IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Review Classification by using Neural Network & TF-IDF

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

[1]Student ,[2]Assistant Professor

[1]Computer Science & Engineering, [2]Computer Science & Engineering

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

As in modern world data is most essential thing for people as well as to the tech jints. The more the data the more the information this information can be in various form images, videos, text, documents, archives etc. As the internet speed & connectivity increases the data will be flowing with the greater speed as compared to the previous times. This data can be used to analyse the patterns in it so by recognizing that this can be used to solve the problem or for drawing conclusions from data. Suppose consider to the reviews: that comes in our day to day life it a thoughtful way of expressing feelings or someone view on anything. If we consider it in the forms of consumer & product provider then this reviews play an important part for both the product provider will get to know how my product is doing & what are drawbacks of my product that needs to be overcome in the upcoming time & consumer can express his opinion on the product with the help of reviews. That same reviews provided by consumers will help other consumer whether to go for this product. It is same as getting opinion on product from bunch of people. So this huge amount of reviews data needs to be classified in such a way so that form which conclusions can be drawn which will help the product provider with their next releases. In that releases they would try to rectify the mistakes that happened previously.

As we know in our day to day life when we need to buy new product or use new services that time we try to get knowledge about that from the existing user like friend, family, relative, etc. so in this process we are trying to understand whether users are satisfied what are pitfalls of using what are usefulness of this. we try to analyze or exciting user try to analyze & explain it to us that same can be done for fewer reviews but for huge amount of reviews it's not possible to categorized & analyze the reviews, so for that the companies use various types of different techniques. That techniques can be various methodologies. Which are as follows

1. Artificial Neural Network
2. Machine learning
3. Various classifications algorithms

App stores square measure digital distribution platforms that enable users to transfer and rate mobile apps. Notable distribution platforms for mobile devices embody Apple and Android app stores, during which users will comment and write reviews of the mobile apps they're victimization. These reviews function a communication between developers and users wherever users will offer relevant data to guide app developers in accomplishing many package maintenance and evolution tasks, like the implementation of latest options, bug fixing, or the development of existing options or functionalities. App developers pay appreciable effort in aggregation and exploiting user feedback to boost user satisfaction. Previous work [10] has shown that more or less one third of the knowledge contained in user reviews is useful for developers. However, processing, analysing and choosing helpful user feedback presents many challenges. 1st of all, app stores embody a considerable body of reviews, which needs an oversized quantity of effort to manually analyse and method. which well-liked apps, like Facebook, received on the average four, thousands of reviews per day. to boot, users typically offer their feedback in sort of unstructured text that's tough to break down and analyze. Thus, developers and analysts got to browse a large amount of textual data to become aware of the comments and needs of their users [10]. In addition, the quality of reviews varies greatly, from useful reviews providing ideas for improvement or describing specific issues to generic praises and complaints (e.g. “You have to be stupid to program this app”, “I love it!”, “this app is useless”).

Binary classifier type as well as Multiclass classifier types.

II. Literature survey :

The text classification domain, victimization completely different approaches and introducing some new techniques during this field. The study [9] works on app review classification victimization 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].Ttend 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 7 categories

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.

