KeAi
CHINESE ROOTS
GLOBAL IMPACT

International Journal of Innovation Studies

SENTIMENT ANALYSIS ON MOBILE REVIEWS BY USING DEEP LEARNING MODELS

Ameya Parkar, Rajni Bhalla
School of Computer Application, Lovely Professional University, India
E-mail: parkar.ameya@gmail.com, rajni.b27@gmail.com

Abstract
The last two decades has been people posting about different aspects of their lives online. Be it their private life, their pets, parents, siblings, kids, etc. It has become a norm in the day-to-day life of different bloggers and people who post videos online. Opinion mining/Sentiment Analysis falls under Natural Language Processing and is used to judge the sentiment expressed by the people. The sentiment could be positive, negative or neutral. We have used deep learning techniques like Recurrent Neural Networks, Convolution Neural Networks amongst others to find sentiments of different users in the domain of mobile reviews. Bi-directional Encoder Representations from Transformers gives the best accuracy score of 98% and an f-score of 97%.

Keywords: Deep Learning, Sentiment Analysis, Mobile reviews

Introduction

Since the recent few years social media has seen a rise in people expressing their opinions and views on different aspects of their lives. Be it posting a video online about their day-to-day life, expressing their opinions on a newly launched mobile phone, expressing their political views on a particular topic and so on. Companies can use this data to judge the quality of their products and services. This process is called as sentiment analysis. A particular review could be categorized as positive, negative or neutral. Some customers also express their opinions by rating the products using stars. With 1 star being the lowest and 5 stars being the highest. It is important to categorize the reviews so as to understand the future action to be taken regarding the product.

Rest of the paper is structured as follows. The next section reviews the literature followed by the proposed methodology, followed by the findings, challenges and the conclusion of the paper.

Literature review

Raman R. et al. [1] used Naïve Bayes classifier to analyse sentiments in application reviews. The reviews were categorized as positive, negative and neutral. They achieved an accuracy score of 85% using the Naïve Bayes classifier.

Lestari N. et al. [2] used word2vec embeddings and Convolution Neural Network technique to detect sentiment in mobile reviews. They achieved an accuracy score of 71% for aspect-based sentiment analysis.

Maroof A. et al. [3] used a hybrid framework of rule based and machine learning and deep learning techniques for aspect-based sentiment analysis. Accuracy score of 79% was achieved using the hybrid technique.

Selamet B. et al. [4] used BERT word embeddings to create vectors. GRU technique was used to perform aspect-based sentiment analysis on Indonesian Google play store reviews on mobile phones. An accuracy score of 93.39% was achieved on the mobile reviews using GRU technique.

Tripathy A. et al. [5] performed document-based sentiment analysis on IMDB dataset of movie reviews using Support Vector Machine technique with an accuracy score of 95%.

Roy G. et al. [6] used six deep learning techniques along with clustering to classify sentiments in Google play store reviews of mobile phones. All the techniques gave a good performance score in classifying the sentiments.

Meng F. et al. [7] used multi head attention mechanism & multi-layer perceptron mixer model to perform sentiment analysis on educational app reviews. They compared the model with standard deep learning techniques to classify sentiments with performance scores of more than 85%.

Suryadi D. et al. [8] used Latent Dirichlet Allocation and SenticNet library to classify sentiments in digital bank applications.

Nurmiati E. et al. [9] used Naïve Bayes and Latent Dirichlet Allocation techniques to classify sentiments and topic modelling on mobile banking reviews in Indonesian language with an accuracy score of 73.60%.

Irianto M. et al. [10] used word2vec word embeddings for extracting features in application reviews. KNN method was used to perform sentiment analysis with an f-score of 91.98%.

Putri C. et al. [11] used term frequency-inverse document frequency for word embeddings and Support Vector Machine to classify sentiments of website cosmetic reviews with an accuracy score of 90.4%.
Alemerien K. et al. [12] used Support Vector Machine technique along with TF-IDF word embeddings to perform sentiment analysis of online reviews with an accuracy score of 97%.

Jamadar J. et al. [13] used topic modelling and performed sentiment analysis on mobile banking reviews. LDA technique along with RoBERT technique was used for sentiment analysis.

Putri M. et al. [14] performed aspect-based sentiment analysis on Indonesian app reviews using IndoBERT lite technique with an accuracy score of 85%.

Table 1: Related Work.

