



Study of Review Classification & Neural Network

¹Mr. Rushikesh B. Bundele, ²Mr. Ravi V. Mante

¹Student, ²Assistant Professor

¹Computer Science & Engineering, ²Computer Science & Engineering

¹Government College of Engineering, Amravati, India, ²Government College of Engineering, Amravati, India

Abstract: Review plays an important role for consumer as well as for manufacturer, seller so this reviews are needed to be analyzed in such a manner so that it can give better results for effective usage the reviews data is huge so to analyze this data there are various types of techniques to be analyzed this review data. This paper studies various approaches & techniques to analyze the data & word embedding is also studied. A system is proposed to use Neural network, word embedding for review classification.

Index Terms – Word Embedding, Classification, Neural Network, Machine Learning.

I. INTRODUCTION

With the increase of web connectivity and low cost internet access, E-commerce industry has ballooned over time. Many products and services are available online at click of a button. This has helped connect both consumers and providers. Consumers now have a host of providers to choose from based on their needs and capacity. This not only give a lot of option to choose from but also gave rise to the problem of which provider to select.

Mobile devices became a crucial consider life as individuals use them for creating calls, taking photos, surfing the web, taking part in games etc. skillfulness of mobile devices have semiconductor diode to plenty of mobile applications on the Apple App Store and Google Play Store, etc., As there area unit several applications within the same class and also the competition is incredibly high within the market, software system corporations and freelance developers have to be compelled to maintain quality of their applications. A channel for application users to voice their opinions concerning the standard of the applications area unit posting their reviews on the App Store and Play Store.

User reviews square measure direct messages from real users square measure victimisation mobile applications and that they square measure necessary for application improvement. Some users describe the issues they face whereas victimisation the applications. Some users counsel new options to reinforce the applications. Some categorical their feelings concerning the applications and provides rating. There are user reviews with content that's useless or shy for the mobile application team to form use of. In general, there may well be loads of such user reviews for a mobile application. particularly for a well-liked application, there may well be many new user reviews per day. it's so tough for a mobile development team to travel through user reviews of their application, one by one, to research issues and opportunities for package sweetening. The question is, however will we tend to scale back this user review analysis effort and improve mobile application quality at an equivalent time?

The information that may be obtained from product and repair reviews isn't solely helpful to customers, however additionally for firms. Knowing what has been announce on the online will facilitate firms improve their merchandise or services. great amount of reviews has been announce on-line.

Companies invariably need to grasp the success rate of their products/apps, and for that, they typically request users to supply feedback or reviews that's later accustomed analyze the impact and quality of their merchandise. However, it's important to research such feedback thanks to the quantity and redundancy.

Therefore, this work investigates associate degree economical thanks to analyze such feedback and solve the issues associated with the classification of app reviews.

Neural networks represent deep learning using artificial intelligence. Certain application scenarios are too heavy or out of scope for traditional machine learning algorithms to handle. As they are commonly known, Neural Network pitches in such scenarios and fills the gap.

Artificial neural networks area unit galvanized from the biological neurons at intervals the build that activate underneath sure circumstances leading to a connected action performed by the body in response. Artificial neural nets encompass numerous layers of interconnected artificial neurons steam-powered by activation functions that facilitate in change them ON/OFF. Like ancient machine rule, here too, there area unit sure values that neural nets learn within the coaching part.

Briefly, each neuron receives a multiplied version of inputs and random weights which is then added with static bias value (unique to each neuron layer), this is then passed to an appropriate activation function which decides the final value to be given out of the neuron. There are various types of hidden layers involve in the artificial neural network, it is working as a feed forward neural network, suppose there are 3 layer so 1st layer output will be acting as a input for 2nd layer and 2nd layer output will be acting as a input for 3rd layer, here each neuron having weight with & bias is also acting in obtaining the output There are various activation functions available as per the nature of input values. Once the output is generated from the final neural net layer, loss function (input vs output) is calculated and back propagation is performed where the weights are adjusted to make the loss minimum. Finding optimal values of weights is what the overall operation is focusing around. Please refer to the following for better understanding-

There are various types of learning techniques which are as follows

- Supervised Learning
- Unsupervised Learning
- Semi supervised Learning

Supervised Learning : In Supervised learning well labeled data is used to train the model so this model perform better on seen data. It is just like a training in presence of teacher suppose consider if the model needs to recognized the image of dog form lots of images so here the model is trained on different animal images that having various types of label attached so that it will give better accuracy while recognizing a different problem rather than on training data. This type is used for to solve classification problem in which there are different classes are there & data need to be classified in the different classes by using the features