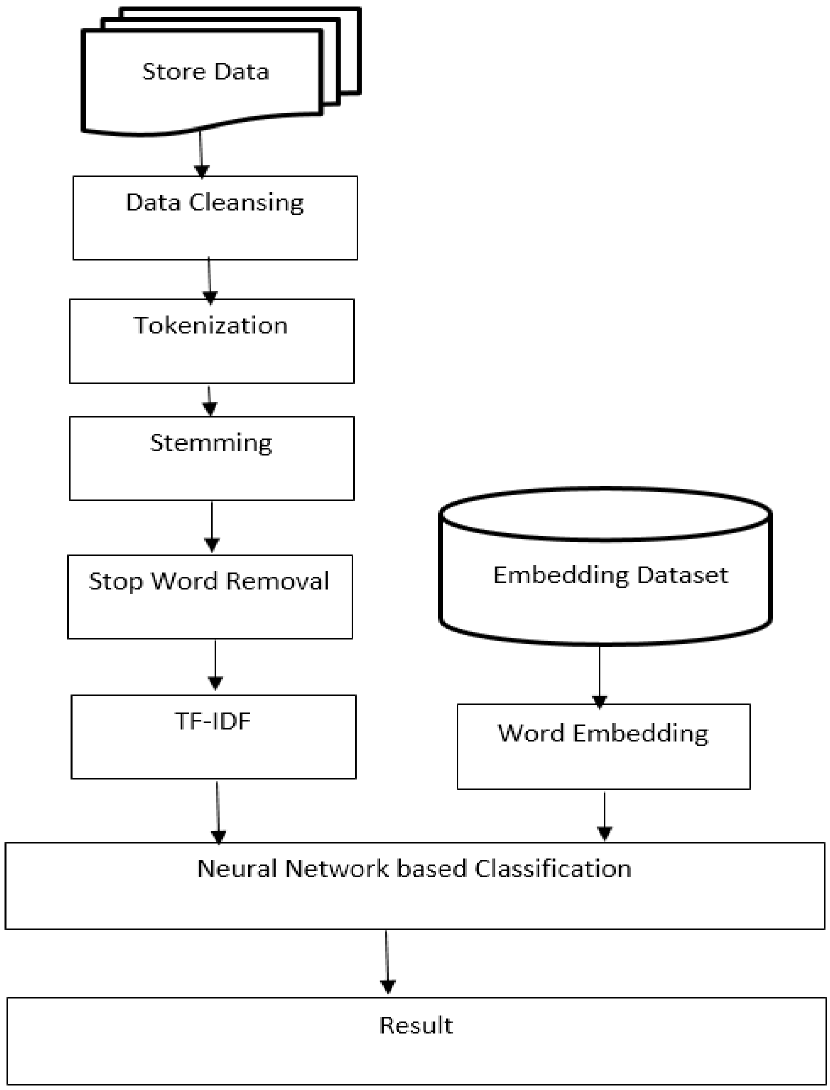
They found that Overall, the Logistic Regression and Neural Network classifiers showed a better precision than the Naive Bayes and SVM models. Furthermore, the Neural Network model had the highest F-measure average among all individual classifers, whereas the SVM and Neural Network models had the highest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

III. Proposed Methodology :

Here to solve the problem of classification the new system is proposed this system uses the Neural network, TFIDF, word embedding this techniques are used over here in this system the following steps will be followed

1. Data Cleaning
2. TF-IDF
3. Neural Network Classifier
4. Word Embedding



1. Data Cleaning

Data Cleaning means the process of identifying the incorrect, incomplete, inaccurate, irrelevant or missing part of the data and then modifying, replacing or deleting them according to the necessity.

1. Get Rid of Extra Spaces
2. Select and Treat All Blank Cells
3. Convert Numbers Stored as Text into Numbers
4. Remove Duplicates
5. Highlight Errors
6. Change Text to Lower/Upper/Proper Case
7. Spell Check
8. Delete all Formatting
9. Remove Stop words

This are all the data cleaning processes that can be applied on data to clean the data the data that gathered form the data set will be inconsistent which needs to be converted in proper format suppose consider where in statement the special symbols, emoji's, or extra data that is irrerevalent needs to be removed here so here in data cleaning process that part will be removed.

In data cleaning the data needs to be in on form so all the data is converted into on case format i.e. into the smaller case.

Here the stop word removal is done in this stage suppose consider the words that are irrelevant to categorization. That time this words are removed.

