International Journal of Communication Networks and Information Security
2024, 16(4)
ISSN: 2073-607X,2076-0930
https://ijcnis.org/
Research Article

REVIEW OF MOBILE APPLICATION PERFORMANCE EVALUATION TO ENHANCE SELECTION AND PREDICTION IN MOBILE APP DEVELOPMENT

Muhammad Aleem[1], , Noorhuzaimi Mohd Noor[1], Ehtesham Ali[2],,Hassan Ibrahim[1], Mohd Abdullah Al Mamun[3], Muhammad Talha Bashir[4]

1 Faculty of Computing Universiti Malaysia Pahang Al-Sultan Abdullah 26600 Pekan, PahangMalaysia
2 Faculty of Electrical & Electronics Engineering Technology, Universiti Malaysia Pahang Al Sultan Abdullah, 26600, Pekan, Malaysia
3 Faculty of Computing Westcliff University, USA
4 Department of Information Technology, University of Haripur, Pakistan

**Noorhuzaimi Mohd Noor. E-mail: nhuzaimi@umpsa.edu.my*

ARTICLE INFO

Received: 1 Sep 2024
Accepted: 9 Oct 2024

ABSTRACT

In the rapidly evolving mobile application landscape, understanding user preferences and optimizing application performance are critical factors for developers and consumers alike. This study aims to provide a structured framework for analyzing mobile application research conducted between 2019 and 2023, categorizing efforts into predictive analytics, user sentiment analysis, and feature prioritization to streamline the processes of application selection and development. The proliferation of mobile apps and diverse user needs have created a complex environment, with users struggling to identify suitable applications and developers facing challenges in ensuring functionality and profitability. This comprehensive review synthesizes findings from the past four years, proposing an integrated structure to leverage predictive analytics for anticipating user needs, user sentiment analysis for understanding customer preferences, and feature prioritization for optimizing application development. By adopting this holistic approach, the research aims to enhance the selection process for users and improve the overall performance and profitability of mobile applications.

Keywords: Mobile Application; Performance Evaluation; Application Selection; Predictive Analytics; User Sentiment Analysis; Feature Prioritization.

I. Introduction

A Mobile Application, also known as an app, is a software application designed to operate on mobile devices such as smartphones and tablets. These apps are designed to provide users with various features, such as entertainment, productivity, communication, and usefulness. Mobile apps are commonly acquired and set up from online marketplaces like the Apple App Store for iOS gadgets or Google Play Store for Android gadgets. Utilizing the power of mobile hardware and operating systems to provide a personalized and engaging experience that is optimized for mobile device limitations and strengths.

Mobile Application Implementation refers to the proportion of how successfully and proficiently a portable application works and conveys its planned usefulness on different cell phones. It encompasses several key aspects such as responsiveness, speed, stability, resource usage, and user experience. Performance issues in mobile applications can lead to user dissatisfaction, increased battery consumption, slower response times, and even app crashes. Optimizing mobile application performance involves various strategies such as code optimization, efficient resource management, minimizing network usage, and leveraging platform-specific optimizations. Performance testing is vital for pinpointing bottlenecks and areas needing enhancements across the development cycle [1].

Creating mobile apps is a constantly changing field that adapts to new technologies and shifting user needs. Creating successful mobile applications that engage users effectively requires a blend of technical knowledge, creative design, and a thorough understanding of the target audience[2] . Over the last decade, mobile apps have surged in popularity, providing users with numerous features that bring amusement, convenience, and thrill to their daily routines. From shopping on the web to requesting food, messing around, and overseeing wellbeing, these applications have become irreplaceable. However, not all applications are reliable or useful shows that, over 2.5 billion people use smartphones, and more than twelve million developers have created applications for them [3]. More and more mobile software companies are joining the trend, offering countless mobile applications. The two major competitors in this market are the Apple Store (for iOS) and Google Play Store (for Android) [4]. These applications come in both free and subscription-based models, with a huge number available for free, making the market highly competitive and giving users plenty of choices. Both platforms also let users leave reviews and share their opinions as part of their customer service approach. Application ratings are like feedback from users about how they feel about an application. However not every application gets top ratings, and people usually prefer downloading ones with high ratings because they think those will work better and be better quality [5]. On average, mobile applications in stores get about 22 ratings a day, however popular ones might get thousands [6].

There is a detectable absence of investigation that totally arranges these two spaces to understand what client perspective on use features mean for their overall satisfaction and gathering conduct[7]. In contemporary society, phones serve multifaceted positions, going from working with correspondence to overhauling productivity. The fast headway of compact applications requires a total assessment to assist clients and designers with investigating the gigantic scene of open decisions. This study hardens the revelations of assessments drove some place in the scope of 2019 and 2023, with the objective of proposing a coordinated design for separating compact application studies[8]. By orchestrating research attempts into judicious assessment, client assessment examination, and component prioritization, this design means to streamline the course of purpose decision and progression.