References	Objective of the paper	Dataset Size	Language	Performance A: Accuracy P: Precision R: Recall F: F-score
[1]	Sentiment Analysis on Application reviews	Motorku app	English	Naïve Bayes A: 85%
[2]	Aspect based sentiment analysis on mobile reviews	MyPertamina app 1776 reviews	Indonesian	CNN A: 71%
[3]	Aspect based sentiment analysis on service industry	Mobile reviews 4165 sentences	English	Rule + ML + DL A: 79%
[4]	Aspect based sentiment analysis on mobile reviews	Google play 7994 reviews	Indonesian	BERT + GRU A: 93.39%
[5]	Document level sentiment analysis on movie reviews	IMDB	English	SVM A: 95%
[6]	Sentiment Analysis using Transformers	Google play 73483 reviews	English	BERT A: 99%
[7]	Sentiment Analysis on mobile reviews	Educational Apps 13000	English	MHA-MLP model F: 86.96%
[8]	Sentiment Analysis on Digital Bank applications	Digital Bank reviews 7861	English	LDA
[9]	Sentiment Analysis on Mobile Banking reviews	Mobile Banking 45000 reviews	Indonesian	Naïve Bayes + LDA A: 73.60%
[10]	Sentiment Analysis on Application reviews	App reviews	Indonesian	KNN F: 91.98%
[11]	Sentiment Analysis on cosmetic product reviews	Cosmetic products 208 reviews	English	SVM A: 90.4%
[12]	Sentiment Analysis on online reviews	Online hotel 33575 reviews	English	SVM A: 97%

[13]	Sentiment Analysis on Mobile Banking reviews	Mobile banking 3000 reviews	English	RoBERT + LDA
[14]	Aspect based sentiment analysis of Indonesian app reviews	5000 Tokopedia reviews	Indonesian	IndoBERT A: 85%

(All values are in percentages)

Proposed Methodology

The dataset chosen for this research study is the mobile reviews posted on the Amazon website. We extracted 71922 records. The dataset contains columns like date of posting the review, the name of the user who posted it, the review itself and the sentiment of the review (positive, negative or neutral).

Python was the language chosen for the research study. Different libraries like nltk, pandas, keras, tensorflow, etc. were used. Data was unprocessed and therefore, preprocessing techniques like tokenization, removal of stop words, removal of unnecessary symbols, lemmatization, etc. Word2vec technique was used to convert text to vectors. The dataset was split into 80% for training and 20% for testing.

Different standard models of deep learning neural networks like Artificial Neural Networks, Recurrent Neural Networks, etc. were used for training the model. The Results section contains more details about the techniques used for this research study.

Overview of Evaluation Metrics

The performance of the model used for this research study was judged on different performance metrics like accuracy, precision, recall and f-score. They are explained as follows:

Accuracy (A)

Accuracy metric is to be used when we are judging the performance of the model in terms of positives as well as negatives under consideration.

A = (True Positive + True Negative) / (True Positive + True Negative + False Positive + False Negative)

Precision (P)

Precision metric is to be used when we are focusing on the class of positives and ignoring the class of negatives.

P = True Positive / (True Positive + False Positive)

Recall (R)

Recall metric is to be used when we are focusing on the class of positive samples and class of false negative samples.

R = True Positive / (True Positive + False Negative)

F-score (F)

F-score is the harmonic mean of precision and recall and works well on balanced as well as unbalanced datasets.

F = (2 * precision * recall) / (precision + recall)

Results and Discussion

We used deep learning classifiers on the Amazon dataset of mobile reviews and the results are summarized in Table 2 below:

Table 2. Performance of Machine Learning classifiers

Techniques	A	P	R	F
Artificial Neural Network	87.2	91.99	92	91.99
Recurrent Neural Network	85.91	86.48	97.64	91.72
Long Short-Term Memory	85.83	87.23	96.39	91.58
Bi-directional Long Short-Term Memory	85.95	90.38	97.26	93.69
Convolution Neural Network	85.46	93.68	87.73	90.61
Bi-directional Encoder Representations from Transformers	91.57	94.01	95.54	94.77

(All values are in percentages)

From Table 2, it is clearly evident that Bi-directional Encoder Representations from Transformers with accuracy of 91.57%, precision of 94.01%, recall of 95.54%, f-score of 94.77% gave the best performance measures. The f-score of all the classifiers were above 90%.

Conclusion

Sentiment analysis is a way to judge the polarity of the user opinions. From the Kaggle website, we downloaded a dataset on mobile phone reviews. We used different machine learning classifiers to train the model and found that Logistic Regression and Support Vector Machine gave the best performance on classifying the mobile reviews. In future work, we will include deep learning algorithms as well as transformers.

Opinion Mining is used to judge the sentiment of reviews. We extracted mobile reviews from the Amazon website and used deep learning techniques to train the model and tested it on the testing dataset with different performance metrics. Bi-directional Encoder Representations from Transformers technique gave the best performance score with an accuracy of 91.57%.

For future scope, we will work on different types in Natural Language Processing such as detection of sarcasm.

References

[1] R. Raman, V. Kumar, B. G. Pillai, D. Rabadiya, P. K. Bhoyar, and N. Kumbhojkar, "Analyzing Sentiments in Motorku X App Reviews using the Naive Bayes Classifier Approach," in *2nd IEEE International Conference on Data Science and Information System, ICDSIS 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/ICDSIS61070.2024.10594505.