	A	B	C	D	E	
1	id	title	rating	reviews_c	tagline	pricing_hint
2	9e4748a9-7eda-4814-83b6-0537d44152b1	Panda Language Translate	4.7	379	Translate your store into multiple languages	7-day free trial
3	d1476138-a608-4bb9-8d39-b30f3ca7617d	Instant Brand Page	4.9	13	A-Z Brand Index Page and Favourites Slider	10-day free trial
4	d6e49a3c-2f9f-4bfa-8c26-5d024faf2241	FAQ Accordion | Help Center	4.5	202	FAQ page, FAQ accordion menu for product Info & refund policy	
5	0ef0087f-3ae5-4dbc-84e0-193b576d82ed	Promote Me | Many apps in one	4.9	18	Spin Wheel,Currency Converter,Quick ATC Buttons & 18	10-day free trial
6	7aac2a1f-ff03-4f38-aeb7-7619403a6f05	Instalify	0	0	Supercharge Your Mobile App Installs	7-day free trial
7	c13bfb7f-8b5a-40c6-a338-dbdec5cfd130	EASY product feed	5	3	Get your products listed on Google Shopping the EASY v	14-day free trial
8	6a71634f-f94f-498d-8713-d2b01ca90917	PosBill Connect	0	0	Connect your Webshop to the solutions of PosBill.	
9	da50f0bf-d116-46a0-b0c1-0d90c05d8ffe	Bulk Fulfill	4.9	76	Reduce wasted time - Automatically sync your tracking numbers	
10	89735c3c-4d25-40f2-8b54-150f59cfb099	Show Price in BTC	5	2	A better currency converter app. Premium features for free!	
11	3c89c108-4858-4b07-893a-460ea5a0d91a	POKY â€' Product Importer	4.9	70	Copy / Import products from any Shop store with a singl	2-day free trial
12	81de0a93-e81f-4ac3-88d7-7ff114b6ac15	Actindo Unified Commerce Suite	0	0	Omni-Channel, Data Hub & Enterprise Commerce ERP Plattform	
13	9e6d517f-01ec-42c8-a660-7bef81b741da	inSites: Insights Bar	5	2	Do you know the Conversion Rate of every Product?	14-day free trial
14	04aa830a-7517-42fa-be5e-84d449a85f65	Product Warnings & Pop Ups	5	2	Adds custom warning labels and pop ups to products pages	
15	1264971d-c9eb-46ec-8acf-cb5469cae1c1	Pinterest Feed Ninja	0	0	Widget to display Pinterest Profile, Board and Pin	14-day free trial
16	c382563a-9cd7-4b62-ae17-50c4073e47f3	DataChamp Excel/CSV Exports	5	9	Easily export orders and more. Scheduled. Customized CSV/Excel	
17	733fd831-b85f-471f-853f-ab43975661c9	Rotating Announcement Bar	4.9	64	Free Shipping Bar, Countdown bar, Promotion Bar for U	7-day free trial
18	b37596a5-6aa4-4a4a-8997-b20560a5ef15	Simple Free Shipping Bar	0	0	Show your free shipping bar to increase sales	
19	d07745e1-cbb3-4cbb-b6d8-9316e8319f5f	Maisie AI Chatbots	4	4	Chatbots for automated sales, marketing & support	30-day free trial
20	3f2139a1-1c1d-4c77-8dda-bfd7e6e5bf13	LinkHaitao Affiliate Network	2	1	Grow sales in China with China's largest affiliate network	
21	65a1d3c7-f878-4aa7-a28d-82d1897e5cd7	Mighty PO Box Blocker	0	0	Warn customers against entering in PO Box addresses.	7-day free trial
22	aad768ae-c00d-4391-8e0c-84694c21ffb2	New: Ad Targeting Engine	5	3	Boost performance with customer intelligence & smart	14-day free trial
23	7932e175-8298-411e-b0b8-01d36475458b	Easy Catalogs â€' Unlimited	5	3	Create unlimited PDF catalogs saved on cloud, print or s	3-day free trial

