

Original Article

Expense Tracker Application using Naive Bayes

Raj Thakare¹, Ninad Thakare², Raj Sangtani³, Shubham Bondre⁴, Amitkumar Manekar⁵

^{1,2,3,4,5}Information Technology, Shri Sant Gajanan Maharaj College of Engineering, Shegaon, Maharashtra, India.

Received: 13 April 2023

Revised: 21 May 2023

Accepted: 04 June 2023

Published: 16 June 2023

Abstract - This study introduces an Expense Tracker mobile application that utilizes the Naive Bayes algorithm for automated expense tracking. The app, developed for Android users using Kotlin and XML in Android Studio, allows manual entry of expenses and automatic detection of bank messages. The Naive Bayes algorithm is employed to classify these messages. The app provides visual representations of expenses through Pie Charts for multiple time frames such as monthly, weekly, yearly etc. It helps users gain insights into their spending habits. With Firebase as the online database, data persistence is ensured even if the app is uninstalled. Overall, the Expense Tracker app offers a user-friendly solution for individuals to manage their finances effectively and make informed decisions about their expenses.

Keywords - Machine learning, Personal finance management, Expense tracking, Predictive modeling, User interface.

1. Introduction

The "Expense Tracker" is a mobile application designed for Android users, utilizing Kotlin and XML languages within the Android Studio development environment. Its primary purpose is to assist users in effectively managing their daily expenses, allowing them to gain a comprehensive understanding of their spending habits and maintain better financial control. The app presents expense data through informative Pie Charts and Bar Graphs, enabling users to visualize and comprehend their expenditure patterns quickly. Furthermore, the application offers manual entry options for expenses, including amount, category, and description, and automated functionality to detect and extract relevant information from bank SMS notifications on the user's mobile device. By utilizing Firebase as the online database, the app ensures that data remains securely stored, even if the application is uninstalled. To access the "Expense Tracker" app, users are required to create an account using their email address and password.

2. Literature Survey

In [1], authors aim to classify spam emails from legitimate ones using Naive Bayes and SVM. SVM showed higher accuracy than NB (Naive Bayes) in diverse sizes of training emails. In [2], The results showed that Naive Bayes had higher accuracy on Spam Data but lower precision than SPAMBASE, where precision was more important for classifying spam emails correctly. The paper [3] focuses on evaluating ML techniques for spam SMS detection and compares 8 different classifiers. The CNNC (Convolutional-Neural-Network-Classifier) achieved the maximum accuracy of 99.19% & 98.25% for the two datasets, surpassing traditional classifiers like SVM and NB. After

analyzing several research studies, in [4], it was found that the supervised machine learning approach is highly adopted and produces consistent and accurate results. NB (Naive Bayes) and SVM algorithms are in high demand compared to other ML algorithms. This [5] survey paper provides an impression of different existing spam filtering systems using ML techniques, summarizing their accuracy and effectiveness in filtering email spam. While all the methods are effective, some have better outcomes, and others implement additional processes to increase accuracy. The paper [6] discusses the problem of detecting spamming activities on Twitter and the need to protect user privacy. The Lfun approach is introduced for spam detection, which involves training classifiers with training data and adding detected spam tweets to the training set over time. The paper proposes [7] a new approach called GADT, which uses a hybrid of genetic algorithm and decision tree for email spam detection, with the added benefit of using PCA to reduce high dimensionality. [8] aimed to classify Hindi Poetries using machine learning algorithms.

The task is challenging due to the language's morphological variance. The paper [9] aimed to classify "Aviation Safety Reports" using ML techniques and [26] proposed a framework to improve the prediction by the selection of correct attributes that contributes to better data extraction. The paper [28] presents a systematic literature review of 13 research papers on SMS spam detection, analysing their proposed techniques, advantages, disadvantages, challenges, and evaluation procedures. In [29] author discusses the difficulty of multi-label classification in NLP and text mining and the achievement of finding a solution for this problem in the less explored research domain of the Bengali language. The authors



suggest using a different word embedding techniques in the future and consider unsupervised or semi-supervised learning models for predicting new labels more cost-effectively. The paper [30] analyses ML and DL (Deep Learning) models for text classification on the 20NewsGroup dataset. The TF-IDF representation outperformed “Word2Vec” and “BERT” embeddings, and LR (Logistic Regression) and a bi-channel CNN model produced excellent results. The paper [31] discusses methods to build an ML model to identify spam emails using popular algorithms, and the random tree outperforms other classification algorithms in spam filtering. According to [32], NB (Naïve Bayes) outperforms LR (logistic regression) and RF (random forest) algorithms in SMS spam classification with high accuracy of 98.445%. Random Forest also performed well and has a shorter running time, but Naïve Bayes is the best choice. This study [33] focuses on developing an expense management system for users to track and manage their expenses effectively. The existing system is limited to desktop software and lacks portability, hindering users from updating expenses and their locations.