Unsupervised Learning: In this types the Model need to perform on its here training is not given on labeled data. Here model needs to understand from lots of data here not labeled special training is done so model needs to understand the features & on the basis of that it need to give the output. This type of technique is mainly used in clustering problem. Clustering means collecting similar feature items on one type

Semi supervised Learning: It uses both types of approaches like supervised & semi-supervised that's why it is known as semisupervised learning. Here some data is well labeled & some data is not well labeled so model need to perform on this whole data to obtain the result so this types of learning is known in semi supervised learning.

There are different types of deep learning processes which are as follows

- Artificial Neural Networks (ANN)
- Convolution Neural Networks (CNN)
- Recurrent Neural Networks (RNN)

Convolutional Neural Network: Deep learning is mainly inspired by the human brain and typically uses a multilayer perceptron (MLP) algorithm for classification. Deep-learning methods such as convolutional neural networks (CNNs) consist of a number of convolutional and sub-sampling layers producing a fully connected layer, which in turn can be used as a robust feature extractor and classifier module.

There are various types of classifier ways to solve Classification problem some are as follows:

- Logistic Regression
- Decision tree Algorithm
- Random Forest Algorithm
- Naive Bayes Classifier
- K-Nearest Neighbor

- Artificial Neural Network

Above mentioned each have their strengths & weaknesses out of which some are Binary classifier types & some are Multiclass classifier types. Out of which Artificial neural network on both way it can be either Binary classifier type as well as Multiclass classifier types.

II. Literature survey :

The text classification domain, victimisation completely different approaches and introducing some new techniques during this field. The study [9] works on app review classification victimisation ensemble algorithms and techniques. The dataset employed in the study was antecedently examined in [3], the dataset contains reviews from Apple's app store and also the Google Play app store. within the study [9], the authors used NB, SVM, LR, and neural network (NN) in varied mixtures for classification. They designed 3 ensemble algorithms A, B, and C. In ensemble A, four classifiers, NB, SVM, LR, and NN, were classified for final prediction; in ensemble B, 3 classifiers, SVM, LR, and NN, were classified, and in ensemble C, the 2 classifiers NB and SVM were classified. the simplest performers from these individual and ensembles algorithms were LR and NN. This study additionally used ensemble models, like RF and AC, that work with numbers of base learners (decision trees) to create final predictions. In another analysis [4], text analysis was performed for mobile app feature requests. They designed MARA (mobile app review analyzer), a example for automatic retrieval of mobile app feature requests from on-line reviews. MARA takes review content as input for feature request mining. The feature request mining rule uses a collection of linguistic rules, that are outlined for supporting the identification of sentences that indicate such requests. The linear discriminant analyser model was accustomed determine topics which will be related to these requests in user reviews.

They used true positive (TP), false positive (FP), true negative (TN), false negative (FN), precision, recall, and Matthews correlation coefficient as evaluation metrics to check the accuracy of the algorithm. Researchers perform analysis on app reviews to facilitate app developers in finding out whether their customers are happy are not, which is also a goal of this study. In study [10], researchers tried to help mobile app developers by performing analysis on user reviews to categorize information that is important for app maintenance and evolution. For classification purposes, they deduced a taxonomy of user review categories that are relevant to app maintenance. The authors merged three techniques, natural language processing, text analysis, and sentiment analysis. By merging these techniques, they achieved desirable results in terms of precision and recall (Precision Score 74% and Recall Score 73%). They also applied these techniques individually to classify user reviews. In another study [11], the authors tried to extract the values of comparison scores of sentiment reviews using different feature extraction techniques, such as word2vec, word2doc, and TF-IDF, with SVM, NB, and decision tree algorithms. In study [11], the authors used grid search algorithms for parameter optimization of machine learning algorithms and feature extraction methods.