Fig. 1 Review Excel Screenshot

	A	B	C	D	E	F	G	H
1	app_id	rating	posted_at	body	helpful_c	develope	developer_reply	
2	b1da53a4-0474-4700-9620-bf386bc033fb	5	August 6, 2020	Great and super fast customer service! Highly customizable app with good, easy to decipher analytics.	0			
3	b1da53a4-0474-4700-9620-bf386bc033fb	5	August 4, 2020	Still setting up my store, and after initially paying for another search app, I made the decision to try Instant Search. Their aesthetic really suits my type of store, and the custom settings provide you with so many options. I am new to all of this, and am so happy I found this app.	0			
4	b1da53a4-0474-4700-9620-bf386bc033fb	5	August 4, 2020	This is an excellent search app, which they have well designed functions to help merchant to uplift customer's online searching experience. Furthermore, they have one of the best customer services supports to provide valuable advise, as well as willing to linkup other App owner to resolve my challenges. Well Done ED!	0			
5	b1da53a4-0474-4700-9620-bf386bc033fb	5	July 30, 2020	A+, great great great customer service! thanks to Matt for the help. We use your service for 3 sites and recommend it to others.	0			
6	b1da53a4-0474-4700-9620-bf386bc033fb	5	July 28, 2020	I'm begginig to use this app, the search engine is intuitive and easy, Its helping me with very complex products that have many variants. Support has been very good.	0			
7	b1da53a4-0474-4700-9620-bf386bc033fb	5	July 6, 2020	Great work guys. Easy app to install and use. Great customer service as well for app questions. 10/10	0			

Fig. 2 App Excel screenshot

Here for proposed system the above data set is used which is from kaggle i.e of shopify app store. So on this data set the data cleaning process will be done & after that step by step procedure.

2. TF-IDF

tf-idf stands for Term frequency-inverse document frequency. The tf-idf weight may be a weight usually utilized in info retrieval and text mining. Variations of the tf-idf weight theme area unit usually utilized by search engines in grading and ranking a document's connectedness given a question . This weight may be an applied mathematics live wont to appraise however necessary a word is to a document in an exceedingly assortment or corpus. The importance will increase proportionately to the amount of times a word seems within the document however is offset by the frequency of the word within the corpus (data-set).

Here during this methodology what quantity a word is vital that finding is completed during this phase. From that frequency of word is given.

This methods is applied on information that's acquire once the info improvement process. By applying this method we tend to we are going to get the necessary words & their frequency.

TF-IDF may be a applied mathematics live that evaluates however relevant a word is to a document in an exceedingly assortment of documents. this is often done by multiplying 2 metrics: what number times a word seems in an exceedingly document, and also the inverse document frequency of the word across a collection of documents

3. Neural network classifier :

Neural networks area unit loosely representative of the human brain learning. a man-made Neural Network consists of Neurons that successively area unit chargeable for making layers. These Neurons also are called tuned parameters.

The output from every layer is passed on to consequent layer. There area unit totally different nonlinear activation functions to every layer that helps within the learning method and also the output of every layer. The output layer is additionally called terminal neurons.

The weights related to the neurons and that area unit chargeable for the general predictions area unit updated on every epoch. The training rate is optimised victimisation varied optimisers. Every Neural Network is given a value operate that is minimised because the learning continues. The simplest weights area unit then used on that the price operate is giving the simplest results

A neuron in an artificial neural network is

1. A set of input values (xi) and associated weights (wi).

2. A function (g) that sums the weights and maps the results to an output (y).

Neurons area unit organized into layers: input, hidden and output. The input layer consists not of full neurons, however rather consists merely of the record's values that area unit inputs to ensuing layer of neurons. ensuing layer is that the hidden layer. many hidden layers will exist in one neural network. The ultimate layer is that the output layer, wherever there's one node for every category. One sweep forward through the network ends up in the assignment of a price to every output node, and also the record is assigned to the category node with the best worth.

Training an Artificial Neural Network

In the coaching part, the right category for every record is thought (termed supervised training), and also the output nodes may be assigned correct values -- one for the node like the right category, and zero for the others. (In apply, higher results are found exploitation values of zero.9 and 0.1, severally.) It's so doable to match the network's calculated values for the output nodes to those correct values, and calculate a slip-up term for every node (the Delta rule). These error terms square measure then wont to regulate the weights within the hidden layers in order that, hopefully, throughout ensuing iteration the output values are nearer to the right values

The Iterative Learning Process

A key feature of neural networks is Associate in Nursing unvaried learning method within which records (rows) area unit conferred to the network one at a time, and therefore the weights related to the input values area unit adjusted every time. On balance cases area unit conferred, the method is commonly recurrent. Throughout this learning part, the network trains by adjusting the weights to predict the proper category label of input samples. Blessings of neural networks embrace their high tolerance to rip-roaring information, furthermore as their ability to classify patterns on that they need not been trained. The foremost common neural network algorithmic program is that the back-propagation algorithmic program planned within the Nineteen Eighties.