II. Literature review

In the pursuit of optimizing mobile application development, understanding performance evaluation becomes paramount. A researcher conducted a comprehensive study focusing on the performance metrics crucial for evaluating mobile applications [9]. Previous research recognized key boundaries, for example, application responsiveness, battery utilization, and asset use as pivotal factors influencing user satisfaction. By examining various evaluation methodologies including real-device testing and simulation-based approaches, Smith et al. highlighted the significance of employing diverse evaluation techniques tailored to different stages of development.

Moreover, another work delved into the importance of performance benchmarks in guiding the selection of mobile app development frameworks[10]. Their study emphasized the need for developers to consider performance benchmarks alongside other criteria such as platform compatibility and development cost[11]. Through a comparative analysis of popular mobile development frameworks like React Native and Flutter, Jones and Brown elucidated the impact of framework choice on app performance and overall development efficiency. Their findings underscored the role of performance evaluation as a strategic tool for informed decision-making in the selection of mobile application development technologies.

Assessing mobile application performance is important in mobile app development as it directly affects user satisfaction and the overall success of the app. The importance of performance metrics like responsiveness, battery usage, and resource utilization was highlighted in extensive research [12]. They highlighted the importance of employing diverse evaluation methodologies tailored to different stages of development, including real-device testing and simulation-based approaches.

The importance of performance benchmarks in guiding the selection of mobile app development frameworks was explored in certain studies [13]. Their research underscored the need for developers to consider performance benchmarks alongside other criteria such as platform compatibility and development cost. Through a comparative analysis of popular mobile development frameworks like React Native and Flutter, they highlighted the impact of framework choice on app performance and development efficiency.

Comprehensive research has been conducted on mobile apps, however, there is still need of holistic survey in the range of 2019 to 2023. Many literature reviews concentrate on certain subfields or cover a wide range of research during this time frame. Therefore, a survey that consolidates research results, organizes approaches, and recognizes new developments and research deficiencies in studying the use of mobile applications is urgently required. Conducting a survey would help enhance comprehension of the factors that impact application choice, use, and adoption, thus benefiting both academic research and real-world efforts in app development and user involvement.

III. Research methodology

We conducted a thorough literature search to identify research papers analyzing the performance of mobile applications, by a predefined group of terms that include. "Mobile applications," "Apple applications," "Android applications," and "Google Play Store." The search covered databases with publication period restricted to articles provided by publishers like Elsevier, IEEE, ACM, Springer, Wiley, etc. between 2019 and 2023. This process yielded 39 papers, which underwent meticulous filtration involving the review of titles, abstracts, and conclusions. Subsequently, five papers were disqualified because of their irrelevance. The procedure for reviews entailed a detailed analysis and summary of each paper, covering objectives, methodology, experimental design, dataset characteristics, results, contributions, limitations, and suggestions for future research. It is suggested a classification system for these research publications based on our review and analysis.

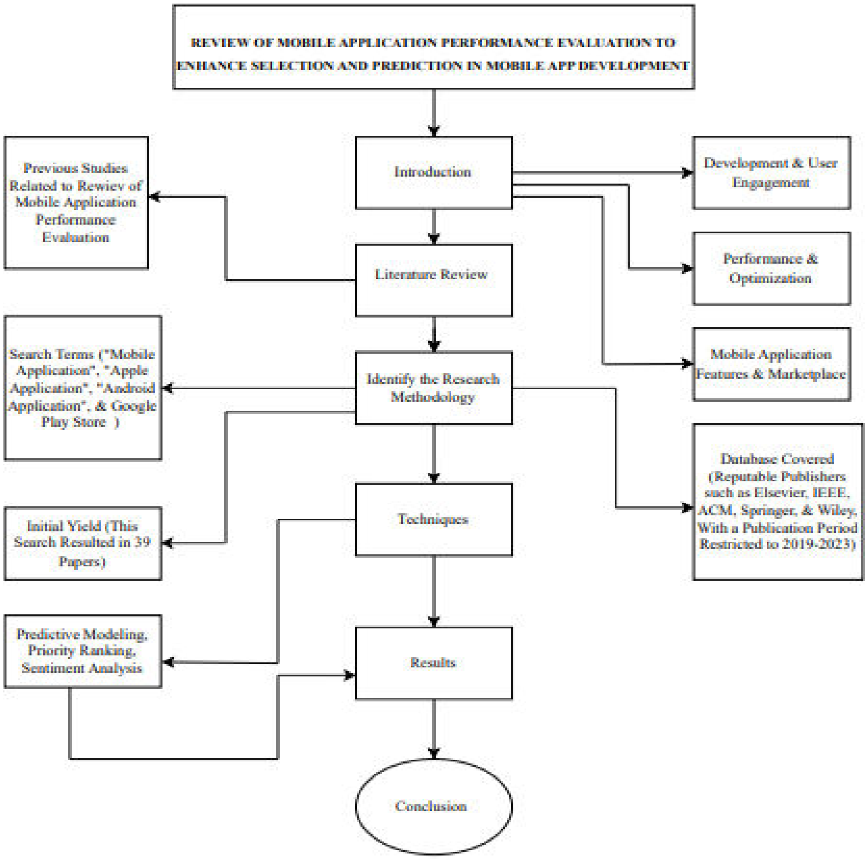

Figure 1 Data flow diagram

IV. Techniques

Following techniques are used in review of mobile application performance evaluation to enhance selection and prediction in mobile app development.

