Odd Posts Identification through the Vocabulary by Semantic Sentiment Analysis Using Machine Learning Algorithm

Dr. Mohammed Ali Alzahrani

College of Computing and Information Technology, Taif University, Saudi Arabia

marzahrani@tu.edu.sa

Abstract— In this work, the importance to detect the odd posts is considered to understand. This work explains the odd posts, and the methods using machine learning algorithms are well explained. In the first part, a survey is conducted on what runs on the internet and why to eliminate the odd posts is required. After explaining the importance of the odd posts as the data is exponentially increasing, to eliminate the odd posts various techniques are used but here are some of the best, TF-IDF is used as the clustering technique, SVM used to eliminate the odd posts and observed as a better solution. However, to eliminate the odd posts using machine learning algorithms RNN produces good results, as RNN is feasible to use for the large scale dataset. The sentiment analysis using machine learning approaches is used to obtain the progressive results which aim to eliminate the odd posts from the data, as the current demand and major source of the data are social media platforms. The accuracy is improved, and the implementation of the model is done on the dataset of the tweets which includes negative, positive and neutral tweets.

Keywords— Odd posts identification, semantic sentiment analysis, machine learning, TF-IDF, Support vector machine, Recurrent neural network.

_____

Introduction:

The last decade had increased the use of the internet data, and the internet-based posts and especially the textual data

increased. The exponential increase in the textual data, the huge issues regarding fair communication, arise in

meanwhile machine learning plays its part, by developing such methods to analyze the textual data. The content

requires fulfilling the term and remaining ethically strong enough to remain in a good category. However, this work

explains to develop the dictionary or the vocabulary to help ease in detecting the odd words first. The current era and

the coming era, directly belong to machine learning. As Bill Gates explained that machine learning may worth ten Microsoft, similarly, Tony Tether Director DARPA explains that machine learning will be next internet. The first spam post was carried out in 1978, as an email of digital marketing purposes, which is still the part of internet marketing. To eliminate the odd posts from the existing posts machine learning algorithms, techniques and models are playing a vital role.

Furthermore, in this work, the analysis between different techniques to conclude a better solution to this problem is also observed. However, from the last two decades, the deep learning models using supervised techniques minimized this issue which is also observed in this work, but still, the problem requires eliminating. The regression models were used two decades before but from the last decade, the classification is also carried out, to eliminate the odd posts or the spam classification. In this survey, the different techniques were analyzed to check which techniques perform better than the other one. Previously the classification was done through some un-supervised and supervised models, now the deep learning neural language is taking place and classifying and producing the results more accurately. The most important techniques used to implement were concluded to eliminate the odd posts and the spam words used to initiate such posts helped in a learning problem. To eliminate the odd posts serious development methods require to integrate through the already developed systems to overcome, spam posts.

Literature Review:
The progressive use of the internet is leading to some of the critical issues on the World Wide Web. The spam posts are the part from the last decade and increasing its number till the year twenty seventeen. There are different types of odd posts some of the odd posts are the negative comments on the social media that might be done for the marketing purpose. Some critical posts also tend towards the odd behavior, intentionally posting the same comment or the message (Rădulescu, Dinsoreanu, & Potolea (2014) pp. 29-35). The very first face of the odd post was the unwanted emails, which were also known as spam emails, between this the other form of the odd posts are the marketing adds on the social media websites, the next phase of the odd posts is the one if you get spam comments, messages, posts on the social media websites, on the professional websites or even in the learning websites (Ruihai Dong, Markus Schaal and Barry Smyth(2013) pp. 1310-1317).

The odd posts are now the part of the internet World Wide Web this directly disturbs the intentions of the working users, this impact in the negative image of the internet, this also impacts the negative image of those social media

platforms. To generate a positive impact on the application, website or any mean of communication, the odd posts must be eliminated properly, to eliminate the odd posts different techniques and the methods are in use (Muhammad, Qamar, & Noureen, (2018, November) pp. 435-443). To detect the odd posts different types of analysis are conducted on the data to develop the smart solutions. However, highly professional websites are developing their whole systems and also developing their models to eliminate such activities, but the problem is still the part and required to eliminate through the proper use of the machine learning models.