Once a network has been structured for a specific application, that network is prepared to be trained. To begin this method, the initial weights area unit chosen arbitrarily. Then the coaching (learning) begins. The network processes the records within the coaching Set one at a time, victimization the weights and functions within the hidden layers, then compares the ensuing outputs against the specified outputs. Errors area unit then propagated back through the system, inflicting the system to regulate the weights for application to following record. This method happens repeatedly because the weights area unit tweaked. throughout the coaching of a network, an equivalent set of information is processed repeatedly because the affiliation weights area unit frequently refined.

Note that some networks ne'er learn. this might be as a result of the input file doesn't contain the particular data from that the specified output comes. Networks conjointly won't converge if there's not enough information to modify complete learning. Ideally, there ought to be enough information on the market to make a Validation Set.

Feed forward, Back-Propagation

The feed forward, back-propagation design was developed within the early Seventies. This co-development was the results of a proliferation of articles and talks at varied conferences that stirred up the whole business. Currently, this synergistically developed back-propagation design is that the hottest model for complicated, multi-layered networks. Its greatest strength is in non-linear solutions to ill-defined issues.

The typical back-propagation network has Associate in nursing input layer, Associate in Nursing output layer, and a minimum of one hidden layer. there's no theoretical limit on the quantity of hidden layers however generally there square measure only 1 or 2. Some studies have shown that the entire range of layers required to unravel issues of any complexness is 5 (one input layer, 3 hidden layers Associate in Nursingd an output layer). every layer is absolutely connected to the succeeding layer.

The coaching method unremarkably uses some variant of the Delta Rule, that starts with the calculated distinction between the particular outputs and therefore the desired outputs. mistreatment this error, affiliation weights square measure inflated in proportion to the error times, that square measure a scaling issue for world accuracy. this implies that the inputs, the output, and therefore the desired output all should be gift at identical process part. the foremost complicated a part of this algorithmic rule is determinative that input contributed the foremost to Associate in Nursing incorrect output and the way should the input be changed to correct the error. (An inactive node wouldn't contribute to the error and would haven't any ought to modification its weights.) to unravel this downside, coaching inputs square measure applied to the input layer of the network, and desired outputs square measure compared at the output layer. throughout the training method, a forward sweep is formed through the network, and therefore the output of every part is computed by layer. The distinction between the output of the ultimate layer and therefore the desired output is back-propagated to the previous layer(s), typically changed by the spinoff of the transfer perform. The affiliation weights square measure unremarkably adjusted mistreatment the Delta Rule. This method yield for the previous layer(s) till the input layer is reached.

Structuring the Network

The variety of layers and therefore the number of process parts per layer are vital selections. To a feedforward, back-propagation topology, these parameters also are the foremost ethereal -- they're the art of the network designer. there's no quantitative answer to the layout of the network for any specific application. There are solely general rules picked up over time and followed by most researchers and engineers applying whereas this design to their issues.

Rule One: because the complexness within the relationship between the input file and therefore the desired output will increase, the amount of the process parts within the hidden layer ought to conjointly increase.

Rule Two: If the method being sculpturesque is divisible into multiple stages, then extra hidden layer(s) is also needed. If the method isn't divisible into stages, then extra layers might merely modify acquisition of the coaching set, and not a real general resolution.

Rule Three: the quantity of coaching Set offered sets Associate in Nursing edge for the amount of process parts within the hidden layer(s). To calculate this edge, use the variety the amount the quantity} of cases within the coaching Set and divide that number by the add of the amount of nodes within the input and output layers within the network. Then divide that result once more by a scaling issue between 5 and 10. Larger scaling factors are used for comparatively less creaky information. If too several artificial neurons are used the coaching Set are memorized, not generalized, and therefore the network are useless on new information sets.

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. Andrean, 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