre-processing steps, or using a different model altogether.

8. Integration: Integrate the sentiment analysis model into the mobile app or

other product to automatically analyse and label user feedback. This can help to identify issues or areas of concern quickly and to take action to address them.

(Liu 2012; Hutto & Gilbert 2014; Zhang, Zhao & LeCun 2015)

Review collecting

To collect user reviews from Google Play Store, a Python program was developed to automate the process. The program collected reviews from six English speaking countries, including the UK, USA, Australia, New Zealand, Canada, and India. The reviews were collected from the most downloaded apps, totaling around 800 apps, with nearly 20 million reviews collected in total.

The program was developed to navigate to the review section of each app and scrape the review text, rating, and other relevant data, such as the date and time of the review. The program also automatically stored the collected reviews in CSV format. Once the reviews were collected, they were imported into a Microsoft SQL database for further analysis. The SQL database was chosen for its ability to handle large volumes of data and its compatibility with the tools used for data analysis.

Overall, the automated program was able to efficiently collect a large volume of reviews from multiple countries, providing a diverse and comprehensive dataset for analysis. The use of a SQL database allowed for efficient storage and retrieval of the collected reviews, making it easy to analyse and draw insights from the data.

Review pre-processing

To conduct NLP analysis on reviews of mobile apps, pre-processing steps need to be taken to clean and transform the text data. The pre-processing steps for conducting NLP analysis on reviews of mobile apps were conducted by a self-developed Python program that utilised the Natural Language Toolkit (NLTK) module (https://www.nltk.org). The following steps were followed:

I. Tokenization: The first step in pre-processing is to tokenize the text, which involves breaking down the text into individual words or tokens. This allows the text to be analysed on a more granular level.

II. Lowercasing: All text is converted to lowercase to ensure that the same word written in different cases is treated as one word.

III. Stop words removal: Stop words such as "the", "and", "a", and "in" are removed from the text since they do not carry significant meaning in the analysis.

IV. Stemming and Lemmatization: This process reduces inflected words to their base or root form. This is done to normalise the text, as different forms of the same word can affect the analysis.

V. Removal of punctuation and special characters: All punctuation marks and special characters like "@", "#", etc. are removed from the text as they do not add value to the analysis.

VI. Removing numbers: Numbers are removed from the text, as they usually do not provide any meaningful information.

VII. After pre-processing, the text data is ready to be analysed using various NLP techniques. These pre-processing steps help to ensure that the text data is clean and ready for analysis, enabling more accurate and useful insights to be generated.

Sentiment analysis

To conduct sentiment analysis on the

collected mobile app reviews, a Python program was utilised, which incorporated the TextBlob module. The TextBlob module was used to determine the overall sentiment of each review by employing the blob.sentiment.polarity function, which assigned a polarity score to each review. The polarity score ranged from [-1] to [1], where a negative score indicated negative sentiment, a positive score indicated positive sentiment, and a score of [0] indicated neutral sentiment.

Firstly, the collected reviews were imported into the program from the database and then each review was passed through the TextBlob module. The blob.sentiment.polarity function analysed the review and generated a polarity score. The program then assigned each review a label of "Positive," "Negative," or "Neutral" based on the polarity score generated.

The sentiment analysis output was stored back to the database, which allowed for easy tracking and analysis of the overall sentiment of each app. The sentiment analysis results were then used to gain insight into the general perception of each app, which could be used to identify areas for improvement in the app's features or design.

By using the TextBlob module in a Python program, it was possible to conduct sentiment analysis on the collected mobile app reviews. This allowed for easy tracking and analysis of the overall sentiment of each app, providing valuable insights into the perception of the apps by the users.

Labelling

Labelling is a critical process in NLP that enables the identification and extraction of relevant information from large amounts of unstructured data, such as mobile app reviews. In the context of mobile app reviews, labelling is particularly useful for identifying concerns related to cybersecurity and privacy. By applying pre-trained models and relevant keywords, the labelling process can help to quickly categorise and prioritise reviews based on the presence of specific security and privacy concerns.

To begin the labelling process, a pre-trained model can be applied to the reviews to identify and extract relevant features and patterns. For instance, a model that has been trained on cybersecurity and privacy related data can be applied to the app reviews to identify keywords and phrases that are commonly associated with security and privacy concerns. This can include terms such as "data breach," "password," "encryption," "data protection," "access control," "vulnerability," and "intrusion detection."

After the model has identified these keywords and phrases, review the output and apply additional labelling to the reviews. This involves manually reviewing a number of reviews and tagging those with relevant labels that correspond to the security and privacy concerns that were identified by the model. For instance, a review that mentions "data breach" might be tagged with a "data breach" label, while a review that mentions "password" might be tagged with a "password security" label.

By applying pre-trained models and relevant keywords, the labelling process can be more efficient and effective, as it enables the identification of potential security and privacy issues in a timely manner. Once the reviews have been labelled, the data can be analysed further to identify patterns and trends in the security and privacy concerns that are mentioned most frequently by users. This can be useful for app developers and security experts to identify areas of improvement and prioritise efforts to address potential vulnerabilities.

Pattern recognition

Pattern recognition is a crucial step in NLP analysis of mobile app user reviews, especially for identifying repeated complaints related to cybersecurity and privacy issues. In this step, a Python program utilising the 'nltk' module was used to process the reviews.

The first step was to extract the relevant keywords related to cybersecurity and privacy issues from a predefined list. This list includes the most commonly used keywords that indicate security or privacy concerns in mobile app reviews. The keywords were tokenized, stemmed, and lemmatized to increase the accuracy of pattern recognition.

Next, the program used regular expression matching to identify patterns in the reviews. The regular expressions were created using the extracted keywords to search for relevant phrases in the reviews. The program then counted the frequency of each pattern to identify the most commonly reported issues related to cybersecurity and privacy.

Finally, the program used clustering algorithms to group similar patterns together. This helped to identify patterns that were not immediately apparent from individual reviews but became apparent when similar reviews were grouped together. The clustering algorithm also helped to identify the most critical issues by grouping together patterns with the highest frequency.

By utilising a self-developed Python program and the 'nltk' module, it is able to accurately identify patterns in the reviews and group similar issues together to identify the most critical concerns.

User requirement extracting

Conducting user requirement extraction through NLP analysis is crucial in understanding users' needs and expectations for mobile app development. One of the primary steps in this process is identifying the keywords that are associated with general software requirements and particularly those related to cybersecurity and privacy concerns.

To start, it's essential to gather a dataset of user feedback from mobile app reviews. This dataset should include a wide range of reviews from different users to ensure that the extracted requirements are representative of the app's user base. Once the dataset is gathered, the next step is to preprocess the data, including removing any irrelevant information such as numbers and special characters and converting the text into lowercase.

The next step is to identify the keywords related to general software requirements. These can include terms such as "user interface," "functionality," "performance," and "usability." It's essential to identify synonyms and related terms for each keyword to ensure that all relevant requirements are captured.

Once the general requirements are identified, the focus should shift towards identifying keywords related to cybersecurity and privacy concerns. These can include terms such as "data protection," "security," "privacy policy," and "encryption." It's essential to identify any specific security or privacy concerns that are unique to the app's domain or industry.

After identifying the relevant keywords, the next step is to apply NLP techniques to extract the user requirements associated with each keyword. This can include sentiment analysis to understand whether the user feedback is positive or negative, topic modelling to identify the themes within the feedback, and entity recognition to identify specific features or aspects of the app that are mentioned in the feedback. By identifying the relevant keywords related to general software requirements and cybersecurity and privacy concerns, developers can ensure that the app meets the users' needs and is secure and private.

Consolidation and visualisation

After conducting NLP analysis on mobile app user reviews, the next step is to consolidate and visualise the findings. This step involves extracting useful insights from the data obtained during the NLP analysis and presenting them in a meaningful and easy-to-understand way. Consolidation and visualisation are crucial in making sense of the vast amount of data generated from the NLP analysis and identifying patterns and trends.

One important aspect of consolidation and visualisation is sentiment analysis. This involves analysing the overall sentiment of the reviews, such as positive, negative, or neutral, and identifying the most commonly used words or phrases that contribute to the sentiment. Visualising the results of sentiment analysis in the form of charts or graphs can provide a clear picture of the app's user sentiment and help identify areas that require improvement.

Another important aspect of consolidation and visualisation is pattern recognition. This involves identifying common themes or topics that occur frequently in the user reviews, such as usability, performance, or customer support. Visualising the results of pattern recognition in the form of word clouds or heat maps can help identify the most important topics and provide insights into the app's strengths and weaknesses.