The proposed system, built using Android and SQL Lite, allows users to register and provides features such as expense tracking, income tracking, and generating customized reports. The system includes modules for adding and listing income and expense details, setting reminders for future expenses, and filtering data by month, year, and date. The research highlights the user-friendly nature of the application. It emphasizes the importance of mobile application testing and future enhancements, including support for different Android versions and viewing deleted details from the history. Overall, the system aims to provide users with a convenient and efficient method of managing their expenses on the go. The article [34] titled "Expense Tracker: An Android Application for Daily Expense Tracking" introduces the development of an Android app called "Expense Tracker." The app serves as a digital record-keeping tool, allowing users to track their daily expenses and income. It provides features such as setting a daily expense limit and issuing warnings when exceeded, thereby promoting better financial management. The app also enables users to save money by adding any remaining funds from daily expenses to their savings. Additionally, it generates monthly expense reports in PDF format. The paper outlines the technology used, including Angular 8 for the front end and SQLite for the back end. The app requires user registration, and expenses can be categorized by week, month, and year. Overall, the Expense Tracker app offers a user-friendly solution for effective expense tracking and financial stability. This research article [35] focuses on developing a mobile application for expense tracking and budget control. The application utilizes computer vision and text recognition algorithms to scan and process receipts captured by the phone's camera. The relevant data from the receipts are extracted and stored in a database, allowing for

expense statistics and analysis. The study successfully implements a working text recognition algorithm and identifies the necessary data for expense tracking. The preservation and storage of the data are also implemented. The research highlights the importance of personal budget control and the convenience of using mobile applications for expense accounting. The future steps involve expanding the supported stores and enhancing the application's functionality. The proposed web application in an article [37], "Online Income and Expense Tracker," aims to assist individuals in managing their income and expenses efficiently. It provides a convenient platform to record and monitor daily or periodic financial transactions, serving as a reminder and indicator for upcoming payments and tasks. The application includes features such as categorizing expenses, capturing photos, adding location information, and setting income and expense limits. By utilizing this tracker, users can reduce manual calculations, track their expenditures, and easily share expenses with others. The research highlights the importance of budget tracking and the need for efficient tools to manage finances effectively in today's fast-paced society. The project's innovative approach and user-friendly modules offer an attractive and efficient income and expense management solution. The research study [38] discusses the development of a mobile application called Expense Manager, designed to track personal and group expenses, maintain monthly incomes, and provide information on debts and contributions. The application aims to eliminate confusion and data inconsistencies in expense recording and splitting, replacing traditional methods like spreadsheets and ledgers. Additional features include user registration through email or social networking accounts, synchronization of profile information, and collection of sample expense data for market analysis using data mining techniques. The paper highlights the benefits of the application in helping users manage their expenses effectively while assisting marketing executives in planning targeted marketing strategies.

3. Activity Diagram

The activity Diagram for Expense Tracker Application, as given in Fig. 1, shows how the expense tracker application goes through multiple activities and phases.

The given activity diagram shows that at first user goes to the authentication activity, and if the user is new, then that user will need to sign up or else log in can be used to enter the application. From home, the user can go and view all expenses activity; another option from home is to add the added expense with details like amount, date, description and category. Charts activity shows the user's statistics for multiple time frames such as weekly, monthly, yearly etc. The user can add or view existing goals from goal activity. By clicking on the profile icon from the home user can go to profile activity, where he can add personal details and also sign out from the app.