In the paper of Ensemble Methods for App Review Classification: An Approach for Software Evolution the researchers are Emitza Guzman, Muhammad El-Halaby, Bernd Bruegge [1] have categorized in below given manner

The definition of taxonomy relies on the classes found in a very previous study [4] that manually analyzed the content of app store user reviews. For the event of their taxonomy, 2 of the authors manually annotated the relevancy to software system evolution of every antecedently outlined class. Overall, nine of the first classes were thought-about relevant for software system evolution. classes were deemed as necessary for software system evolution once they gave info concerning aspects of the app that required to be improved or enforced. in addition, classes that highlighted the app options or practicality that satisfy users were conjointly contemplated as relevant to software system evolution as a result of they thought-about that this info notifies developers concerning aspects of the app that area unit necessary for users and concerning options that area unit being actively used [14].Tend to thought-about general praise and criticism as classes relevant to software system evolution as a result of they provide info concerning the user acceptance and this information would possibly have an effect on software system evolution. They renamed a number of the first classes into terms they thought-about additional descriptive and changed a number of the previous definitions for higher clarity throughout the annotation of truth set The taxonomy they have arrived at consists of the following 7 categories:

- Bug report: Reviews that report a problem, such as faulty behavior of the application or of a specific feature.
- Feature strength: Reviews that identify an aspect about an existing feature that users are satisfied with.
- Feature shortcoming: Reviews that identify an aspect about an existing feature that users are unsatisfied with.
- User request: Reviews that ask for a missing feature, functionality or content, as well as reviews that ask for the improvement of an existing feature.
- Praise: Reviews where users express general appreciation with the application. It focuses on general judgment, unlike feature strength which emphasizes on the positive feedback about a specific feature.
- Complaint: Reviews where users express general dissatisfaction with the application. In contrast with feature shortcoming which focuses on the negative feedback about a specific existing feature, general complaint concentrates on general judgment.
- Usage scenario: Reviews where users describe workarounds, use cases and scenarios involving the app.

They compared the performance of four different classification algorithms and their ensembles. More concretely, we compared the performance of Naive Bayes, Support Vector Machines (SVMs), Logistic Regression and Neural Networks, as well as the performance of combinations of the predictions of these classifiers. The choice of the classifiers for our experiment was motivated by the effectiveness of the algorithms when categorizing text.

They found that Overall, the Logistic Regression and Neural Network classifiers showed a better precision than the Naive Bayes and SVM models. Furthermore, the Neural Network model had the highest F-measure average among all individual classifiers, whereas the SVM and Neural Network models had the highest recall values.

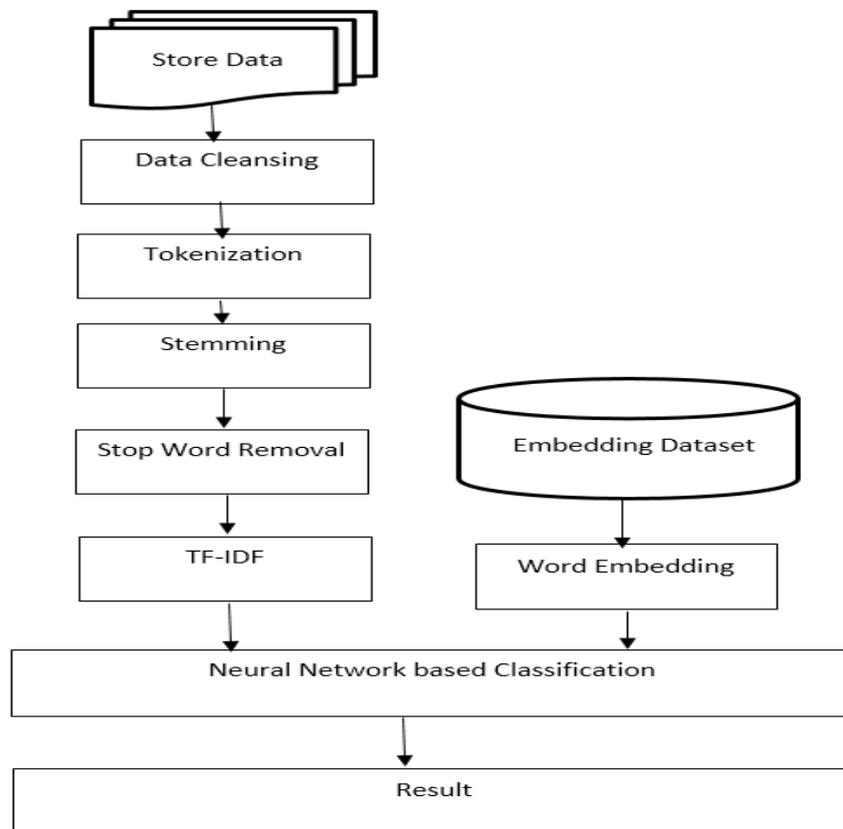
In this analysis the researchers had develop the MARA [15] could be a image developed to mine for and retrieve feature requests from on-line reviews of mobile apps. The system is meant to : 1) retrieve all the reviews obtainable for associate degree app (Review retrieval), 2) mine the content of the reviews for distinctive sentences or fragments of sentences expressing feature requests (Feature requests mining), 3) summarize such content (Feature requests summarization), and 4) gift it during a easy manner (Feature requests visualization). throughout the review retrieval part, an internet crawler extracts the page sources that compose the reviews of a given app (raw reviews) and parses their content. Of interest to the current work is that the actual content of reviews, however data related to every review is additionally collected for additional analysis. Such meta-data includes the date the review was announce, the rating the user gave, the device associated to the review, the version of the app utilized by the user, and also the title the user associated the review with. additional analysis may embrace a written account analysis of the evolution of feature requests for apps. each the review's content and also the review's meta-data area unit keep, the content being normalized to scale back the noise within the final results. in addition, the reviews content is split into sentences. For that, they used LangPipe, a language process tool that supports sentence rendering.