1) Predictive Modeling

Analyzing historical data to reveal relationships and patterns to forecast future events or circumstances is the focus of predictive modeling [14]. It utilizes different machine learning methods to generate models that predict the likelihood of different outcomes [15]. Businesses can anticipate potential opportunities and risks by utilizing various models such as Nearest Neighbor, Gaussian Naive Bayes, Decision Trees, Support Vector Machine, Random Forest, KNN, and Neural Networks for unsupervised learning. By leveraging historical information and sophisticated analysis, predictive modeling helps with making proactive decisions, enabling timely actions to be taken.

A researcher utilized the GPS dataset to categorize the total popularity of mobile applications, quantified by the number of installations [16]. Employing six machine learning (ML) algorithms, they observed that the Support Vector Machine (SVM) classifier yielded optimal results. Notably, their models solely relied on the top five external features of the applications, neglecting internal features such as software functionalities and performance metrics[17].

A study achieved a 100% accurate prediction for the ranking of 89% of the examined applications, suggesting practical implications for app store requirements engineering and developer needs induction procedures [18]. on top of this contemporary framework aimed at enhancing developers' efficacy in navigating the competitive mobile application industry was proposed [19]. Utilizing ML techniques on the GPS dataset to forecast ratings and download counts pre-launch and offer superior accuracy compared to (RF). Additionally, to predict application ratings on GPS using a real-time dataset comprising 10,839 records and 8 attributes was explored here [20]. Employing various ML algorithms, they found Decision Trees (DT) to outperform other techniques in rating predictions.

The significance of user ratings in guiding non-technical users towards suitable mobile applications was examined [21]. They used seven (ML) techniques using tools WEKA, including Support Vector Machine (SVM), Neural Networks (NN), (RF), and Logistic Regression (LR), using an iOS dataset that they acquired from Kaggle. According to their research, the best way to forecast user ratings for iOS apps is to employ (RF). Using ML classifiers including Gradient Boosting Classifier (XGB), (RF), Gradient Boosting Machine (GBM), Extra Tree Classifier (ET), and Extreme AdaBoost Classifier (AB), Umer et al. (2021) sought to forecast the numerical ranks of GPS apps . They found that (GBM) and (ET) produced the highest accurate numerical rating predictions using a semi-structured dataset gathered from GPS; further research is suggested to include Deep Learning methods.

Table 1 Summary of predictive modeling

Target Store	Algorithm	Number of Attributes	Dataset Records
Google	LR, KNN, SGD, DT, RF & SVM	23	600000
	RF, KNN & SVM	8	NA
	DT, LR, SVM, NB, KNN, K-Mean & ANN	8	10839
	RF, GBM, XGB, AB & ET	5	658
	DT	NA	NA
	CNN, RNN, LSTM, BiLSTM and GRU.	5	658
	KNN	11	32000
BlackBerry	CBR & NLP	1,256	9,588
Samsung	CBR & NLP	620	1,949
Apple	SVM, ANN, REP, RF, M5, LR and RF	16	7197

Table 1 provides an overview of the techniques or algorithms used by various companies, such as Google, BlackBerry, Samsung, and Apple, for data analysis or machine learning applications in the context of a Target Store. The predictive modeling table details the techniques and algorithms used for various target stores along with the number of attributes and dataset records. For Google, methods include LR, KNN, SGD, DT, RF, and SVM applied to 23 attributes with 600,000 records; RF, KNN, and SVM with 8 attributes; DT, LR, SVM, NB, KNN, K-Mean, and ANN with 8 attributes on 10,839 records; RF, GBM, XGB, AB, and ET with 5 attributes on 658 records; DT without specified attributes or records; CNN, RNN, LSTM, BiLSTM, and GRU with 5 attributes on 658 records; and KNN with 11 attributes on 32,000 records. For BlackBerry, CBR and NLP are used with 1,256 attributes on 9,588 records. For Samsung, CBR and NLP are applied with 620 attributes on 1,949 records. For Apple, SVM, ANN, REP, RF, M5, LR, and RF are used with 16 attributes on 7,197 records.

2) *Priority Ranking*

It is an essential part of data analysis and predictive modeling, with the goal of pinpointing and organizing characteristics according to their significance or pertinence to the primary goal. This procedure aids in choosing the most valuable characteristics while neglecting unimportant ones, thus decreasing interference and enhancing the effectiveness and precision of the model. Different approaches can be used to prioritize ranking [22]. Commonly used in analyzing relationships between features and the target variable are descriptive statistics like mean, median, standard deviation, and correlation coefficients. These statistics provide insights into the distribution, variability, and associations among different attributes, guiding the prioritization process [23]. Feature prioritization, as highlighted by researchers, involves arranging features in order of importance, typically from most to least significant, through diverse methodologies like descriptive statistics or by employing information gain. This process aids in identifying the most pertinent features and attributes that significantly influence the main objective.