Finally, consolidation and visualisation also play a crucial role in software requirement engineering. By identifying the most commonly mentioned features or functionalities in the user reviews, app developers can gain insights into the users' needs and preferences and prioritise their development efforts accordingly. Visualising the results of software requirement engineering in the form of feature maps or prioritisation matrices can help ensure that the app's development aligns with the users' needs and expectations.

Overall, consolidation and visualisation are essential steps in making sense of the vast amount of data generated during NLP analysis of mobile app user reviews. By presenting the findings in a clear and meaningful way, app developers can gain insights into the users' sentiments, identify patterns and trends, and prioritise their development efforts accordingly to improve the app's overall user experience.

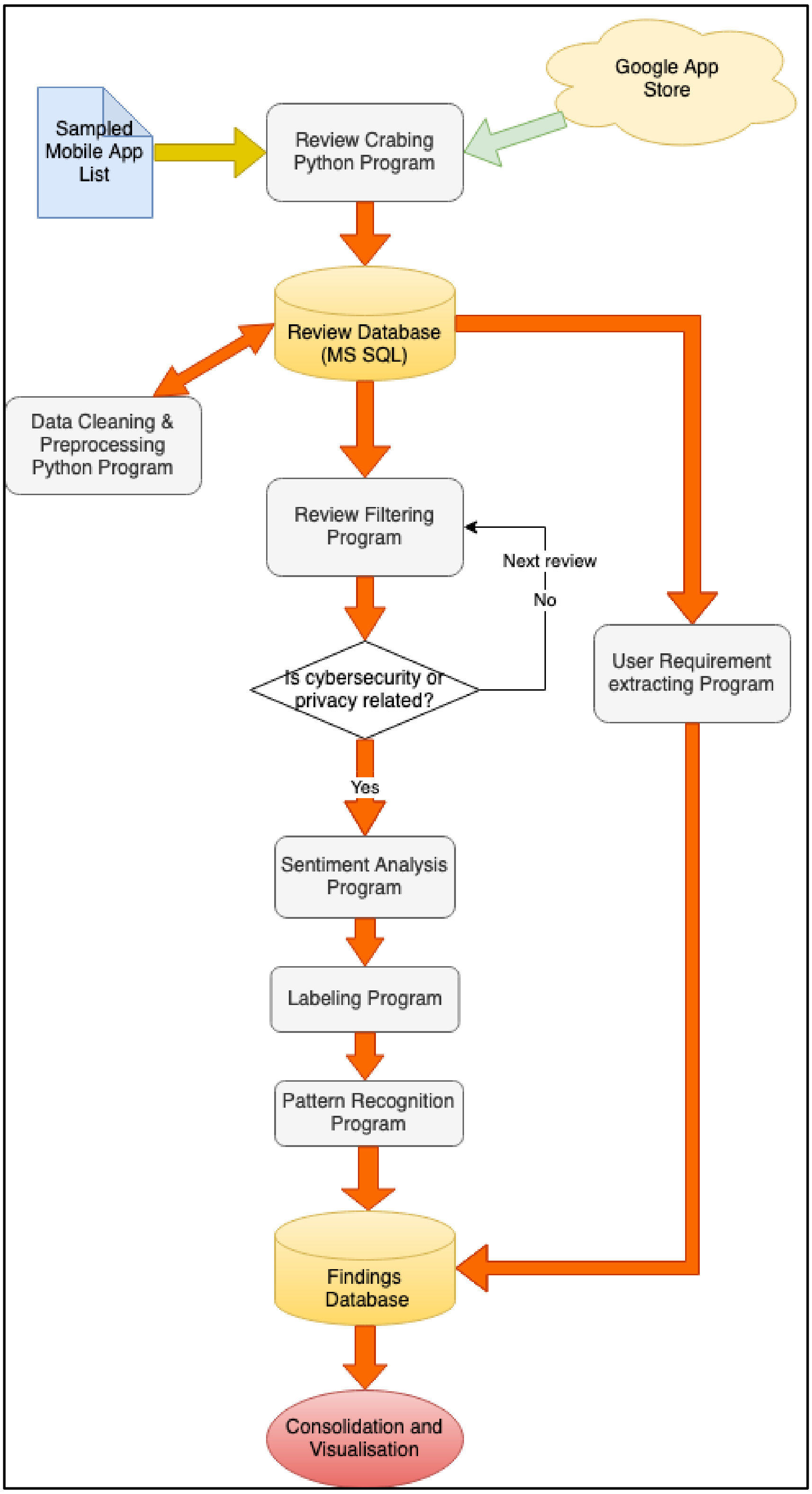
Conclusion

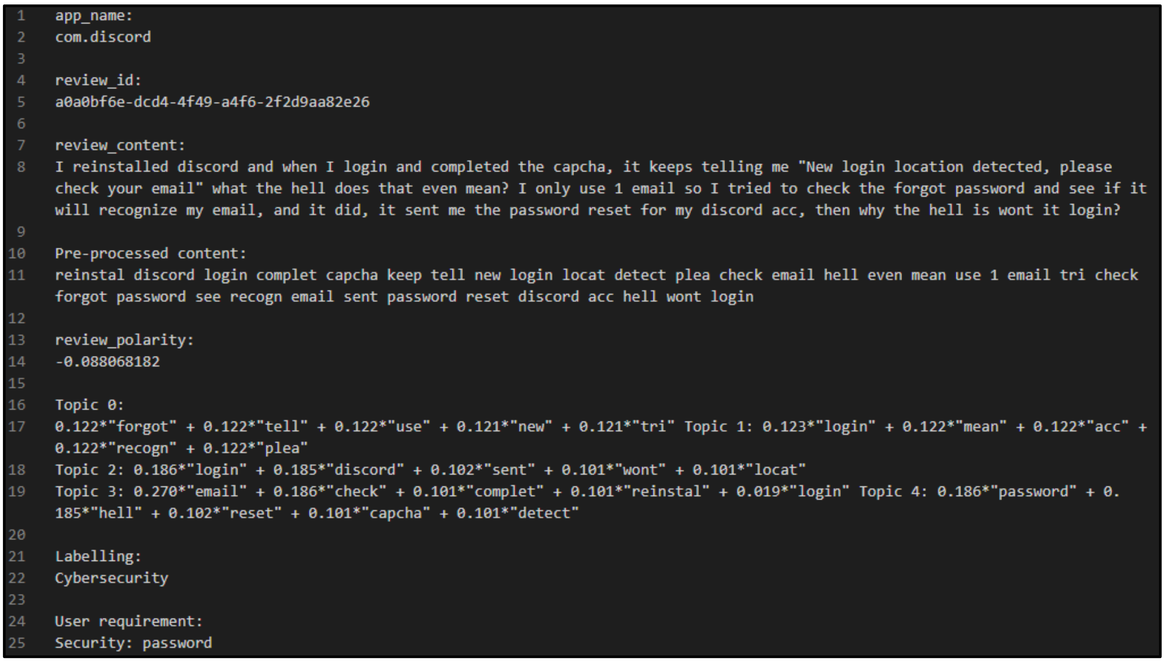
In conclusion, the use of a Python program and the Natural Language Toolkit module enabled the efficient collection, pre-processing, sentiment analysis, and labelling of mobile app reviews. The program successfully collected nearly 20 million reviews from six English-speaking countries and stored them in a Microsoft SQL database. Pre-processing steps such as tokenization, lowercasing, stop words removal, stemming and lemmatization, and removal of punctuation, special characters, and numbers ensured that the text data was clean and ready for analysis. Sentiment analysis using the TextBlob module provided valuable insights into the overall sentiment of each app, which could be used to identify areas for improvement. Finally, the labelling process enabled the identification and extraction of relevant information from large amounts of unstructured data, such as cybersecurity and privacy concerns. Overall, these techniques allowed for the efficient and effective analysis of mobile app reviews, providing valuable insights for app developers and stakeholders.

VIII. Project Design

Figure 1 demonstrated a user review of the "Discord" app as an example of how NLP techniques can be utilised for analysis purposes. The review is pre-processed to remove unnecessary words and analysed to identify the user's sentiment and the topics mentioned in the review. The analysis suggests that the user is having difficulty logging into the app and is frustrated with the app's email verification process. The user's requirement from the app is improved security, particularly in password-related issues. This review analysis presented the NLP analysis process used to extract user requirements from user reviews for app developers to make necessary improvements to their apps.