[2] N. I. Lestari, S. M. Taib, W. Wibowo, I. A. Aziz, and M. R. Habibi, "Aspect-Based Sentiment Analysis for Mobile App Review Using Convolutional Neural Network (CNN) and Word2Vec," in *2024 IEEE 7th International Conference on Electrical, Electronics, and System Engineering: Dissemination and Advancement of Engineering Education using Artificial Intelligence, ICEESE 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/ICEESE62315.2024.10828541.

[3] A. Maroof, S. Wasi, S. I. Jami, and M. S. Siddiqui, "Aspect-Based Sentiment Analysis for Service Industry," *IEEE Access*, vol. 12, pp. 109702–109713, 2024, doi: 10.1109/ACCESS.2024.3440357.

[4] Selamet, A. Bejo, and L. M. Putranto, "Aspect-Based Sentiment Analysis on PLN Mobile Application Using Word Embedding Bert and GRU Classification Algorithm: a Case Study of Play Store Comments in 2023," in *2024 7th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, IEEE, Dec. 2024, pp. 777–783. doi: 10.1109/ISRITI64779.2024.10963459.

[5] A. Tripathy, C. C. Rao, S. S. Sen, and M. Tripathy, "Document Level Sentiment Analysis Using Aspect Based Approach for Movie Reviews," in *Proceedings of the 2024 International Conference on Artificial Intelligence and Emerging Technology, Global AI Summit 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 987–992. doi: 10.1109/GlobalAISummit62156.2024.10947829.

[6] G. T. Roy and D. Biswas, "Exploring Transformer and Recurrent Neural Models for Sentiment Analysis of Mobile App Reviews," in *2nd International Conference on Information and Communication Technology, ICICT 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 209–213. doi: 10.1109/ICICT64387.2024.10839678.

[7] F. Meng, M. Qi, and S. Bao, "Innovative Sentiment Analysis of Mobile App Review Texts Based on MHA-MLP-Mixer," in *Proceedings - 2024 IEEE Smart World Congress, SWC 2024 - 2024 IEEE Ubiquitous Intelligence and Computing, Autonomous and Trusted Computing, Digital Twin, Metaverse, Privacy Computing and Data Security, Scalable Computing and Communications*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 1150–1157. doi: 10.1109/SWC62898.2024.00183.

[8] D. Suryadi and K. O. Padlan, "Leveraging Topic Modeling and Sentiment Analysis to Improve Digital Bank Applications," in *2024 5th International Conference on Data Analytics for Business and Industry, ICDABI 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 86–90. doi: 10.1109/ICDABI63787.2024.10799999.

[9] E. Nurmiati, M. Qomarul Huda, and S. Mitsalina, "Sentiment Analysis and Topics Modeling on Mobile Banking Reviews of Sharia Bank in Indonesia Using Naive Bayes and Latent Dirichlet Allocation," in *2024 12th International Conference on Cyber and IT Service Management, CITSM 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/CITSM64103.2024.10775521.

[10] M. D. R. Irianto, M. D. Purbolaksono, and Bunyamin, "Sentiment Analysis of Livin' by Mandiri Application Reviews Using Word2Vec Feature Extraction and KNN Method," Institute of Electrical and Electronics Engineers (IEEE), Sep. 2024, pp. 236–241. doi: 10.1109/icodsa62899.2024.10651870.

[11] C. Putri, C. Salim, W. K. Veronica Dwiyanti, S. Achmad, and J. V. Moniaga, "Sentiment Analysis of Makeup Alley Website Cosmetic Reviews Using SVM and TF-IDF," in *Proceedings of 2024 International Conference on Information Management and Technology, ICIMTech 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 334–339. doi: 10.1109/ICIMTech63123.2024.10780915.

[12] K. Alemerien, A. Al-Ghareeb, and M. Z. Alksasbeh, "Sentiment Analysis of Online Reviews: A Machine Learning Based Approach with TF-IDF Vectorization," *Journal of Mobile Multimedia*, vol. 20, no. 5, pp. 1089–1116, 2024, doi: 10.13052/jmm1550-4646.2055.

[13] J. Jamadar, K. Karnik, A. Birari, and Y. Patil, "User Perception of Mobile Banking: Application of Sentiment Analysis and Topic Modelling Approaches to Online Reviews," Institute of Electrical and Electronics Engineers (IEEE), Sep. 2024, pp. 1–6. doi: 10.1109/tqcebt59414.2024.10545236.

[14] N. A. Putri Masaling, H. Soeparno, Y. Arifin, and F. L. Gaol, "Utilizing Sentiment Insights for Software Evalution with Aspect-Based Sentiment Analysis in App Reviews Using IndoBERT-Lite," in *2024 7th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, IEEE, Dec. 2024, pp. 977–982. doi: 10.1109/ISRITI64779.2024.10963515.