Doc2vec as a simple extension to Word2vec. Many types of texts used this feature. They are word n-gram, sentence, paragraphs or document [7]. Refer to the embedding of the word sequence; we used the term document embedding that supported by Doc2vec. The classification method was done with 3 classifier NB (Naive Bayes), SVM (Support Vector Machine), and DT (Decision Tree). SVM can be used to create the highest accuracy results in text classification problems [1]. The NB has high accuracy than other followed by the DT classifier. This research dedicated to select the best feature extraction and choosing the best model for multiclass classification by comparing the TF-IDF, Word2vec, Doc2vec feature extraction and increase the accuracy using hyperparameter optimization. Hyperparameter optimization used to search the best parameter that produces the best classification accuracy.

III. Proposed Methodology :

The review dataset will need to be pre-processed and then tokenized. In pre-processing, all the stop words will be removed. After that the root words of the remaining word will then be found. This is called stemming. Then, the tokenization is performed. Then the further processing is done as follows viz.

- 1) Pre-processing and Tokenization of reviews data
- 2) Further TF-IDF is applied
- 3) Word embedding is used
- 4) Neural Network classification is applied



Pre-processing & Tokenization

The reviews obtained from the user first need to be preprocessed first for to do that there are various preprocessing techniques needed to be applied on the reviews so that by applying these techniques we get a clean data for further processing for review classification, some of preprocessing techniques are as follows Tokenization, punctuation removal, lower-case conversion, removal of numeric values, and stopword removal. Finally, the stemming technique was used to get the root-form of each feature in the reviews.

TF-IDF

TF-IDF which stands for **Term Frequency – Inverse Document Frequency**. It is one of the most important techniques used for information retrieval to represent how important a specific word or phrase is to be given for document. Let's take an example, we have a string or **Bag of Words (BOW)** and we have to extract information from it, then we can use this approach, TF-IDF do not convert directly raw data into useful features. Firstly, it converts raw strings or dataset into vectors and each word has its own vector. Then we'll use a particular technique for retrieving the feature like Cosine Similarity which works on vectors, etc. As we know, we can't directly pass the string to our model. So, TF-IDF provides numeric values of the entire document for us.

Word Embedding

Word embeddings focus on learning distributed vector representations of words by leveraging the contextual information in large corpora using neural network architectures. For each word (target word) to be refined, the nearest neighbor ranking is first applied to select a set of words both semantically and sentimentally similar to the target word for refinement. This can be accomplished by first calculating the semantic similarity between the target word and the other words in the lexicon based on the cosine distance of their pre-trained vectors. The top-k most similar words are then selected as the nearest neighbors, and ranked according to their intensity scores. Based on weights assigned to these nearest neighbors, the intensity score is calculated.

Neural Network classification

Classification involves predicting which class an item belongs to. Some classifiers are binary, resulting in a yes/no decision. Others are multi-class, able to categorize an item into one of several categories. A neural network is a series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. So the above output of phases is feed to neural network classifier to obtain the classification. Artificial Neural Network is work in both ways either as a Binary classifier or as a Multi-class classifier. There are different types of Feed-forward neural network & back-propagation neural network their involving the Input layer, Hidden layer, Output layer.