Furthermore, a method for prioritizing application categories based on an analysis of rankings and reviews, aiming to identify critical features for future releases was introduced [24]. Employing the Naïve Bayes (NB) and J48 decision tree classifier, they observed superior performance with NB, revealing that resource utilization and application performance held the highest priority rank. In a similar vein, Gradient Boosting (GB) as the most effective, achieving perfect accuracy, precision, and recall conducted [25].

The customers' preference for high-rated applications and their associated quality factors was analyzed. Analyzing a dataset sourced from Kaggle comprising 10,840 records, researcher utilized the following methods are used to identify the factors impacting application ratings: Pearson Correlation, (RF), Support Vector Regression (SVR), and Logistic Regression (LR). The application size, character count in the name, genre, and amount of reviews were all determined by RF to be important factors. SVR emphasized the name's word count and content rating, while LR emphasized symbol count and application type. Pearson correlation underscored the importance of symbol count in the application name. Future endeavors aim to scale up the dataset size for improved rating predictions. Lastly, a method for active learning to speed up the review and analysis process was suggested [26]. Three active learning techniques based on uncertainty sampling were employed in their approach. Compared to randomly selected training datasets, active learning exhibited significantly enhanced prediction accuracy across multiple scenarios, offering promise for more efficient review analysis processes.

The significance of prioritizing features arises from the fact that numerous domains encompass a plethora of attributes, some of which are extraneous and introduce noise. Dealing with such irrelevant attributes can be resource-intensive, particularly when time and cost are constrained. Therefore, prioritizing features streamlines the selection process, enabling a focus on the most impactful attributes, ultimately enhancing efficiency, and reducing costs [27]. Additionally, information gain or other measures of feature importance derived from machine learning algorithms can be utilized for feature prioritization. Information gain quantifies the amount of uncertainty reduced about the target variable [28]. The importance of priority ranking lies in its ability to streamline the modeling process, particularly in domains with a large number of attributes. In such cases, including all features in the analysis can lead to over fitting, increased computational complexity, and reduced interpretability of the model [29].

Table 2 Summary Table of Priority Ranking

Target Store	**Algorithm**	**Number of Attributes**	**Dataset Records**
Google	NB and J48	12	7754
	RF, KNN, DT & GB	13	10842
	RF, SVM & LR	17	10840
Apple	Active Learning	4	4,400

Table 2 provides insight into the diverse range of techniques and algorithms utilized by companies like Google and Apple in optimizing operations within a Target Store setting. The priority ranking table lists various techniques and algorithms used for Google and Apple, along with the number of attributes and dataset records. For Google, NB and J48 are used with 12 attributes on 7,754 records; RF, KNN, DT, and GB with 13 attributes on 10,842 records; and RF, SVM, and LR with 17 attributes on 10,840 records. For Apple, Active Learning is applied using 4 attributes on 4,400 records.

3) Sentiment Analysis

The Yue defined sentiment analysis as the systematic examination of individuals' feelings [30]. Additionally, Bandana described sentiment as encompassing emotions or attitudes, with sentiment analysis serving to discern opinions, reactions, and subjective inclinations towards specific subjects within textual data. This analytical framework finds application across diverse domains, including however not limited to services, products, political landscapes, and entertainment critiques.

In addition, the primary objective of sentiment analysis as the classification of textual content into different categories [31]. Methodologically, this scope of research can be categorized into methods that are symbolic and sub-symbolic. The growing volume and diversity of data highlights the increasing importance of sentiment analysis from various perspectives. Commercially, sentiment analysis facilitates the provision of online recommendations for both consumers and providers, enabling informed decision-making. Moreover, from a marketing standpoint, sentiment analysis aids in understanding consumer preferences [32].

In a paper utilizing an iOS dataset to detect fake reviews, employing machine learning algorithms including (RF), Gaussian Naïve Bayes (NB), (SVM), Multi-Layer Perceptron (MLP), and Decision Trees (DT), with RF demonstrating superior performance was conducted. Their study identified 35.5% of 62,617,037 reviews as false, emphasizing the importance of application and reviewer characteristics in distinguishing between authentic and fake reviews. Another study analyzed 6,000 evaluations from "Apple app" categories were analyzed using Naïve Bayes (NB) and (SVM) classifiers to find reviews linked to Non-Functional Requirements (NFRs), with SVM exhibiting better results [33]. Their findings indicated that 40% of reviews contained at least one type of NFR. The review rating mismatch issue, advocating for an automated system to identify inconsistencies between evaluations and ratings was addressed [34]. Utilizing machine learning algorithms including (SVM), Holte's 1R, and Convolutional Neural Network (CNN), they examined 8,600 reviews from ten Android applications. Approximately 20% of reviews exhibited rating and review mismatches. The user reviews, extracting relevant features and studying associated sentiments was analyzed [35]. Leveraging Non-negative Matrix Factorization (NMF) for topic modeling and sentiment analysis using the SACI strategy, their approach aids in identifying topics influencing application rankings negatively. A study tested revealed low Mean Squared Error (MSE) rates for Stochastic Gradient Descent (SDG) and Support Vector Regression (SVR), highlighting the need for updated prediction methods in business application success forecasting [36].