Workflow of NLP Analysis


Google App Store
Sampled Mobile App List
Review Crabing Python Program
Review Database (MS SQL)
Data Cleaning & Preprocessing Python Program
Review Filtering Program
Next review
No
Is cybersecurity or privacy related?
Yes
User Requirement extracting Program
Sentiment Analysis Program
Labeling Program
Pattern Recognition Program
Findings Database
Consolidation and Visualisation



```
app_name:
com.discord

review_id:
a0a0bf6e-dcd4-4f49-a4f6-2f2d9aa82e26

review_content:
I reinstalled discord and when I login and completed the capcha, it keeps telling me "New login location detected, please
check your email" what the hell does that even mean? I only use 1 email so I tried to check the forgot password and see if it
will recognize my email, and it did, it sent me the password reset for my discord acc, then why the hell is wont it login?

Pre-processed content:
reinstal discord login complet capcha keep tell new login locat detect plea check email hell even mean use 1 email tri check
forgot password see recogn email sent password reset discord acc hell wont login

review_polarity:
-0.088068182

Topic 0:
0.122*"forgot" + 0.122*"tell" + 0.122*"use" + 0.121*"new" + 0.121*"tri" Topic 1: 0.123*"login" + 0.122*"mean" + 0.122*"acc" +
0.122*"recogn" + 0.122*"plea"
Topic 2: 0.186*"login" + 0.185*"discord" + 0.102*"sent" + 0.101*"wont" + 0.101*"locat"
Topic 3: 0.270*"email" + 0.186*"check" + 0.101*"complet" + 0.101*"reinstal" + 0.019*"login" Topic 4: 0.186*"password" + 0.
185*"hell" + 0.102*"reset" + 0.101*"capcha" + 0.101*"detect"

Labelling:
Cybersecurity

User requirement:
Security: password
```

Figure 1. - Output of the review analysis through the NLP program in Python.

IX. Results and analysis

This project collected nearly 20 million reviews of the 780 most downloaded Google apps across 45 categories from six English-speaking countries. The goal is to analyse these reviews using NLP technologies. The study aimed to answer three research questions, and the following are the findings:

General findings

This research project analysed a total of almost 20 million reviews and further analysed a targeted subset of 2.7 million reviews to identify user requirements and cybersecurity/privacy related issues. The analysis found that cybersecurity is a more frequently mentioned issue than privacy, with over 230,000 reviews mentioning cybersecurity and just over 52,000 mentioning privacy shown in Table 1.

Table 1. - No. of User Reviews Identified with Cybersecurity / Privacy Concerns

All Review	19,844,888	100%
Cybersecurity	230,733	1.16%
Privacy	52,819	0.27%
Both	3,875	0.02%

Furthermore, the percentage of reviews mentioning cybersecurity or privacy concerns is relatively low, with cybersecurity issues mentioned in only 1.16% of reviews and privacy issues in only 0.27%. A small proportion of reviews, just 0.02%, mentioned both cybersecurity and privacy concerns.

The study also identified 241 patterns in the reviews that shed light on how users interact with the apps. Finally, the analysis revealed over 2.7 million user requirements, indicating what users expect from the apps.

The mobile review findings above reveal the top 15 mobile apps with cybersecurity and privacy issues. Table 2 shows the top 15 mobile apps with cybersecurity issues and the number of reviews that identified issues, as well as the percentage of the total reviews with identified issues.

Table 2. - Top 15 Mobile Apps with Cybersecurity Issues

Rank	App Name	Identified Reviews	Percent
1	Clash of Clans	5,345	2.28%
2	Coin Master	5,148	2.19%
3	Azure Authenticator	4,795	2.04%
4	Tinder	4,413	1.88%
5	Lotus InTouch	3,401	1.45%
6	Facebook	3,084	1.31%
7	PayPal	2,957	1.26%
8	Lords Mobile	2,636	1.12%
9	Bumble	2,542	1.08%
10	OLX South Asia	2,460	1.05%
11	One Boost	2,283	0.97%
12	Pokemon Go	2,081	0.89%
13	FIFA Mobile	1,951	0.83%
14	Roblox	1,807	0.77%
15	ADP Mobile	1,708	0.73%
	Total	46,611	19.87%

The top five mobile apps with cybersecurity issues include Clash of Clans, Coin Master, Azure Authenticator, Tinder, and Lotus InTouch. These apps have the highest number of reviews that identified cybersecurity issues, accounting for almost 10% of the total reviews with identified issues.

Table 3 shows the top 15 mobile apps with privacy issues and the number of reviews that identified issues, as well as the percentage of the total reviews with identified issues. The top five mobile apps with privacy issues were OLX South Asia, Meesho Supply, Flipkart, AJIO, and Amazon Shopping. These apps have the highest number of reviews that identified privacy issues, accounting for over 20% of the total reviews with identified issues.

Table 3. - Top 15 Mobile Apps with Privacy Issues

Rank	App Name	Identified Reviews	Percent
1	OLX: Buy & Sell Near You	3,637	6.42%
2	Meesho: Resell, Work From Home	2,165	3.82%
3	Flipkart Online Shopping App	1,730	3.05%
4	AJIO Online Shopping App	1,361	2.40%
5	Amazon India Shopping	1,336	2.36%
6	Myntra - Online Shopping App	1,310	2.31%
7	Cash App	1,238	2.18%
8	Ola. Get rides on-demand	1,152	2.03%
9	Swiggy Food Order & Delivery	1,110	1.96%
10	Bajaj Finserv	1,102	1.94%
11	PhonePe – UPI Payments, Recharges & Money Transfer	1,043	1.84%
12	Paytm - Mobile Recharge, UPI Payments & Bank App	1,010	1.78%
13	JioMart: Indian Online Grocery	949	1.67%

14	Zomato - Online Food Delivery & Restaurant Reviews	927	1.64%
15	Google Pay (Tez) - a simple and secure payment app	863	1.52%
	Total	20,933	36.92%

The cybersecurity findings reveal that the Shopping and Finance app categories have the highest number of identified reviews with 29,366 and 27,855, respectively in Table 4.

Table 4. - Mobile App Review Findings for Cybersecurity by App Categories

Ranking	Category	Identified Reviews	Percent
1	Shopping	29,366	12.52%
2	Finance	27,855	11.87%
3	Tools	19,646	8.37%
4	Entertainment	15,697	6.69%
5	Communication	15,272	6.51%
6	Food & Drink	12,926	5.51%
7	Business	12,830	5.47%
8	Social	11,876	5.06%
9	Casual	8,850	3.77%
10	Strategy	8,121	3.46%
	Others	72,169	30.76%

These categories are followed by Tools, Entertainment, Communication, and Food & Drink. Shopping is the most affected app category with 12.52% of its reviews indicating cybersecurity issues, followed by Finance with 11.87% and Tools with 8.37% in Figure 2.

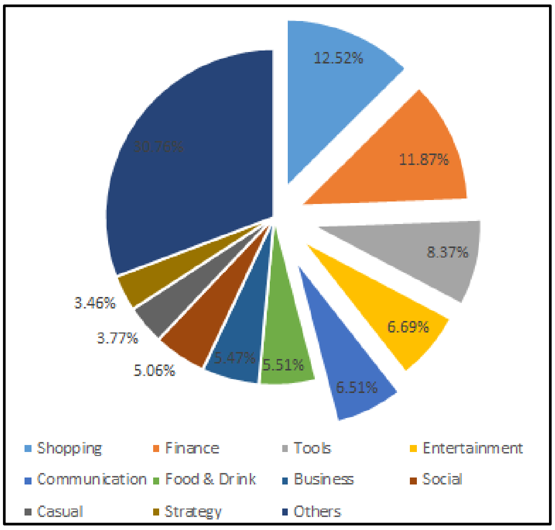


Figure 2. - Cybersecurity Issues by App Categories

On the other hand, the privacy findings show that Shopping has the highest number of identified reviews with 17,883, followed by Finance with 12,013, and Food & Drink with 4,633 in Table 5.

Table 5. - Mobile App Review Findings for Privacy by App Categories

Ranking	Category	Identified Reviews	Percent
1	Shopping	17,883	31.54%
2	Finance	12,013	21.19%
3	Food & Drink	4,633	8.17%
4	Communication	2,953	5.21%
5	Maps & Navigation	2,863	5.05%
6	Entertainment	2,621	4.62%
7	Tools	2,004	3.53%
8	Travel & Local	1,871	3.30%
9	Business	1,169	2.06%
10	Social	1,028	1.81%

	Others	7656	13.50%

Figure 3 for privacy issues shows that Shopping has the highest percentage of privacy issues with 31.54%, followed by Finance with 21.19% and Food & Drink with 8.17%.