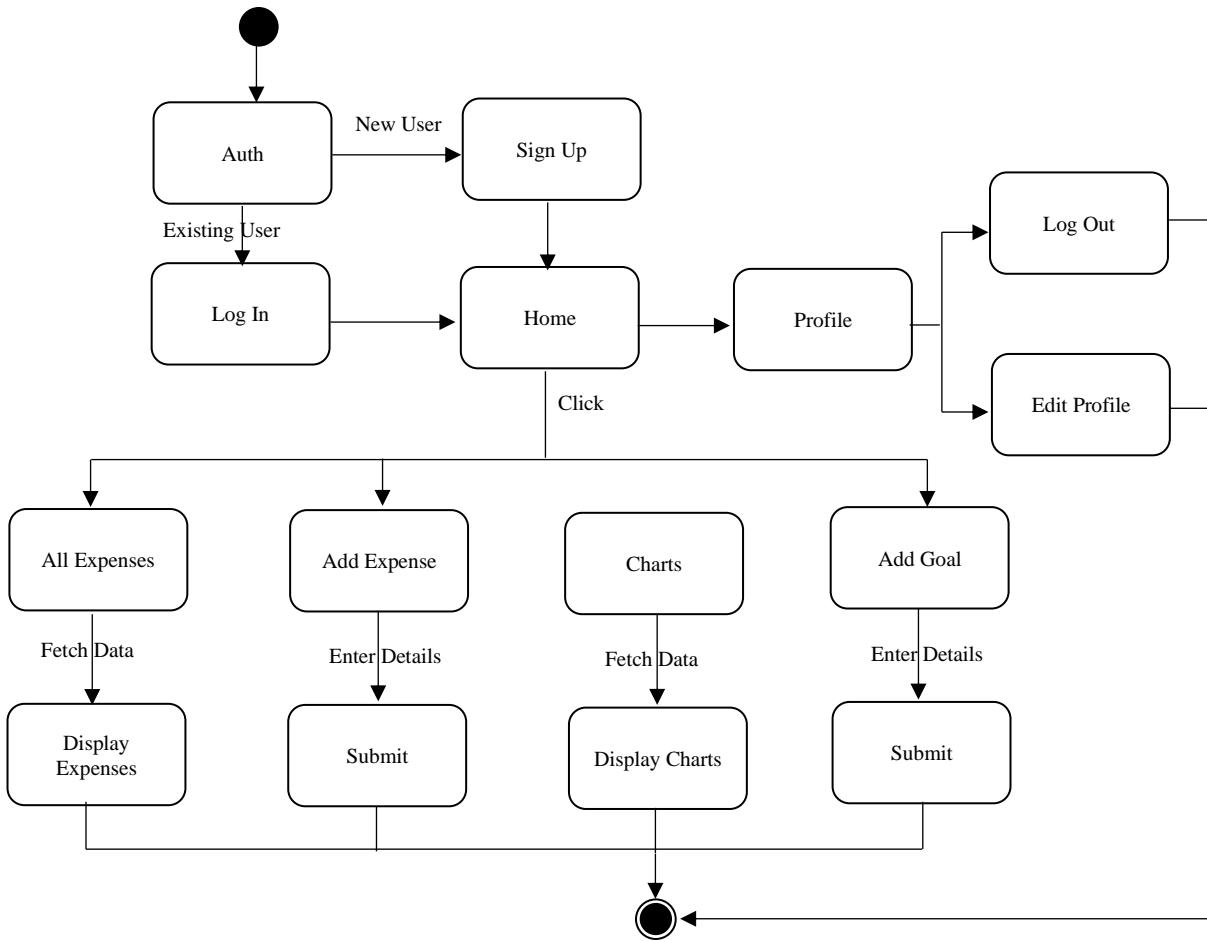


Fig. 1 Activity Diagram

4. Modules

4.1. Requirement Analysis

The first step in developing an expense tracker application is to perform requirement analysis. This involves identifying the user's needs and requirements for the application, as well as the features and functionality that will be included in the application.

4.2. User Authentication

The next module in the sequence is user authentication. This module is responsible for handling user registration, login, and logout functionality. It ensures that only authorized users can access the application.

4.3. Income & Expense Module

The income & expense module is the core of the expense tracker application. It provides users with the ability to add and manage their income and expenses. This module includes features such as adding income, adding expenses, categorizing expenses, and viewing income and expense reports.

4.4. User Statistics Module

The user statistics module provides users with an overview of their financial situation. It displays various statistics and insights related to their income and expenses, such as total income, total expenses, monthly income and expense trends, and budget tracking.

4.5. ML Implementation Module

Machine learning is used to read all device messages and then feed them to the API of the machine learning model. This machine-learning model will give us bank-related messages. And then, with the help of programming libraries, we will extract the amount from the message.

4.6. Testing

Testing is an important module in the development process. It ensures the application functions correctly and meets the user's needs and requirements. Testing can include functional testing, unit testing, integration testing, and user acceptance testing.