Artificial intelligence and Machine learning are playing a vital role, developing structured models using unsupervised, supervised and deep neural networks, to obtain the efficiency as well. While classifying the odd posts various classification algorithms are also used like the KNN, decision tree, SVM, shortest path algorithm (Mateen, Iqbal, Aleem, & Islam. (2017) pp 466-471). There are different techniques and models to identify the odd words and then to identify the odd posts, and then to eliminate them. Various classification models using K nearest neighbor to classify the spam words are used. Similarly, in the small data sets, the decision tree also provides better results, however in the supervised extraction support vector machine provides better results (Jain, Kapoor, Gulyani, & Dubey,(2019) pp 293-307). To identify the active users in the environment, which are posting messages, a support vector machine is applied to identify. To identify the odd comments, natural language processing is used to eliminate odd words.

Natural language processing using machine learning as supervised techniques proposed solution to the spam post-detection methods. Natural language processing is used to detect the spam tags and protect the spam tags by detecting such tags to achieve this first a model is required to construct (Bhatt, & Saini (2016) pp 975). Natural language processing is used to detect the spam comments, the model is prepared using a vocabulary, which first developed as a spam dictionary, using that dictionary the spam words were detected and then the weight of the comment is considered as spam. To extract features from the reviews and then to construct the decision on the reviews, to develop strong relation and meaningful extraction from the sentences a large number of features are required to extract (Green & Spezzano, (2017)).

Vector space document model is a method to eliminate the specific words, this is implemented using the learning technique of the vocabulary, which leads to learning the specific words, and similarly, this method was used to detect spam words from the Facebook comments which then used to identify the spam comments (Matthias, &

Vogeley (2019)). Support vector machine is used in supervised learning to eliminate the spam users from the professional websites, in this case, the features of the spam users were concluded and detected through the machine learning model, which lead to eliminating the spam users (Mahatab, & Jabeen, (2018, November) pp. 1-7). Furthermore, the new trend in odd posts is the spam hashtags are used to boost the images and the data which is then used in the search, to detect such hashtags and to avoid any spam hashtags data natural language processing is used(Tang, Qin, & Liu. (2015, September)).

Literature Review Matrix:

Author Date	Authors Name	Research Question	Methods: Study, types	Major Findings	Limitations Gaps	Conclusions	Implications for future research & practice
2014	Rădulescu, Dinsoreanu, & Potolea.	How to identify the odd comments?	work is of high demand, with all new techniques using NLP, to process and eliminate the spam comments.	work impact to improve the working of the social media platforms.	The methods are fine but require implementing on large scale datasets.	The work is on real-time problem which eliminates the real time issue.	The method should be applied on some large scale data or on some social media platform, to obtain the results.
2013	Ruihai Dong, Markus Schaal and Barry Smyth	How to use information in business intelligence?	To extract the reviews, the positive and the negative neural networks were observed.	In the future this work could be used in different review based websites to develop a positive sense.	Too many features are used to extract the reviews, inefficient for large scale data-set.	This work concludes a better technique to extract positive and negative reviews and to differentiate.	The work would be used in business intelligence as its next step, to make decision of products on the review based.
2016	(Bhatt, & Saini)	How to Detect the spam comments from the Facebook?	Vector space document model is used to eliminate the spam words and the comments from Facebook.	Top spam words are analyzed and detected.	Detailed explanation of the model is not represented.	A better approach to analyze the spam words using VSDM.	To develop a vocabulary using the model and then to develop a tool using the vocabulary, to eliminate odd comments.
2017	Green, & Spezzano.	Spam accounts in the professional platforms ?	Detect the spam users in Wikipedia, using different	The features were concluded from the users profile and the behavior	Features are extracted by implementing various	Machine learning framework is developed	This work is in high demand in future the method used to implement is a progressive way to

| 2018	Mahatab, & Jabeen	How to Search the spam posts using tags?	models like SVM, logistic regression, k-nearest neighbor, random forest.	is detected.	methods.	using different methods to eliminate the spam accounts.	approach this problem.
			A model is developed to eliminate the tags, using five different steps including cosine similarity, and equivalency.	As the trend of tags is taking place to overcome spam issue and odd posts tags should also be focused.	The model to check the tags may be analyzed with new models.	The tags detection or the spam posts is a good approach to analyze the tags.	To identify the spam tags this method and model can be used with some alterations.

Research Methods:

    1.   Survey to analyze the demand:

To analyses the importance of the elimination of the odd comments first, the need is to analyze the social media application. Surveying to analyze different social media applications that directly increasing the online data (luna, & Pennock(2018)). According to the survey, the one minute on the social media applications conducted and concluded the facts from the different social media applications are explained.