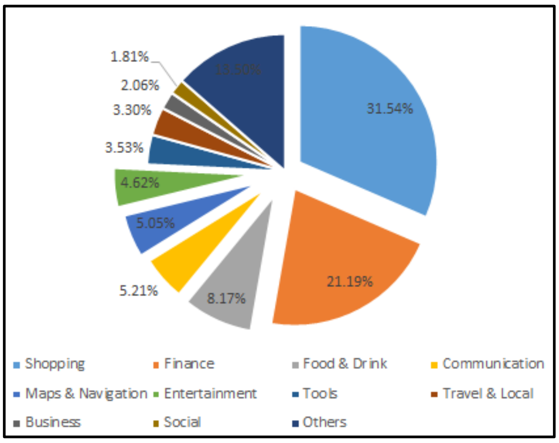


Figure 3. - Privacy Issues by App Categories

The use of NLP to analyse user reviews of mobile apps has provided valuable insights into cybersecurity and privacy issues. The findings indicate that shopping and finance apps have the highest percentage of such issues, highlighting the need for developers to prioritise security measures in these categories. Additionally, NLP can help identify specific areas of concern within each category, allowing developers to target their efforts more effectively. Overall, the use of NLP in analysing user reviews can play an important role in improving the cybersecurity and privacy of mobile apps.

RQ1: Sentiment Analysis

The Sentiment Analysis of user reviews for mobile apps has revealed some interesting findings. Out of the total 2,751,465 reviews analysed, 49.15% were positive, 18.78% are neutral, and 32.07% are negative. The average review polarity index is 0.059 approximately , indicating that the overall sentiment of the reviews was slightly positive.

The Figure 4 below provides a visual representation of the sentiment distribution in the user reviews of mobile apps. It shows that the majority of the reviews are negative or neutral, which indicates that there is room for improvement in terms of user satisfaction with mobile apps.

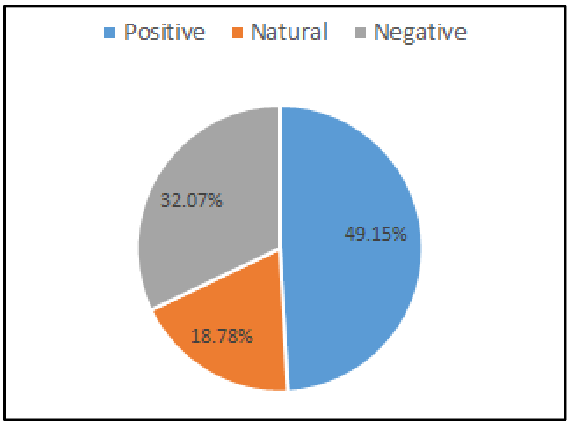
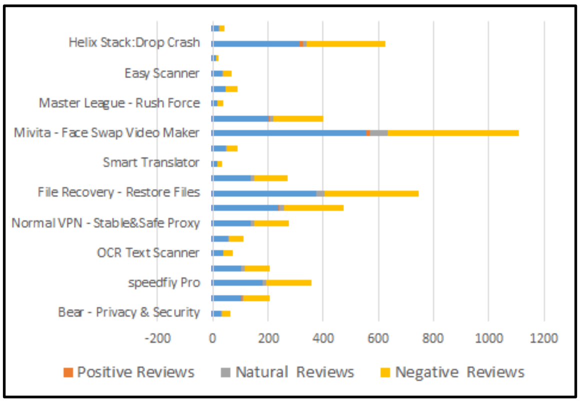


Figure 4. - Overall Sentiment Analysis Result

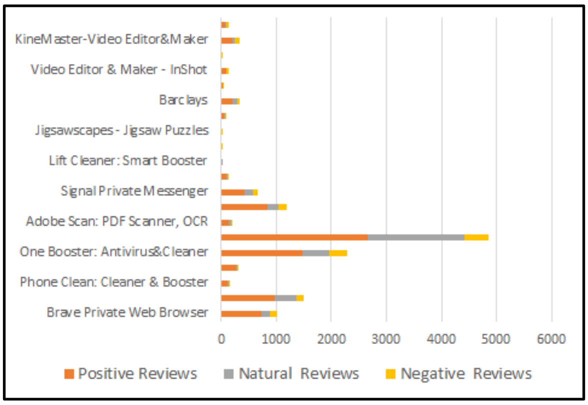
The Sentiment Analysis results for mobile app user reviews with identified cybersecurity /privacy issues are presented in two tables. Table 6 shows the top 20 mobile apps with negative reviews rate, and the second table displays the top 20 mobile apps with positive reviews. In the first table, the app "Bear - Privacy & Security" has the highest number of negative reviews rate, with a total of 29 out of 31 reviews being negative. Similarly, the app "iFace: AI Cartoon Photo Editor" has 98 negative reviews out of a total of 104 reviews. The app "Speedify Pro" has 163 negative reviews out of a total of 179 reviews.

Table 6. - Top 20 Negative Reviews Mobile Apps



On the other hand, Table 7 shows the top 20 mobile apps with positive reviews rate. The app "Brave Private Web Browser" has the highest number of positive reviews, with 725 positive reviews out of a total of 1005 reviews. The app "Phone Guardian VPN: Safe WiFi" has 970 positive reviews out of a total of 1505 reviews. The app "Phone Clean: Cleaner & Booster" has 105 positive reviews out of a total of 142 reviews.

Table 7. - Top 20 Positive Reviews Mobile Apps



Overall, the Sentiment Analysis results indicated that some mobile apps with cybersecurity/privacy issues received significantly more negative reviews than positive ones, while others received predominantly positive reviews.

According to the study against 45 app categories shown in Table 8, the top category with the most positive sentiment is Books & Reference, with an average polarity score of 0.097. This suggests that users generally have a favourable opinion of mobile apps in this category. The Business category follows closely behind, with an average polarity score of 0.069.

On the other hand, the bottom categories with the most negative sentiment are Others, Card, and Auto & Vehicles, with average negative polarity scores of -0.249, -0.199, and -0.189, respectively. This implies that users tend to express negative opinions about mobile apps in these categories.

Table 8. - Sentiment Analysis Results for Mobile App User Reviews by Category

Ranking	Category	Total	Positive	Natural	Negative	Avg. Polarity
1	Books & Reference	524	286	110	128	0.097
2	Business	13887	5948	4135	3804	0.070
3	Weather	139	67	33	39	0.053
4	Communication	18030	8123	4226	5681	0.046
5	Education	1220	533	269	418	0.027
...						
41	Puzzle	2807	566	258	1983	-0.174
42	Role Playing	679	172	32	475	-0.183
43	Simulation	974	185	140	649	-0.184
44	Auto & Vehicles	714	139	171	404	-0.189
45	Card	1135	202	181	752	-0.199

The sentiment analysis results for user reviews of mobile apps show that some app categories have significantly more positive sentiment than others. The top five categories with the highest average polarity are Books & Reference, Business, Weather, Communication, and Education, while Card, Auto & Vehicles, Simulation, Role Playing, Puzzle, Trivia, News & Magazines, Dating, Board, and Sports have more negative reviews than positive. Interestingly, the Tools and Medical categories have a large number of reviews but low average polarity. Developers should focus on improving user experience and addressing privacy and security concerns, especially for categories with negative reviews.

RQ2: Pattern recognition

Table 9 shows the frequency of patterns identified through NLP analysis of mobile app user reviews. The patterns listed are primarily related to security concerns, with the most top three frequent patterns being "ad", "share", "password", which were mentioned 1,010,505 times in total, occupying 71% of total identified patterns. The figure suggests that users are primarily concerned about the security of their personal information when using mobile apps. The fact that patterns related to security concerns were mentioned so frequently in user reviews indicates that users are paying close attention to how mobile apps handle their data and protect their privacy.