5. Result

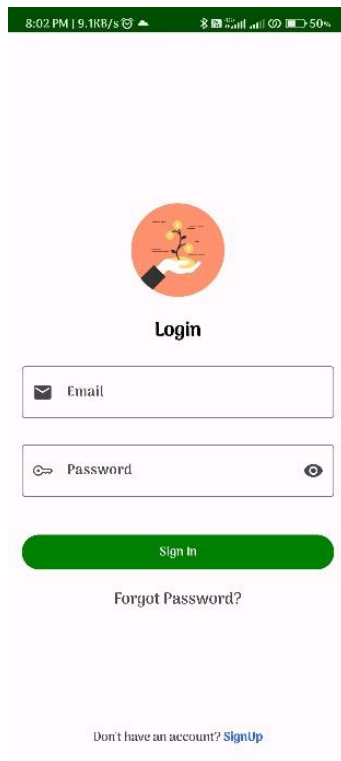


Fig. 2 Authentication

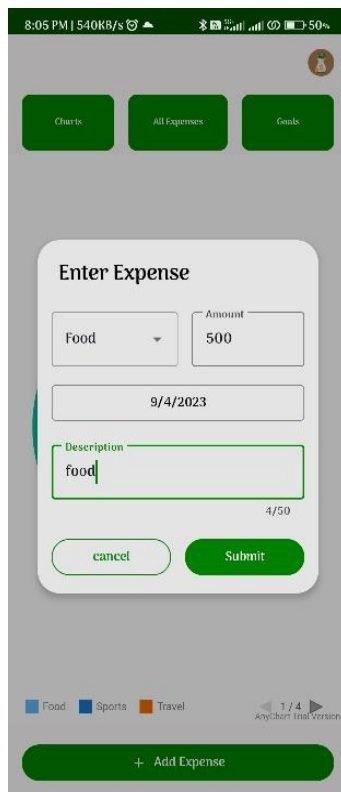


Fig. 3 Add Expense



Fig. 4 All Expenses

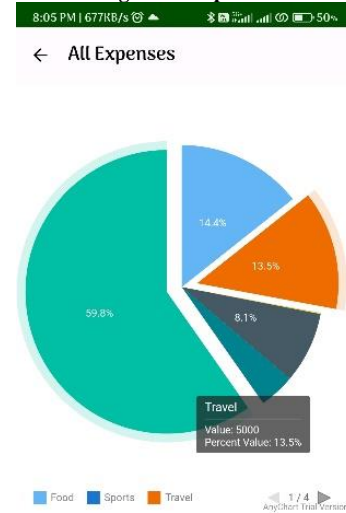


Fig. 5 Pie Chart

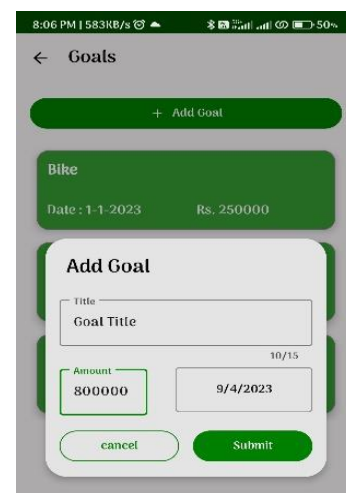


Fig. 6 Add Goal

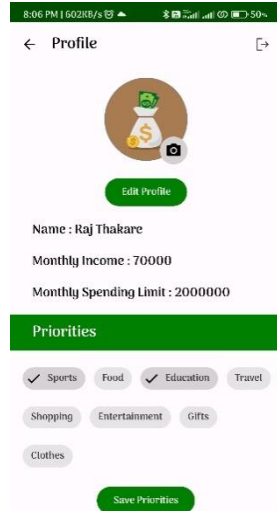


Fig. 7 Profile

6. Conclusion

We developed an Android application that helps users track their daily expenses by allowing them to manually enter their expenses or automatically detect and read expense messages from the bank using a Naive Bayes algorithm. The algorithm was trained using supervised machine learning to accurately classify SMS as "Bank SMS" or "Normal SMS". Overall, the proposed system has the potential to help people manage their finances better and make informed decisions about their spending habits. By automatically detecting and reading expense messages from the bank, the system can help users keep track of their expenses more accurately and efficiently. Moreover, the ability to classify SMS messages as "Bank SMS" or "Normal SMS" can help users filter out irrelevant messages and focus on the ones that are relevant to their financial management.