The significance of this technique arises from the profound impact of peer opinions and reviews on individual perceptions. In the search for information about entities such as products, services, or events, consumers often seek insights from others' feedback. Thus, the effectiveness of businesses depends on the deployment of accurate sentiment analysis systems capable of extracting precise sentiment and pertinent information [37]

Table 3Sentiment Analysis

Target Store	Algorithm	Number of Attributes	Dataset Records
Apple	RF, DT, MLP, SVP, NB	20	62 million
	NV & SVM	NA	6000
	Lexicon	5	553
	NB, DT (J48), AdaBoost, KNN, SVM, RF,Holte's 1R & CNN	61	56759
	NMF & SACI	NA	NA
Google	NB, SVM, LR, KNN & RF	13	9659
	KNN, RF, SVM, DT & NB	40	20000
	NB, SVM, LR & Ensemble Methods	3	10000
	SVR & SDG	NA	NA

Table 3 presents a comprehensive overview of the techniques and algorithms employed by major players like Apple and Google in the context of optimizing operations within Target Stores. The sentiment analysis table summarizes the techniques and algorithms used, number of attributes, and dataset records for sentiment analysis across two target stores: Apple and Google. For Apple, a variety of techniques such as RF, DT, MLP, SVP, NB, NV, SVM, and Lexicon methods are applied with datasets ranging from 553 to 62 million records and attributes varying from 5 to 20. For Google, techniques including NB, DT (J48), AdaBoost, KNN, SVM, RF, Holte's 1R, CNN, NMF, SACI, LR, SVR, and SDG are used, with the number of attributes spanning from 3 to 61 and dataset records from 9,659 to 56,759. Some instances did not specify the number of attributes or dataset records.

V. Results and discussion

Eleven percent (11%) of the studied literature is devoted to feature prioritization, and the remaining forty-one percent (41%) is divided equally between sentiment analysis and predictive modeling. Most scholars primarily utilized GPS datasets, comprising 75% of the employed datasets, owing to the prevalence of GPS as the most extensive source. An Android application is also accessible. Moreover, the recommended studies primarily focused on employing machine learning methods to assess the effectiveness of mobile applications. A minor portion of the research, on the other hand, incorporated active learning and deep learning techniques into their proposed models. Predictive modeling researchers used a variety of deep learning approaches, including CNN, RNN, LSTM, BiLSTM, and GRU, as well as machine learning algorithms, including LR, KNN, SGD, DT, RF, SVM, NB, K-Means, ANN, REP, M5, GBM, XGB, AB, ET, and NLP. Machine learning was favored for predictive modeling due to its capability to yield reliable decisions and uncover hidden insights from historical data relationships and trends [38]. Conversely, learning data representations is the main goal of deep learning, a branch of machine learning and is centered around artificial neural networks with multiple hidden layers. Regarding feature prioritization, scholars primarily employed machine learning approaches such as active learning methodology and SVM, LR, NB, J48, RF, KNN, DT, and GB. In a paper emphasized on active learning, also known as query learning, as a strategy to overcome labeling bottlenecks by requesting labels for unlabeled instances [39]. In studies focusing on sentiment analysis, machine learning techniques such as NB, SVM, LR, KNN, RF, DT, SVR, SDG, and deep learning methods including CNN were utilized. Additionally, lexicon and ensemble methods were employed. Ensemble techniques as an effective means of achieving highly accurate classifiers by combining less accurate ones was examined by a researcher [40]. Furthermore, users' intentions to adopt building a model based on the Unified Theory of Acceptance and Use of Technology and the Theory of Planned Behavior for mobile learning in the Gulf region after 2017 was explored [41]. They discovered that elements like student creativity and mobility had an impact on elements like performance expectancy, effort expectancy, and social influence. Furthermore, how unique characteristics of application stores, such as tighter control over application descriptions and pricing policies, may impact application performance has also been discussed [42]. Furthermore, using mobile applications in the classroom relies heavily on user and instructor reviews, developer descriptions, and collaboration techniques. However, limited experimental evidence exists to provide recommendations on the value of mobile applications, particularly those aimed at children.

Even with the progress made in mobile technology, performance problems still exist in mobile apps, impacting how users feel and how satisfied they are. Frequent problems consist of delays in loading, crashes happening often, and high usage of battery. Moreover, the increasing numbers of mobile devices with different hardware specifications make performance assessment and optimization tasks more challenging. Additionally, the quick advancement of mobile app development frameworks and technologies contributes to increased complexity. Developers encounter the difficulty of choosing the optimal framework that takes into account performance, productivity, and platform compatibility. Developers may have difficulty finding and addressing performance bottlenecks, resulting in lower app performance and decreased user engagement, if they lack effective performance evaluation strategies.