Table 9. - Pattern Identified of Security Concerns

Detected Pattern	Frequency	Percent
ad	823,908	57.60%
share	99,747	6.97%
password	86,850	6.07%
complaint	63,887	4.47%
fake	52,851	3.69%
scam	50,196	3.51%
cheat	29,503	2.06%
hack	28,148	1.97%
fraud	25,398	1.78%
spam	20,622	1.44%
attack	15,403	1.08%
upset	15,121	1.06%
loss	13,782	0.96%
virus	9,461	0.66%
shock	8,673	0.61%
destroy	8,140	0.57%

Table 10 shows the frequency of patterns identified through NLP analysis of mobile app user reviews, those reviews were identified with multiple issues. The figure suggests that users are primarily concerned about security issues related to advertising and sharing of personal data. The top five most frequent patterns identified all involve advertising, with "ad,fake" being the most common pattern identified. This indicates that users are likely experiencing fake advertisements within the app, which can be both frustrating and potentially harmful.

Table 10. - Multiple patterns Identified of Security Concerns

Detected Pattern	Frequency
ad, fake	5,475
ad, share	5,393
ad, scam	4,056
scam, fake	2,321
ad, spam	2,291
ad, complaint	1,965
spam, ad	1,822
complaint, ad	1,471
ad, cheat	992

ad, destroy	904

In addition, the two tables present the frequency of patterns identified through NLP analysis of mobile app user reviews related to security concerns. Table 9 shows that users are primarily concerned about the security of their personal information, with "ad," "share," and "password" being the top three frequent patterns mentioned. This indicates that users are paying close attention to how mobile apps handle their data and protect their privacy. Moreover, patterns related to security concerns were mentioned frequently in user reviews, suggesting that users are highly concerned about their privacy and security when using mobile apps.

Table 10 presents multiple patterns identified through NLP analysis of mobile app user reviews that were identified with multiple issues. This table shows that users are primarily concerned about security issues related to advertising and sharing of personal data. The most frequent patterns identified involve advertising, with "ad, fake" being the most common pattern identified. This suggests that users are likely experiencing fake advertisements within the app, which can be both frustrating and potentially harmful. Moreover, several patterns involve scams, spam, and other types of fraudulent activity, indicating that users are concerned about the safety of their personal information and the potential for financial loss or identity theft.

Overall, the data presented in both tables indicates that mobile app users are highly concerned about cybersecurity and privacy. They pay close attention to how their personal data is handled and protected by mobile apps.

RQ3: Software requirement extraction

The NPL analysis of mobile app user reviews for software user requirements extraction has revealed some interesting insights into user requirements. The total number of reviews analysed was 2,717,947, and the data was categorised into different aspects of the app, including Content, Design, Features, Performance, Usability, Security, and Privacy.

The results in Figure 5 show that the most critical user requirement is related to Features, with 63.5% of the total reviews mentioning this aspect. This indicates that users are looking for apps that provide comprehensive and innovative features. Performance is also an important aspect, with 19.3% of the reviews mentioning it. Usability and Content are also significant, with 3% and 6.2% of the reviews respectively. Interestingly, Security and Privacy are not major concerns for most users, with only 6.1% and 1.3% of the reviews mentioning these aspects.

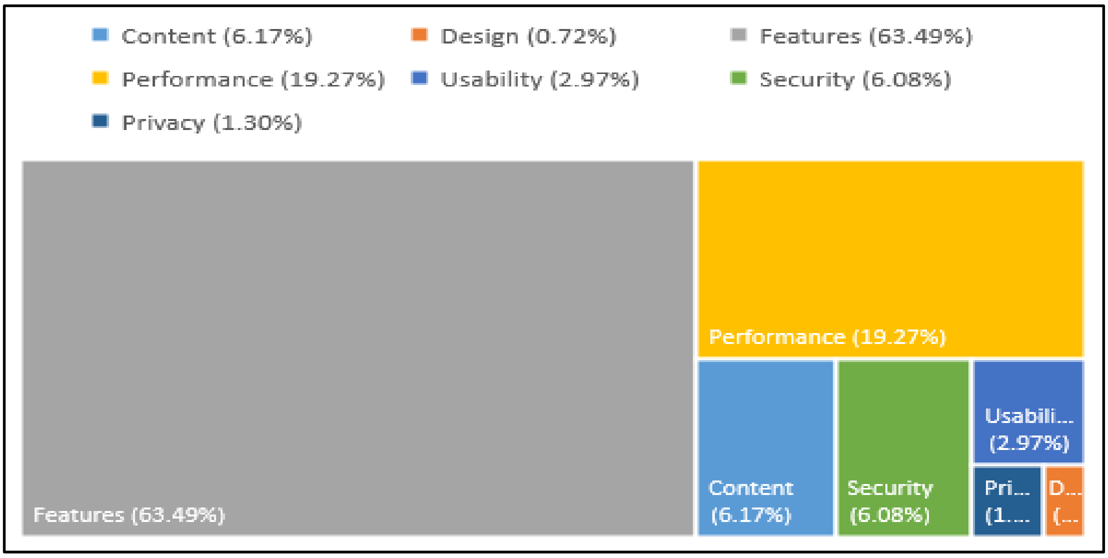


Figure 5. - User Requirements Extraction by Aspects

Table 11 lists the top 15 mobile applications with user requirement extraction as a result of this study. The ranking is based on the total number of requirements, as well as the number of security and privacy requirements. OLX, a shopping app, topped the list with a total of 5593 requirements, of which 2460 were related to security and 3637 were related to privacy. Clash of Clans, a strategy game, ranked second with 5369 requirements, of which 5345 were security requirements.

Table 11. - Top 15 Apps with User Requirement Extracted

Ranking	App Name	Category	Total	Security	Privacy
1	OLX: Buy & Sell Near You with...	Shopping	5593	2460	3637
2	Clash of Clans	Strategy	5369	5345	25
3	Coin Master	Casual	5228	5148	98
4	Tinder: Dating app. Meet. Chat	Dating	4879	4413	544
5	Microsoft Authenticator	Business	4855	4795	69
6	PayPal - Send, Shop, Manage	Finance	3743	2957	848
7	YONO SBI: Banking & Lifestyle	Finance	3549	3401	168
8	Facebook	Social	3295	3084	239
9	Flipkart Online Shopping App	Shopping	3245	1622	1730
10	Meesho: Online Shopping App	Shopping	3228	1221	2165
11	Amazon India Shop, Pay, miniTV	Shopping	2835	1595	1336
12	Bumble - Dating. Friends. Bizz	Dating	2830	2542	319
13	Lords Mobile: Kingdom Wars	Strategy	2720	2636	100
14	Myntra - Fashion Shopping App	Shopping	2512	1270	1310
15	AJIO - House Of Brands	Shopping	2416	1136	1361

Other notable apps on the list include Coin

Master, Tinder, Microsoft Authenticator, PayPal, and Facebook. Shopping apps were particularly well represented, with six of the top 15 apps falling into that category. The study highlights the importance of considering user requirements, particularly related to security and privacy, when designing and developing mobile applications.

Figure 6 displays the extracted user requirements by app categories, including overall, cybersecurity and privacy requirements. Shopping and Finance apps have the most requirements extracted, with 47249 and 39868, respectively. Among the app categories, shopping apps have the highest number of cybersecurity requirements at 29366, while communication apps have the highest number of privacy requirements at 2953. The results show that users have a diverse range of requirements across various app categories, with shopping and finance apps having the most extensive requirements.