IV. Conclusion :

Classification of review can be done by using different methods all this methods in deep learning are mentioned above in the proposed methodology. Used various factors in studying all this methodology the proposed methodology is resulting better. There are various feature scope to this studies. As deep learning is the new possibilities for problem to solve in a better ways

References

- [1] E. Guzman, M. El-Haliby, and B. Bruegge, "Ensemble methods for app review classification: An approach for software evolution (N)," in Proc. 30th IEEE/ACM Int. Conf. Automat. Softw. Eng. (ASE), Nov. 2015, pp. 771–776.
- [2] S. Panichella, A. Di Sorbo, E. Guzman, C. A. Visaggio, G. Canfora, and H. C. Gall, "How can I improve my app? Classifying user reviews for software maintenance and evolution," in Proc. IEEE Int. Conf. Softw. Maintenance Evol. (ICSME), Sep. 2015, pp. 281–290.
- [3] S. M. Isa, R. Suwandi, and Y. P. Andrian, *Optimizing the Hyperparameter of Feature Extraction and Machine Learning Classification Algorithms*. London, U.K.: The Science and Information Organization, 2019.
- [4] F. Rustam, I. Ashraf, A. Mehmood, S. Ullah, and G. Choi, "Tweets classification on the base of sentiments for US airline companies," *Entropy*, vol. 21, no. 11, p. 1078, Nov. 2019.
- [5] V. Svetnik, A. Liaw, C. Tong, J. C. Culberson, R. P. Sheridan, and B. P. Feuston, "Random forest: A classification and regression tool for compound classification and QSAR modeling," *J. Chem. Inf. Comput. Sci.*, vol. 43, no. 6, pp. 1947–1958, Nov. 2003.
- [6] F. F. Bocca and L. H. A. Rodrigues, "The effect of tuning, feature engineering, and feature selection in data mining applied to rainfed sugarcane yield modelling," *Comput. Electron. Agricult.*, vol. 128, pp. 67–76, Oct. 2016. Reinald Kim Amplayo and Seung-won Hwang, "Aspect Sentiment Model for Micro Reviews", 2017 IEEE International Conference on Data Mining.
- [7] Kim Schouten and Flavius Frasincar, "Survey on Aspect-Level Sentiment Analysis", *IEEE Transactions on Knowledge and Data Engineering*, VOL. 28, NO. 3, MARCH 2016.
- [8] M. Hu and B. Liu, "Mining opinion features in customer reviews," in Proc. 19th Nat. Conf. Artif. Intell., 2004, pp. 755–760.
- [9] M. Hu and B. Liu, "Mining and summarizing customer reviews," in Proc. 10th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2004, pp. 168–177.
- [10] C. Long, J. Zhang, and X. Zhut, "A review selection approach for accurate feature rating estimation," in Proc. 23rd Int. Conf. Comput. Linguistics, 2010, pp. 766–774.
- [11] Z. Hai, K. Chang, and J.-J. Kim, "Implicit feature identification via co-occurrence association rule mining," in Proc. 12th Int. Conf. Comput. Linguistics Intell. Text Process. 2011, vol. 6608, pp. 393–404.

- [12] B. Liu, M. Hu, and J. Cheng, "Opinion observer: Analyzing and comparing opinions on the web," in Proc. 14th Int. Conf. World Wide Web, 2005, pp. 342–351.
- [13] Z. Li, M. Zhang, S. Ma, B. Zhou, and Y. Sun, "Automatic extraction for product feature words from comments on the web," in Proc. 5th Asia Inf. Retrieval Symp. Inf. Retrieval Technol., 2009, pp. 112–123.
- [14] Y. Zhao, B. Qin, S. Hu, and T. Liu, "Generalizing syntactic structures for product attribute candidate extraction," in Proc. Conf. North Am. Chapter Assoc. Comput. Linguistics: Human Lang. Technol., 2010, pp. 377–380.
- [15] G. Qiu, B. Liu, J. Bu, and C. Chen, "Expanding domain sentiment lexicon through double propagation," in Proc. 21st Int. Joint Conf. Artif. Intell., 2009, pp. 1199–1204