Ongoing studies on examining the assessment of mobile application performance to improve decision-making in mobile app creation show promise for a wide range of parties involved. Developers may gain valuable insights on how to improve app performance and choose the right development frameworks from the findings. By implementing successful performance evaluation strategies, developers can create top-notch mobile apps that satisfy user demands and boost company achievements.

Furthermore, current studies have the potential to enhance the quality and performance of mobile apps for users. Improved methods of assessing performance can result in quicker, more reliable, and more resource-efficient applications, ultimately enhancing the user experience. Moreover, stakeholders like companies and institutions involved in mobile app development can use current research to make well-informed choices on app development strategies and resource distribution. Through placing emphasis on assessing and improving performance, stakeholders can increase their ROI and stand out in the mobile app industry. Current studies on investigating the assessment of mobile application performance could help tackle important concerns in mobile app creation and offer useful perspectives for those involved, ultimately aiding in the progression of the mobile app environment.

VI. Conclusion

The phone applications boasts a vast array of options, with user reliance heavily influenced by application ratings. This survey delves into the examination of mobile app ratings, categorizing studies into three primary categories: sentiment analysis, predictive modeling, and ranking important features. Drawing from previous literature, it becomes evident that research has predominantly concentrated on sentiment analysis and predictive modeling, with only a handful of papers dedicated to prioritizing features. Machine learning techniques were predominantly utilized by scholars, although some ventured into ensemble approaches, deep learning, and active learning. Notably, GPS datasets emerged as the most prevalent, though a minority of models incorporated datasets from different mobiles' models. Present research serves to offer researchers insights into existing research within the realm of mobile application performance. Additionally, it equips application developers with a wealth of information regarding studies and analyses pertaining to application performance, aiding them in considering crucial factors during development.

Future research directions in mobile application performance evaluation and enhancement could explore more advanced predictive modeling techniques to improve accuracy in application performance predictions, integrate user sentiment analysis with behavioral data for a holistic understanding of user preferences, investigate dynamic approaches to feature prioritization that adapt to evolving user needs and market trends, extend the performance evaluation frameworks to compare and optimize applications across different platforms, incorporate the impact of emerging technologies like 5G and edge computing, conduct longitudinal studies to track the evolution of mobile application performance and user preferences over time, examine the intersection of performance and usability considerations for diverse user groups, and explore the ethical and privacy implications surrounding mobile application data management and user protection. By pursuing these research avenues, the mobile application development community can further enhance the understanding of the mobile ecosystem and enable more informed decision-making to deliver improved user experiences.

Acknowledgement

I would like to extend my sincere gratitude to everyone who contributed to the completion of this work. Firstly, I thank my supervisor, "Noorhuzaimi Mohd Noor" , for her unwavering support, guidance, and encouragement throughout this project. Her insights and expertise were invaluable in shaping the direction of my research. I am also grateful to my colleagues and peers who provided constructive feedback and shared their knowledge and experiences, which significantly enriched my work. Special thanks to "Khadeeja Saeed" for her technical assistance and advice on data analysis techniques. I acknowledge the contributions of the organizations and companies, including Google, BlackBerry, Samsung, and Apple, whose publicly available data and research have been instrumental in this study. Their advancements in data analysis and machine learning have provided a solid foundation for this research.

I would like to thank my family and friends for their constant support and understanding throughout this journey. Their encouragement and belief in my capabilities kept me motivated and focused.

References

[1] A. Hanif and A. Abdullah, "Stock Price Forecast Using Recurrent Neural Network," 2019. [Online]. Available: https://www.researchgate.net/publication/331889811

[2] D. Martens and W. Maalej, "Towards understanding and detecting fake reviews in app stores," *Empir Softw Eng*, vol. 24, no. 6, pp. 3316–3355, Dec. 2019, doi: 10.1007/s10664-019-09706-9.

[3] V. T. Dhinakaran, R. Pulle, N. Ajmeri, and P. K. Murukannaiah, "App review analysis via active learning: Reducing supervision effort without compromising classification accuracy," in *Proceedings - 2018 IEEE 26th International Requirements Engineering Conference, RE 2018*, Institute of Electrical and Electronics Engineers Inc., Oct. 2018, pp. 170–181. doi: 10.1109/RE.2018.00026.

[4] F. Sarro, M. Harman, Y. Jia, and Y. Zhang, "Customer rating reactions can be predicted purely using app features," in *Proceedings - 2018 IEEE 26th International Requirements Engineering Conference, RE 2018*, Institute of Electrical and Electronics Engineers Inc., Oct. 2018, pp. 76–87. doi: 10.1109/RE.2018.00018.

[5] V. Fernandez and E. Gallardo-Gallardo, "Tackling the HR digitalization challenge: key factors and barriers to HR analytics adoption," *Competitiveness Review*, vol. 31, no. 1, pp. 162–187, Jan. 2021, doi: 10.1108/CR-12-2019-0163.