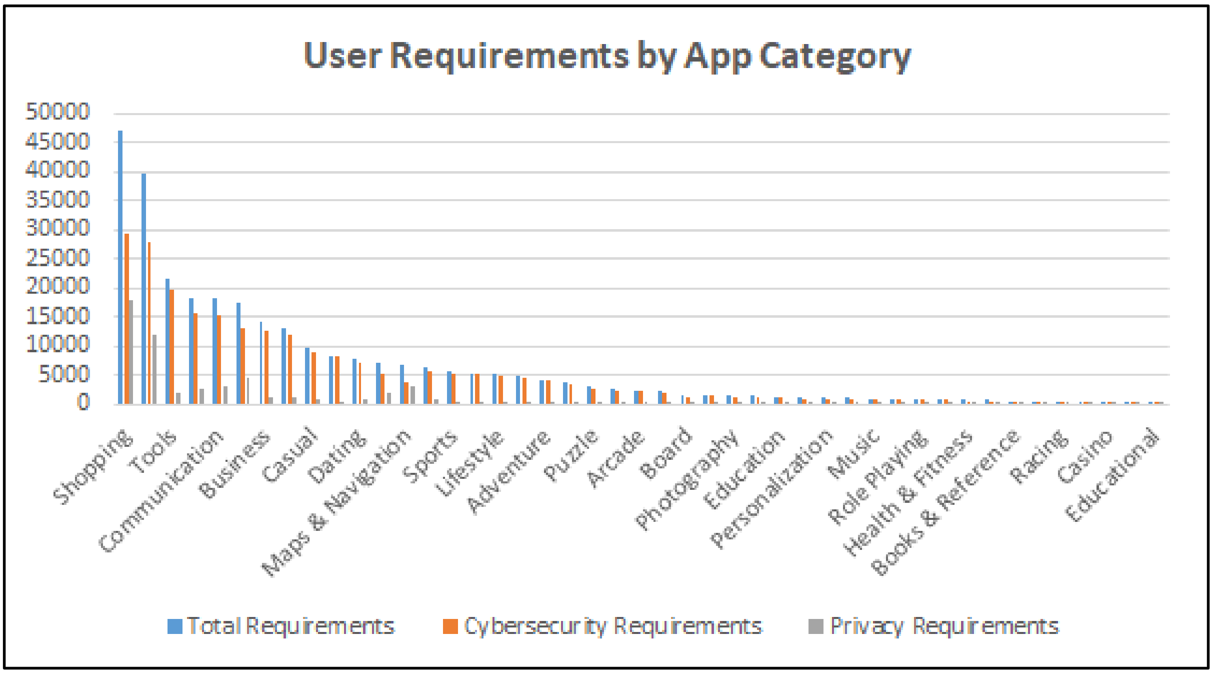


Figure 6. - User Requirements by App Category

Table 12 shows that the most frequent cybersecurity-related keywords matched for user requirements are 'password' (69,666 reviews), 'attack' (12,604 reviews), 'virus' (9,083 reviews), and 'fake' (57,286 reviews). Meanwhile, the most frequent privacy related keywords matched are 'fraud' (28,084 reviews), 'leak' (1,655 reviews), 'consent' (2,895 reviews), and 'breach' (1,960 reviews).

Table 12. - Cybersecurity and Privacy Keywords Matched for User Requirements in Mobile App Reviews

	Keywords	No. of Matched
Security	virus	9,083
	trojan	218
	threat	1,116
	tamper	493
	password	69,666
	jailbreak	193

	hijack	1,078
	fake	57,286
Privacy	fraud	28,084
	untrust	478
	leak	1,655
	consent	2,895
	breach	1,960

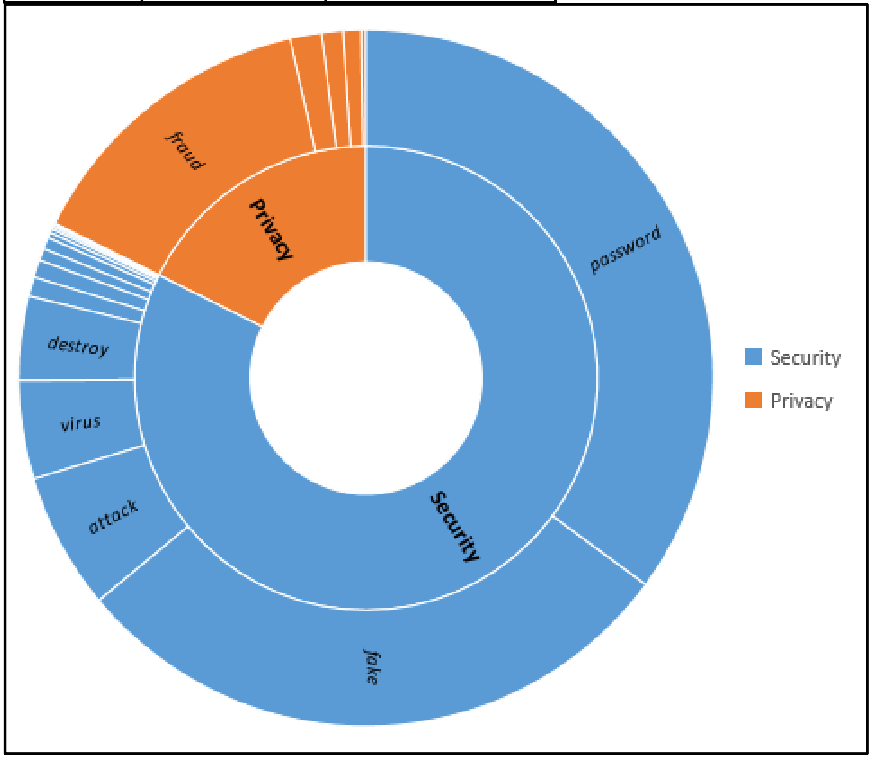


Figure 7. - Frequence of cybersecurity/privacy keywords matched in user reviews

The overall results show that NLP plays a crucial role in extracting user requirements effectively, as the large amounts of data from user reviews can be analysed and categorised effectively. This allows app developers to identify common user concerns and make necessary improvements to their app's security and privacy features. In summary, the data provides valuable insights into user requirements and highlights the importance of using NLP to extract and analyse user feedback.

Interestingly, some keywords are found to relate to both cybersecurity and privacy concerns in Figure 7. For example, 'untrust' (478 reviews), 'fraud' (17,773 reviews), and 'fake' (57,286 reviews) can be considered both cybersecurity and privacy issues. Moreover, 'attack' and 'destroy' (167 reviews) can be considered as both cybersecurity and privacy threats, as user data can be damaged or deleted in an attack.

NLP accuracy analysis by manual verification

The NLP analysis results for sentiment analysis and cybersecurity/privacy issue identification were manually verified using a sample of 400 mobile app user reviews. The purpose of the manual verification was to determine the accuracy of the NLP analysis results.

The process involved selecting a sample of 400 reviews randomly and manually analysing each review according to its contents. The results of the manual verification were then compared to the NLP analysis results to evaluate the accuracy of the NLP analysis.

The results showed that the NLP analysis had an overall accuracy rate of 93% for sentiment analysis and 74% for cybersecurity/privacy issue identification. These results indicate that the NLP analysis is highly accurate in determining user sentiment but may need some improvement in identifying cybersecurity/privacy related concerns.

The high accuracy rate for sentiment analysis suggests that the NLP algorithm used for this analysis is effective in identifying positive and negative sentiment in mobile app user reviews. However, the lower accuracy rate for cybersecurity/privacy issue identification suggests that the algorithm may not be capturing all user concerns related to these issues. Keyword definition is a key issue for effective cybersecurity and privacy issue identification through NLP analysis. The accuracy of NLP analysis results can be impacted by the quality of the keyword selection process, particularly in the area of cybersecurity and privacy concerns. Therefore, it is crucial to have a clear and comprehensive understanding of relevant keywords and their definitions.

To improve the accuracy of pattern recognition for cybersecurity/privacy issue identification, the first step is to refine the set of keywords used in the process. This can be achieved by conducting a thorough review of the current set of keywords, identifying missing or redundant keywords, and identifying potential new keywords based on user concerns that were not previously captured. The refined list of keywords should then be tested by running the pattern recognition process on a sample of mobile app user reviews. The results of the pattern recognition process can then be analysed and compared to the previous results. Adjustments can be made to the list of keywords as needed, and retesting should be conducted until the desired level of pattern recognition accuracy, such as over 75%, is achieved To ensure the accuracy of the refined list of keywords, consultation with experts in the field, review of related research, and a pilot study are also necessary steps.

Overall, the NLP analysis results for sentiment analysis and cybersecurity/privacy issue identification were highly accurate, with an overall accuracy rate of 93% and 74%, respectively. These results suggest that the NLP analysis is an effective tool for analysing large datasets of mobile app user reviews and can provide valuable insights into user sentiment and concerns related to cybersecurity and privacy.

Improve keywords for cybersecurity/privacy issue identification

The initial round of cybersecurity and privacy issue recognition in the project had a low accuracy rate, which was attributed to the predefined keywords. The importance of keywords in NLP recognition was studied and two articles (Paolanti & Frontoni, 2020; Chen et al., 2021) were consulted for guidance. After analysing the articles and refining the list of keywords, pilot runs were conducted several times until a satisfactory accuracy rate was achieved. This process of refining keywords for NLP pattern recognition is a common strategy employed by researchers and practitioners to improve the accuracy of NLP models.

The process of refining the keywords involved conducting a thorough review of the current set of keywords, identifying missing or redundant keywords, and identifying potential new keywords based on user concerns that were not previously captured. We consulted with experts in the field, reviewed related research, and conducted a pilot study to refine the list of keywords.