APPLICATION	MINUTE THREAD 2018		INTERNET THREAD 2017	
	Amount	Action	**Amount**	Action
Facebook	973000	Logins	900000	Logins
WhatsApp	38000000	Messages	29000000	Messages
Google	3700000	Search Queries	3500000	Search Queries
YouTube:	4300000	Video Viewed	4100000	Video Viewed
Playstore:	375000	Apps Downloaded	342000	Apps Downloaded
Instagram:	174000	Scrolling	46200	Scrolling
Twitter:	481000	Tweet Sent	452000	Tweet Sent
Netflix:	266000	Hours Watched	70,017	Hours Watched
Snapchat:	2400000	Snaps Created	1800000	Snaps Created
Messenger:	25000	Gif Sent	15000	Gif Sent
Email Sent	187000000	Email sent	156000000	Email sent
Tinder:	1100000	Swipes	990000	Swipes
Spent Online:	862823 $	Spent	751522	Spent
Twitch:	936073	Views	785782	Views

The change occurred in the usage of the internet in one minute, is so clear and explains the demand to develop a system to analyses the odd posts on the internet.
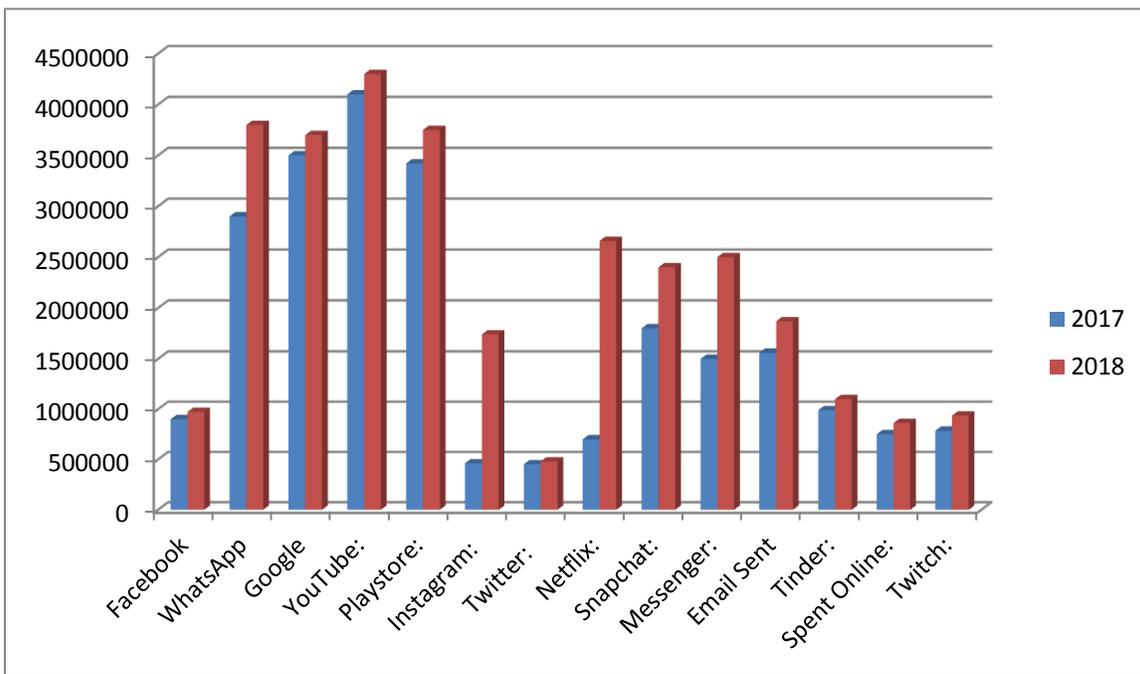
The analysis conducted on the one minute on the internet clears the question regarding why to eliminate the odd posts, the user demands the social media platforms use and an increase in the number of users, is observed (Kasemsap, (2019)pp-824-847). In most of the social media applications, twenty percent of the increase of users and posting a post is observed.

Analysis and Techniques:

To eliminate the odd comments from social media applications, different techniques are being in use. In the preprocessing of the feature extraction, semantic sentiment analysis holds the best results. To eliminate the odd posts from the data semantic sentiment analysis is conducted using different techniques, to improve the results as overall. In a survey conducted and the most reliable techniques were selected to overcome the issue of the odd post.

1. **Tf-IDF as clustering technique:**

   Eliminate the odd posts TF-IDF can be used is also known as term frequency-inverse document frequency, it checks the importance of the word. This technique is often used in text-mining, information retrieval and observed as a better technique else then other techniques used to compare the results it uses the weight of the words and then compares it with the vocabulary to produce high-end results.

$$TF = \frac{Number\ of\ time\ appears\ in\ document}{Total\ number\ of\ documents}$$

By using "TF" the occurrence of the word would be checked, and can be compared with the vocabulary while checking the document, this method may take time for long documents and may compare the term in less time for the short documents (Zhang, & Ge (2019, March)).