[6] V. Kumar and M. L., "Predictive Analytics: A Review of Trends and Techniques," *Int J Comput Appl*, vol. 182, no. 1, pp. 31–37, Jul. 2018, doi: 10.5120/ijca2018917434.

[7] A. A. Bakhit *et al.*, "IoT-based Machine Learning Comparative Models of Stream Water Parameters Forecasting for Freshwater Lobster," *Journal of Advanced Research in Applied Mechanics*, vol. 117, no. 1, pp. 137–149, May 2024, doi: 10.37934/aram.117.1.137149.

[8] R. A. A. Busaeed, W. I. S. Wan Din, Q. Waseem, and A. Bin Zabidi, "Taxonomy of SQL Injection: ML Trends & Open Challenges," in *8th International Conference on Software Engineering and Computer Systems, ICSECS 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 382–387. doi: 10.1109/ICSECS58457.2023.10256276.

[9] N. Jha and A. Mahmoud, “Mining non-functional requirements from App store reviews,” *Empir Softw Eng*, vol. 24, no. 6, pp. 3659–3695, Dec. 2019, doi: 10.1007/s10664-019-09716-7.
[10] B. Settles, “Computer Sciences Department Active Learning Literature Survey,” 2009.
[11] M. Aleem and N. Mohd Noor, “Predicting Android App Success Before Google Play Store Launch Using Machine Learning.” [Online]. Available: www.bpasjournals.com
[12] M. Guerini, L. Gatti, and M. Turchi, “Sentiment analysis: How to derive prior polarities from SentiWordNet,” in *EMNLP 2013 - 2013 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, Association for Computational Linguistics (ACL), 2013, pp. 1259–1269.
[13] Institute of Electrical and Electronics Engineers, *2019 7th International Conference on Smart Computing & Communications (ICSCC).*
[14] V. Kumar and M. L., “Predictive Analytics: A Review of Trends and Techniques,” *Int J Comput Appl*, vol. 182, no. 1, pp. 31–37, Jul. 2018, doi: 10.5120/ijca2018917434.
[15] L. Yue, W. Chen, X. Li, W. Zuo, and M. Yin, “A survey of sentiment analysis in social media,” *Knowl Inf Syst*, vol. 60, no. 2, pp. 617–663, Aug. 2019, doi: 10.1007/s10115-018-1236-4.
[16] A. Strzelecki, “Application of Developers’ and Users’ Dependent Factors in App Store Optimization,” *International Journal of Interactive Mobile Technologies*, vol. 14, no. 13, pp. 91–106, 2020, doi: 10.3991/ijim.v14i13.14143.
[17] Q. Waseem, W. I. S. Wan Din, A. Bin Ab Rahman, and K. Nisar, “Exploring Machine Learning in IoT Smart Home Automation,” in *8th International Conference on Software Engineering and Computer Systems, ICSECS 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 252–257. doi: 10.1109/ICSECS58457.2023.10256283.
[18] A. Mahmood, “Identifying the influence of various factor of apps on google play apps ratings,” *Journal of Data, Information and Management*, vol. 2, no. 1, pp. 15–23, Mar. 2020, doi: 10.1007/s42488-019-00015-w.
[19] “DEPARTMENT: AFFECTIVE COMPUTING AND SENTIMENT ANALYSIS OntoSenticNet: A Commonsense Ontology for Sentiment Analysis.” [Online]. Available: www.computer.org/inteligent
[20] R. Stahlbock, *ICDATA’19 : proceedings of the 2019 International Conference on Data Science*.
[21] K. Tabianan, D. Arputharaj, M. N. Bin, A. Rani, and S. Nagalingham, “Data Analysis and Rating Prediction on Google Play Store Using Data-Mining Techniques,” 2022.
[22] M. Umer, I. Ashraf, A. Mehmood, S. Ullah, and G. S. Choi, “Predicting numeric ratings for Google apps using text features and ensemble learning,” *ETRI Journal*, vol. 43, no. 1, pp. 95–108, Feb. 2021, doi: 10.4218/etrij.2019-0443.
[23] “UPCommons Portal del coneixement obert de la UPC”, doi: 10.1108/CR-12.
[24] C. Herodotou, C. Mangafa, and P. Srisontisuk, “An Experimental Investigation of ‘Drill-and-Practice’ Mobile Apps and Young Children,” *International Journal of Interactive Mobile Technologies*, vol. 16, no. 7, pp. 116–136, 2022, doi: 10.3991/ijim.v16i07.27893.
[25] B. Butcher and B. J. Smith, “Feature Engineering and Selection: A Practical Approach for Predictive Models,” *Am Stat*, vol. 74, no. 3, pp. 308–309, Jul. 2020, doi: 10.1080/00031305.2020.1790217.
[26] Southeast University (Bangladesh), Southeast University (Bangladesh). Department of Computer Science and Engineering, Southeast University (Bangladesh). Department of Electrical and Electronic Engineering, and Institute of Electrical and Electronics Engineers. Bangladesh Section, *ICCIT 2019 : 22nd International Conference on Computer and Information Technology : December 18-20, 2019*.
[27] T. G. Dietterich, “Ensemble Methods in Machine Learning.” [Online]. Available: http://www.cs.orst.edu/~tgd
[28] Universitas Atma Jaya Yogyakarta, Institute of Electrical and Electronics Engineers. Indonesia Section. Computer Society Chapter., Institute of Electrical and Electronics Engineers. Indonesia Section., and Institute of Electrical and Electronics Engineers, *2019 International Conference of Artificial Intelligence and Information Technology (ICAIIT 2019) : 13-15 March 2019, Platinum Adisucipto Yogyakarta Hotel & Conference Center*.
[29] F. Sarro, M. Harman, Y. Jia, and Y. Zhang, “Customer rating reactions can be predicted purely using app features,” in *Proceedings - 2018 IEEE 26th International Requirements Engineering Conference, RE 2018*, Institute of Electrical and Electronics Engineers Inc., Oct. 2018, pp. 76–87. doi: 10.1109/RE.2018.00018.
[30] *2017 15th International Conference on ICT and Knowledge Engineering (ICT & KE).* IEEE, 2017.
[31] V. T. Dhinakaran, R. Pulle, N. Ajmeri, and P. K. Murukannaiah, “App review analysis via active learning: Reducing supervision effort without compromising classification accuracy,” in