The use of NLP technology in cybersecurity and privacy issue identification is gaining increasing attention due to its potential to automate the identification of potential threats and vulnerabilities in large amounts of unstructured data. However, achieving high accuracy rates in NLP models requires careful consideration of the keywords used in the pattern recognition process. The refinement of keywords through consultation with experts, review of related research, and pilot runs is a common approach to improving accuracy rates in NLP models. By adopting this approach, the project was able to achieve a satisfactory accuracy rate in identifying cybersecurity and privacy issues, which is crucial for ensuring the security and privacy of mobile app users.

X. Conclusion

The Mobile Apps Cybersecurity and Privacy Analysis Through User Feedback Reviews project was conducted with the aim of identifying potential cybersecurity and privacy issues in mobile apps through analysis of user feedback reviews. The project utilised natural language processing techniques to automate the identification of potential issues and vulnerabilities in large amounts of unstructured data.

Throughout the course of the project, several challenges were encountered, including low accuracy rates in initial rounds of issue recognition, difficulties in identifying relevant keywords, and limitations in the amount and quality of available data. However, by employing a range of strategies, including refining the list of keywords, consulting with experts, and conducting pilot studies, these challenges were successfully addressed.

The findings of the project highlighted the need for increased attention to be paid to cybersecurity and privacy issues in mobile apps. The analysis revealed a significant number of potential vulnerabilities, including issues with data collection, data storage, and data sharing. The findings also demonstrated the importance of user feedback as a source of information for identifying potential issues, and the potential of NLP techniques to automate the analysis of large amounts of user feedback data.

The project has several implications for mobile app developers, regulatory bodies, and users. Mobile app developers need to pay closer attention to cybersecurity and privacy issues, and to implement appropriate measures to protect user data. Regulatory bodies can play a key role in ensuring that mobile apps are developed and deployed in a manner that prioritises user security and privacy. Finally, users need to be aware of potential vulnerabilities in mobile apps and take steps to protect their personal information.

Overall, the Mobile Apps Cybersecurity and Privacy Analysis Through User Feedback Reviews research project has provided valuable insights into the potential cybersecurity and privacy risks associated with mobile apps. By employing a range of advanced techniques and strategies, the project has demonstrated the potential of NLP techniques to automate the analysis of large amounts of user feedback data, and to identify potential vulnerabilities and issues that may otherwise have gone undetected. The findings of the project have important implications for mobile app developers, regulatory bodies, and users, and serve as a call to action for increased attention to cybersecurity and privacy issues in the mobile app industry.

Reference

1. Abouelmehdi, K., Beni-Hssane, A., Khaloufi, H. and Saadi, M., 2017. Big data security and privacy in healthcare: A Review. *Procedia Computer Science*, *113*, pp.73-80.

2. Abualigah, L., Alfar, H.E., Shehab, M. and Hussein, A.M.A., 2020. Sentiment analysis in healthcare: a brief review. *Recent advances in NLP: the case of Arabic language*, pp.129-141.

3. ACHAR, S., PATEL, H. and HUSSAIN, S., 2022. DATA SECURITY IN CLOUD: A REVIEW. *Asian Journal of Advances in Research*, pp.76-83.

4. Kaluarachchi, T., Reis, A. and Nanayakkara, S., 2021. A review of recent deep learning approaches in human-centered machine learning. *Sensors*, *21*(7), p.2514.

5. Malavolta, I., Ruberto, S., Soru, T. and Terragni, V., 2015, May. Hybrid mobile apps in the google play store: An exploratory investigation. In *2015 2nd ACM international conference on mobile software engineering and systems* (pp. 56-59). IEEE.

6. Andraško, J., Mesarčík, M. and Hamuľák, O., 2021. The regulatory intersections between artificial intelligence, data protection and cyber security: challenges and opportunities for the EU legal framework. *AI & SOCIETY*, pp.1-14.

7. Aristodemou, L. and Tietze, F., 2018. The state-of-the-art on Intellectual Property Analytics (IPA): A literature review on artificial intelligence, machine learning and deep learning methods for analysing intellectual property (IP) data. *World Patent Information*, *55*, pp.37-51.

8. Balducci, B. and Marinova, D., 2018. Unstructured data in marketing. *Journal of the Academy of Marketing Science*, *46*, pp.557-590.

9. Bhavani, A. and Kumar, B.S., 2021, April. A review of state art of text classification algorithms. In *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 1484-1490). IEEE.

10. *Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A. and Agarwal, S., 2020. Language models are few-shot learners. Advances in neural information processing systems, 33, pp.1877-1901.*

11. Cambria, E. and White, B., 2014. Jumping NLP curves: A review of natural language processing research. *IEEE Computational intelligence magazine*, *9*(2), pp.48-57.

12. Chen, J.H., Su, M.C., Azzizi, V.T., Wang, T.K. and Lin, W.J., 2021. Smart project management: interactive platform using natural language processing technology. *Applied Sciences*, *11*(4), p.1597.

13. Chen, X., Ding, R., Xu, K., Wang, S., Hao, T. and Zhou, Y., 2018. A bibliometric review of natural language processing empowered mobile computing. *Wireless Communications and Mobile Computing*, *2018*.

14. Chiche, A. and Yitagesu, B., 2022. Part of speech tagging: a systematic review of deep learning and machine learning approaches. *Journal of Big Data*, *9*(1), pp.1-25.

15. Conger, S., Pratt, J.H. and Loch, K.D.,

2013. Personal information privacy and emerging technologies. *Information Systems Journal*, *23*(5), pp.401-417.

16. *Devlin, J., Chang, M.W., Lee, K. and Toutanova, K., 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.*

17. Dogra, V., Verma, S., Chatterjee, P., Shafi, J., Choi, J. and Ijaz, M.F., 2022. A complete process of text classification system using state-of-the-art NLP models. *Computational Intelligence and Neuroscience*, *2022*.

18. Dreisbach, C., Koleck, T.A., Bourne, P.E. and Bakken, S., 2019. A systematic review of natural language processing and text mining of symptoms from electronic patient -authored text data. *International journal of medical informatics*, *125*, pp.37-46.

19. Eberendu, A.C., 2016. Unstructured Data: an overview of the data of Big Data. *International Journal of Computer Trends and Technology*, *38*(1), pp.46-50.

20. Abd Elaziz, M., Al-qaness, M.A., Ewees, A.A. and Dahou, A. eds., 2019. Recent Advances in NLP: The Case of Arabic Language.

21. Fu, S., Chen, D., He, H., Liu, S., Moon, S., Peterson, K.J., Shen, F., Wang, L., Wang, Y., Wen, A. and Zhao, Y., 2020. Clinical concept extraction: a methodology review. *Journal of biomedical informatics*, *109*, p.103526.

22. Genc-Nayebi, N. and Abran, A., 2017. A systematic literature review: Opinion mining studies from mobile app store user reviews. *Journal of Systems and Software*, *125*, pp.207-219.

23. Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J. and Chen, T., 2018. Recent advances in convolutional neural networks. *Pattern recognition*, *77*, pp.354-377.

24. Hirschberg, J. and Manning, C.D., 2015. Advances in natural language processing. *Science*, *349*(6245), pp.261-266.

25. Hutto, C. and Gilbert, E., 2014, May. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the international AAAI conference on web and social media* (Vol. 8, No. 1, pp. 216-225).

26. Donnelly, L.F., Grzeszczuk, R. and Guimaraes, C.V., 2022, April. Use of natural language processing (NLP) in evaluation of radiology reports: an update on applications and technology advances. In *Seminars in Ultrasound, CT and MRI* (Vol. 43, No. 2, pp. 176-181). WB Saunders.

27. Kang, Y., Cai, Z., Tan, C.W., Huang, Q. and Liu, H., 2020. Natural language processing (NLP) in management research: A literature review. *Journal of Management Analytics*, *7*(2), pp.139-172.

28. Chary, M., Parikh, S., Manini, A.F., Boyer, E.W. and Radeos, M., 2019. A review of natural language processing in medical education. *Western Journal of Emergency Medicine*, *20*(1), p.78.

29. Kastrati, Z., Dalipi, F., Imran, A.S., Pireva Nuci, K. and Wani, M.A., 2021. Sentiment analysis of students’

feedback with NLP and deep learning: A systematic mapping study. *Applied Sciences*, *11*(9), p.3986.