References

- [1] Thae Ma Ma, Kunihito Yamamori, and Aye Thid, "A Comparative Approach to Naive Bayes Classifier and Support Vector Machine for Email Spam Classification," *2020 IEEE 9th Global Conference on Consumer Electronics, IEEE*, pp. 324-326, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [2] Ziyang Mohammed et al., "A Comparative Study for Spam Classifications in Email Using Naïve Bayes and SVM Algorithm," *Journal of Emerging Technologies and Innovative Research*, vol. 6, no. 5, 2019. [[Google Scholar](#)] [[Publisher Link](#)]
- [3] Mehul Gupta et al., "A Comparative Study of Spam SMS Detection using Machine Learning Classifiers," *Proceedings of 2018 Eleventh International Conference on Contemporary Computing, IEEE*, Noida, India, pp. 1-7, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [4] Mansoor Raza et al., "A Comprehensive Review on Email Spam Classification using Machine Learning Algorithms," *International Conference on Information Networking, IEEE*, pp. 327-332, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [5] H. Bhuiyan et al., "A Survey of Existing E-Mail Spam Filtering Methods Considering Machine Learning Techniques," *Journal Name, Global Journal of Computer Science and Technology*, 2018. [[Google Scholar](#)] [[Publisher Link](#)]
- [6] Rutuja Katpatal, and Aparna Junnarkar, "An Efficient Approach of Spam Detection in Twitter," *Proceedings of the International Conference on Inventive Research in Computing Applications*, pp. 1240-1243, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [7] Ahmed I. Taloba, and Safaa S. I. Ismail, "An Intelligent Hybrid Technique of Decision Tree and Genetic Algorithm for E-Mail Spam Detection," *IEEE Ninth International Conference on Intelligent Computing and Information Systems, IEEE*, pp. 99-104, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [8] Kaushika Pal, and Biraj. V. Patel, "Automatic Multiclass Document Classification of Hindi Poems using Machine Learning Techniques," *International Conference for Emerging Technology*, Belgaum, India, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [9] V. de Vries, "Classification of Aviation Safety Reports using Machine Learning," *International Conference on Artificial Intelligence and Data Analytics for Air Transportation, IEEE*, pp. 1-6, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [10] J. Phani Prasad, and T. Venkatesham, "Classification of E-mail Spam with Supervised Machine Learning-Naïve Bayesian Classification," *Advances and Applications in Mathematical Sciences*, vol. 20, no. 12, pp. 3087-3092, 2021. [[Google Scholar](#)] [[Publisher Link](#)]
- [11] Aditya Shrivastava, and Rachana Dubey, "Classification of Spam Mail using Different Machine Learning Algorithms," *International Conference on Advanced Computation and Telecommunication, IEEE*, pp. 1-10, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [12] Kothapally Nithesh Reddy, and Vijayalakshmi Kakulapati, "Classification of Spam Messages using Random Forest Algorithm," *Journal of Xidian University*, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [13] N. Krishnaveni, and V. Radha, "Comparison of Naive Bayes and SVM Classifiers for Detection of Spam SMS Using Natural Language Processing," *ICTACT Journal on Soft Computing*, vol. 11, no. 2, pp. 2260-2265, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [14] G. Bharath et al., "Detecting Fake News Using Machine Learning Algorithms," *International Conference on Computer Communication and Informatics*, 2021.