To compare that how important the term for that inverse document frequency is used, some of the words may have less importance and some may have more importance while comparing it with vocabulary, for that purpose weight of the frequent terms matters.

$$IDF = Log e^{(\frac{Total\ Number\ of\ document}{number\ of\ terms\ t\ in\ document})}$$

We can use this technique to stop odd posts, as it tackles such tasks using the document as input and considers the term frequency and inverse document frequency (González, I., & Mateos, A. (2019, September)213-223). This technique produces high results using it for odd post-elimination as it first it consider words as input and compares it with the vocabulary of the odd words, this technique process by comparing words and produces the required results. Tf-IDF is a feature selection technique and commonly used in machine learning tasks, this technique produces better results than most of the techniques.

2. **SVM as a semi-supervised technique:**

Support vector machine is used to eliminate the odd posts it is a semi-supervised technique that is used even when the data is labeled and even when the data is not labeled (Yousef, (2019)). The support vector machine model is used for the data analysis which is then used for classification. To eliminate the odd posts from the data, the support vector machine will work better, then the TF-IDF.

The basic working of the SVM is that is constructing the vector and differentiate between the data, like while taking a data set and making a decision of having odd posts or not, SVM will construct the vector between the odd posts and normal posts, by doing so support vector machine ease the classification process.

$$. a * b - c = 0$$

In this equation "a" is the normalized vector of the hyperplane similarly, "b" represents any hyperplane and the "c" is the offset of the hyperplane, the above equation satisfies the normal margin between the hyperplanes.

$$a * b - c = 1$$

In the above equation of support vector machine, the "1" explains that the hyperplane is a label with "1" anything above this boundary or on this boundary is in one class(Badal, & Franzén,(2019))..

$$a * b - c = -1$$

In this equation, the support vector machine, construct a boundary which includes anything on this boundary and below this boundary, is of the other class of hyperplane.
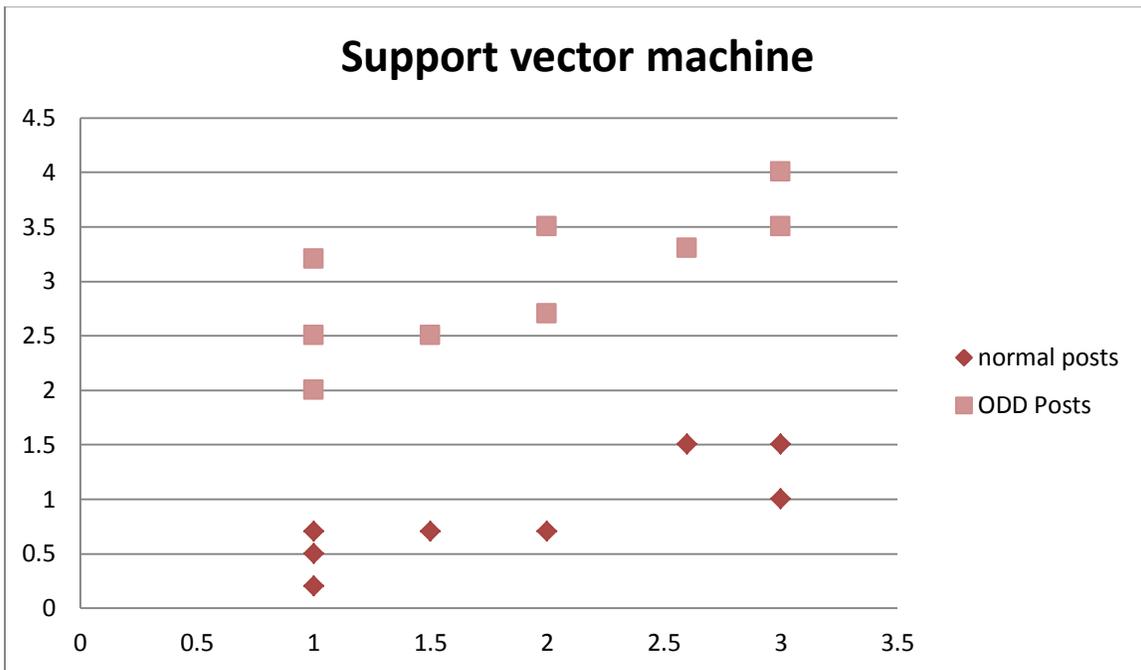
SVM constructs a vector and differentiates the data, as in the above figure the odd posts and the normal posts are in different quadrant this is because the SVM develops the vector between them and differentiates the required information (Chathavalappil, & Satyanarayana, (2019)).