Proceedings - 2018 IEEE 26th International Requirements Engineering Conference, RE 2018, Institute of Electrical and Electronics Engineers Inc., Oct. 2018, pp. 170–181. doi: 10.1109/RE.2018.00026.

[32] B. T. Magar, S. Mali, and E. Abdelfattah, “App Success Classification Using Machine Learning Models,” in *2021 IEEE 11th Annual Computing and Communication Workshop and Conference, CCWC 2021*, Institute of Electrical and Electronics Engineers Inc., Jan. 2021, pp. 642–647. doi: 10.1109/CCWC51732.2021.9376021.

[33] O. Lengkong and R. Maringka, “Apps Rating Classification on Play Store Using Gradient Boost Algorithm,” in *2020 2nd International Conference on Cybernetics and Intelligent System, ICORIS 2020*, Institute of Electrical and Electronics Engineers Inc., Oct. 2020. doi: 10.1109/ICORIS50180.2020.9320756.

[34] Amity University. Dubai Campus, Institute of Electrical and Electronics Engineers. UAE Section, and Institute of Electrical and Electronics Engineers, *Abstract proceedings of International Conference on Computation, Automation and Knowledge Management (ICCAKM-2020) : 9th-10th January 2020*.

[35] D. Mota and R. Martinho, “An Approach to Assess the Performance of Mobile Applications: A Case Study of Multiplatform Development Frameworks,” in *International Conference on Evaluation of Novel Approaches to Software Engineering, ENASE - Proceedings*, Science and Technology Publications, Lda, 2021, pp. 150–157. doi: 10.5220/0010497401500157.

[36] G. Arther Sandag, F. Gara, and Irfan, “Android Application Market Prediction Based on User Ratings Using KNN,” in *2020 2nd International Conference on Cybernetics and Intelligent System, ICORIS 2020*, Institute of Electrical and Electronics Engineers Inc., Oct. 2020. doi: 10.1109/ICORIS50180.2020.9320812.

[37] W. Luiz *et al.*, “A feature-oriented sentiment rating for mobile app reviews,” in *The Web Conference 2018 - Proceedings of the World Wide Web Conference, WWW 2018*, Association for Computing Machinery, Inc, Apr. 2018, pp. 1909–1918. doi: 10.1145/3178876.3186168.

[38] J. Schmitt, J. Bönig, T. Borggräfe, G. Beitinger, and J. Deuse, “Predictive model-based quality inspection using Machine Learning and Edge Cloud Computing,” *Advanced Engineering Informatics*, vol. 45, Aug. 2020, doi: 10.1016/j.aei.2020.101101.

[39] S. Sadiq, M. Umer, S. Ullah, S. Mirjalili, V. Rupapara, and M. Nappi, “Discrepancy detection between actual user reviews and numeric ratings of Google App store using deep learning,” *Expert Syst Appl*, vol. 181, Nov. 2021, doi: 10.1016/j.eswa.2021.115111.

[40] R. Aralikatte, G. Sridhara, N. Gantayat, and S. Mani, “Fault in your stars: An Analysis of Android App Reviews,” Aug. 2017, doi: 10.1145/3152494.3152500.

[41] S. Ranjan and S. Mishra, “Comparative Sentiment Analysis of App Reviews,” Jun. 2020, [Online]. Available: http://arxiv.org/abs/2006.09739

[42] B. Lakimoski and B. Lakimoskon, “The role of performance benchmarks in selecting mobile app development frameworks.” [Online]. Available: https://www.researchgate.net/publication/337437805