- [15] Abdullah-All-Tanvir et al., "Detecting Fake News using Machine Learning and Deep Learning Algorithms," *2019 7th International Conference on Smart Computing & Communications*, East West University, Dhaka, Bangladesh, pp. 1-5, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [16] Zulfikar Alom, Barbara Carminati, and Elena Ferrari, "Detecting Spam Accounts on Twitter," *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, IEEE*, pp. 1191-1198, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [17] Prachi Gupta, Ratnesh Kumar Dubey, and Dr. Sadhna Mishra, "Detecting Spam Emails/Sms Using Naive Bayes and Support Vector Machine," *International Journal of Scientific & Technology Research*, vol. 8, no. 11, 2019. [[Publisher Link](#)]
- [18] Manmohan Singh et al., "Email Spam Classification by Support Vector Machine," *International Conference on Computing, Power and Communication Technologies, IEEE*, pp. 878-882, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [19] Shradhanjali, and Toran Verma, "E-Mail Spam Detection and Classification Using SVM and Feature Extraction," *International Journal of Advance Research, Ideas and Innovations in Technology*, vol. 3, no. 3, pp. 1491-1495, 2017. [[Google Scholar](#)] [[Publisher Link](#)]
- [20] Kriti Agarwal, and Tarun Kumar, "Email Spam Detection using integrated approach of Naïve Bayes and Particle Swarm Optimization," *Proceedings of the Second International Conference on Intelligent Computing and Control Systems*, pp. 685-690, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [21] Nikhil Kumar, Sanket Sonowal, and Nishant, "Email Spam Detection Using Machine Learning Algorithms," *Proceedings of the Second International Conference on Inventive Research in Computing Applications*, pp. 108-113, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [22] Megha Tope, "Email Spam Detection using Naive Bayes Classifier," *International Journal of Scientific Development and Research*, vol. 4, no. 6, pp. 1-7, 2019. [[Google Scholar](#)] [[Publisher Link](#)]
- [23] Muhammad Azam et al., "Feature Extraction based Text Classification using K-Nearest Neighbor Algorithm," *International Journal of Computer Science and Network Security*, vol. 18, no. 12, pp. 95-101, 2018. [[Google Scholar](#)] [[Publisher Link](#)]
- [24] R. Deepa, and Kiran N Lalwani, "Image Classification and Text Extraction Using Machine Learning," *Proceedings of the Third International Conference on Electronics Communication and Aerospace Technology, IEEE*, pp. 680-684, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [25] Sandy Kurniawan, and Indra Budi, "Indonesian Tweets Hate Speech Target Classification using Machine Learning," *Fifth International Conference on Informatics and Computing*, pp. 1-5, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [26] U. Murugavel, and R. Santhi, "K-Nearest Neighbor Classification of E-Mail Messages for Spam Detection," *ICTACT Journal on Soft Computing*, vol. 11, no. 1, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [27] Zeeshan Bin Siddique et al., "Machine Learning-Based Detection of Spam Emails," *Scientific Programming*, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [28] Samadhan Nagre, "Mobile SMS Spam Detection using Machine Learning Techniques," *Journal of Emerging Technologies and Innovative Research*, vol. 5, no. 12, pp. 548-550, 2018. [[Publisher Link](#)]
- [29] Wahiduzzaman Akanda, and Ashraf Uddin, "Multi-label Bengali article classification using ML-KNN algorithm and Neural Network," *Proceedings of the International Conference on Information and Communication Technology for Sustainable Development*, pp. 466-471, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [30] C. M. Suneera, and Jay Prakash, "Performance Analysis of Machine Learning and Deep Learning Models for Text Classification," *IEEE 17th India Council International Conference*, pp. 1-6, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [31] S. Nandhini, and K.S Jeen Marseline, "Performance Evaluation of Machine Learning Algorithms for Email Spam Detection," *International Conference on Emerging Trends in Information Technology and Engineering, IEEE*, pp. 1-4, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [32] Paras Sethi, Vaibhav Bhandari, and Bhavna Kohli, "SMS Spam Detection and Comparison of Various Machine Learning Algorithms," *International Conference on Computing and Communication Technologies for Smart Nation, IEEE*, pp. 28-31, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [33] Y.Anitha, R.Ranjini, and S.Gomathi, "Easy App for Expense Manager Using Android," *International Journal of Trend in Research and Development*, vol. 3, no. 2, pp. 65-67, 2016. [[Google Scholar](#)] [[Publisher Link](#)]
- [34] Atiya Kazi et al., "Expense Tracker," *IRE Journals*, vol. 4, no. 11, pp. 19-21, 2021. [[Google Scholar](#)] [[Publisher Link](#)]
- [35] Denis E. Yurochkin, Anton A. Horoshiy, and Saveliy A. Karpukhin, "Development of an Application for Expense Accounting," *IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering*, Moscow, Russian Federation, pp. 753-757, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [36] Jayamala Kumar Patil et al., "A Novel Method of Traffic Rule Compliance Detection Using Morphological Operations at Toll Booths," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 10, no. 1, pp. 148-159, 2023. [[CrossRef](#)] [[Publisher Link](#)]

- [37] S. Chandini et al., "Online Income and Expense Tracker," *International Research Journal of Engineering and Technology*, vol. 6, no. 3, pp. 4119-4124, 2019. [[Google Scholar](#)] [[Publisher Link](#)]
- [38] N.Zahira Jahan, and K.I. Vinodhini, "Personalized Expense Managing Assistant Using Android," *International Journal of Computer Techniques*, vol. 3, no. 2, 2016. [[Google Scholar](#)] [[Publisher Link](#)]