30. Khan, M.T., Durrani, M., Ali, A., Inayat, I., Khalid, S. and Khan, K.H., 2016. Sentiment analysis and the complex natural language. *Complex Adaptive Systems Modeling*, *4*, pp.1-19.

31. Khurana, D., Koli, A., Khatter, K. and Singh, S., 2023. Natural language processing: State of the art, current trends and challenges. *Multimedia tools and applications*, *82*(3), pp.3713-3744.

32. Kocaleva, M., Stojanov, D., Stojanovic, I. and Zdravev, Z., 2016. Pattern recognition and natural language processing: State of the art. *Tem Journal*, *5*(2), pp.236-240.

33. Kreimeyer, K., Foster, M., Pandey, A., Arya, N., Halford, G., Jones, S.F., Forshee, R., Walderhaug, M. and Botsis, T., 2017. Natural language processing systems for capturing and standardizing unstructured clinical information: a systematic review. *Journal of biomedical informatics*, *73*, pp.14-29.

34. Conneau, A. and Lample, G., 2019. Cross-lingual language model pretraining. *Advances in neural information processing systems*, *32*.

35. Lavanya, P.M. and Sasikala, E., 2021, May. Deep learning techniques on text classification using Natural language processing (NLP) in social healthcare network: A comprehensive survey. In *2021 3rd International Conference on Signal Processing and Communication (ICPSC)* (pp. 603-609). IEEE.

36. Le Glaz, A., Haralambous, Y., Kim-Dufor, D.H., Lenca, P., Billot, R., Ryan, T.C., Marsh, J., Devylder, J., Walter, M., Berrouiguet, S. and Lemey, C., 2021. Machine learning and natural language processing in mental health: systematic review. *Journal of Medical Internet Research*, *23*(5), p.e15708.

37. Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 5(1), 1-167.

38. Chen, N., Lin, J., Hoi, S.C., Xiao, X. and Zhang, B., 2014, May. AR-miner: mining informative reviews for developers from mobile app marketplace. In *Proceedings of the 36th international conference on software engineering* (pp. 767-778).

39. Mathew, L. and Bindu, V.R., 2020, March. A review of natural language processing techniques for sentiment analysis using pre-trained models. In *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 340-345). IEEE.

40. Mathews, S.M., 2019. Explainable artificial intelligence applications in NLP, biomedical, and malware classification: a literature review. In *Intelligent Computing: Proceedings of the 2019 Computing Conference, Volume 2* (pp. 1269-1292). Springer International Publishing.

41. Mulligan, S.P., Freeman, W.C. and Linebaugh, C.D., 2019. Data protection law: An overview. *Congressional Research Service*, *45631*, p.25.

42. Sharma, A., Chakraborty, S. and Kumar, S., 2022. Named entity recognition in natural language processing: A systematic review. In

Proceedings of Second Doctoral Symposium on Computational Intelligence: DoSCI 2021 (pp. 817-828). Springer Singapore.

43. O'Loughlin, K., Neary, M., Adkins, E.C. and Schueller, S.M., 2019. Reviewing the data security and privacy policies of mobile apps for depression. *Internet interventions*, *15*, pp.110-115.

44. Paolanti, M. and Frontoni, E., 2020. Multidisciplinary pattern recognition applications: A review. *Computer Science Review*, *37*, p.100276.

45. Qiu, X., Sun, T., Xu, Y., Shao, Y., Dai, N. and Huang, X., 2020. Pre-trained models for natural language processing: A survey. *Science China Technological Sciences*, *63*(10), pp.1872-1897.

46. Rahman, M.R., Mahdavi-Hezaveh, R. and Williams, L., 2020, November. A literature review on mining cyberthreat intelligence from unstructured texts. In *2020 International Conference on Data Mining Workshops (ICDMW)* (pp. 516-525). IEEE.

47. Jabangwe, R., Edison, H. and Duc, A.N., 2018. Software engineering process models for mobile app development: A systematic literature review. *Journal of Systems and Software*, *145*, pp.98-111.

48. Rehman, I.U., Sobnath, D., Nasralla, M.M., Winnett, M., Anwar, A., Asif, W. and Sherazi, H.H.R., 2021. Features of mobile apps for people with autism in a post COVID-19 scenario: Current status and recommendations for apps using AI. *Diagnostics*, *11*(10), p.1923.

49. Roy, P.K., Saumya, S., Singh, J.P., Banerjee, S. and Gutub, A., 2022. Analysis of community question-answering issues via machine learning and deep learning: State-of-the-art review. *CAAI Transactions on Intelligence Technology*.

50. Salloum, S., Gaber, T., Vadera, S. and Sharan, K., 2022. A systematic literature review on phishing email detection using natural language processing techniques. *IEEE Access*.

51. Santos, R., Groen, E.C. and Villela, K., 2019, March. An Overview of User Feedback Classification Approaches. In *REFSQ Workshops* (Vol. 3, pp. 357 -369).

52. Sarker, I.H., Furhad, M.H. and Nowrozy, R., 2021. Ai-driven cybersecurity: an overview, security intelligence modeling and research directions. *SN Computer Science*, *2*, pp.1-18.

53. Sarker, I.H., Hoque, M.M., Uddin, M.K. and Alsanoosy, T., 2021. Mobile data science and intelligent apps: concepts, AI-based modeling and research directions. *Mobile Networks and Applications*, *26*, pp.285-303.

54. Scherm, H., Thomas, C.S., Garrett, K.A. and Olsen, J.M., 2014. Meta-analysis and other approaches for synthesizing structured and unstructured data in plant pathology. *Annual Review of Phytopathology*, *52*, pp.453-476.

55. Shaik, T., Tao, X., Li, Y., Dann, C., McDonald, J., Redmond, P. and Galligan, L., 2022. A review of the trends and challenges in adopting natural language processing methods for education feedback analysis. *IEEE Access*.

56. *Sharir, O., Peleg, B. and Shoham, Y., 2020. The cost of training nlp models: A concise overview. arXiv preprint arXiv:2004.08900.*

57. Shivaprasad, T.K. and Shetty, J., 2017, March. Sentiment analysis of product reviews: a review. In *2017 International conference on inventive communication and computational technologies (ICICCT)* (pp. 298-301). IEEE.

58. Sousa, S. and Kern, R., 2022. How to keep text private? A systematic review of deep learning methods for privacy-preserving natural language processing. *Artificial Intelligence Review*, pp.1-66.

59. Srisopha, K., Phonsom, C., Li, M., Link, D. and Boehm, B., 2020, June. On building an automatic identification of country-specific feature requests in mobile app reviews: Possibilities and challenges. In *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops* (pp. 494-498).

60. *Sun, T., Gaut, A., Tang, S., Huang, Y., ElSherief, M., Zhao, J., Mirza, D., Belding, E., Chang, K.W. and Wang, W.Y., 2019. Mitigating gender bias in natural language processing: Literature review. arXiv preprint arXiv:1906.08976.*

61. Sun, S., Luo, C. and Chen, J., 2017. A review of natural language processing techniques for opinion mining systems. *Information fusion*, *36*, pp.10-25.

62. Sworna, Z.T., Mousavi, Z. and Babar, M.A., 2022. NLP methods in host-based intrusion detection Systems: A systematic review and future directions. *arXiv preprint arXiv:2201.08066*.

63. Reyes-Menendez, A., Saura, J.R. and Alvarez-Alonso, C., 2018. Understanding# WorldEnvironmentDay user opinions in Twitter: A topic-based sentiment analysis approach. *International journal of environmental research and public health*, *15*(11), p.2537.

64. Tavakoli, M., Zhao, L., Heydari, A. and Nenadić, G., 2018. Extracting useful software development information from mobile application reviews: A survey of intelligent mining techniques and tools. *Expert Systems with Applications*, *113*, pp.186-199.

65. Thakur, K., Qiu, M., Gai, K. and Ali, M.L., 2015, November. An investigation on cyber security threats and security models. In *2015 IEEE 2nd international conference on cyber security and cloud computing* (pp. 307-311). IEEE.

66. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I., 2017. Attention is all you need. *Advances in neural information processing systems*, *30*.

67. Wylde, V., Rawindaran, N., Lawrence, J., Balasubramanian, R., Prakash, E., Jayal, A., Khan, I., Hewage, C. and Platts, J., 2022. Cybersecurity, data privacy and blockchain: A review. *SN Computer Science*, *3*(2), p.127.

68. Zhang, X., Zhao, J. and LeCun, Y., 2015. Character-level convolutional networks for text classification. *Advances in neural information processing systems*, *28*.