The support vector machine also works better in the non-linear classification, linear classification. SVM best performs in labeled text data, however, it also performs in handwritten data and is also used in image data but as our purpose of using SVM is to implement it for textual data, and for textual odd posts detection SVM works properly.

3. Sequence to sequence odd post analysis:

Elimination of the odd posts can be carried out using a supervised learning process there are different types of neural networks, which can be used to check data sequence by sequence. Such supervised learning techniques can also perform using hierarchical representation (Mariño, Alborodo, Fredlund, & Herranz, (2019)). To obtain high-end results recurrent neural networks, artificial neural network and long short-term memory networks are used to eliminate the odd posts from the data. The recurrent neural network is a method that uses the class of the artificial neural network this connects the node to the graph of data with the temporary sequence. The temporary sequence exhibit the dynamic behavior, RNN uses a feedforward neural network, which uses internal memory which processes the sequence of inputs.

To process a large scale dataset RNN produces high-quality results even if the data is unsegmented, its process and eliminates the data, from the finite impulse and as well as infinite impulse(Khosravi, Koury, Machado, & Pabon, J. J. G. (2018)). The RNN is directly referring to the two classes of the general structure which is finite impulse and the infinite impulse the finite impulse recurrent network can be unrolled by the direct acyclic graph. However, the infinite impulse recurrent network cannot be unrolled by the direct acyclic graph.

$$Rt = W f(R(t-1)) + W(Rx) x$$

The above equation represents the simple formula of the RNN, in this "Rt" is the hidden state which is required to find, f is the function for previous hidden state, "$R(t-1)$)" this turns from the previously hidden place to the next hidden place, "Rx" represent the next move from the hidden place to the output. RNN produces better results in the elimination of the odd posts as RNN is capable of dealing with more complex tasks, and RNN is capable of finding the hidden phases and aspects from the data. Using labeled data, and for supervised learning, RNN produces better results.

4. Text sentiment analysis for the posts:

In-text sentiment analysis to eliminate the odd posts, sentiment analysis can use to classify the posts that it is an odd or a normal post. Sentiment analysis extracts the features from the data and develops a strong decision on the features extracted from the data (Petrucci, & Dragoni, (2018)). There are different approaches to use sentiment analysis, like lexical analysis, hybrid approach and the machine learning approach. Using machine learning approach is the best way to have clear outputs, it is subdivided into

different phases mainly supervised, unsupervised, and now semi-supervised techniques are also the part for the sentiment analysis using machine learning techniques.

However, in this work, unsupervised, supervised and semi-supervised, sentiment analysis techniques are used and explained to obtain high-quality results to eliminate the odd posts(Recupero, Dragoni, Buscaldi, Alam, Cambria, Kessler, & STLab, (2018)). In the unsupervised machine learning approach, TF-IDF produces high-quality results, in semi-supervised techniques and linear approach SVM produces the best results, and using a supervised approach for the sentiment analysis the RNN approach to eliminate the odd posts, by extracting features is a better way to implement the analysis.

5. **Analysis of the techniques:**

To analyze the odd posts from the data, various techniques are available, but to choose the most effective to have the high-end results a proper analysis is required to take place, to eliminate the odd posts, tf-idf can be used in the regression tasks and performs well, with a limit of having small data set, SVM performs better with the semi-supervised learning tasks with a medium scale data sets and after applying some pre-processing techniques SVM also performs better in large scale data (Kusupati, Singh, Bhatia, Kumar, Jain, & Varma. (2018)). However, RNN performs better in supervised learning decision making tasks RNN also performs better in Big data, and on large scale data.

Classifiers	TF-IDF	SVM	RNN
Type of Learning	Un-supervised learning	Semi-supervised learning	Supervised learning
Size of Data-Set	Normal data sets	Large scale data-set	Big data and large scale data set.
	Best to analyze the numeric statistical data	Best to use for the categorical data.	Useful in time series prediction, it uses the convolutional layers.

6. **Data set:**

The tweets data set is taken from the kaggle the tweets were regarding the airlines-tweets-sentiments. The dataset having 1091 records, the data-set having negative positive and neutral tweets, the tweets language were English.

7. **Tweets analysis using RNN:**

To improve the accuracy and to eliminate the odd posts, RNN is implemented to achieve the high accuracy. The elimination process of the odd posts, word embedding is used firstly using word2vec pre-processing is carried out.

93
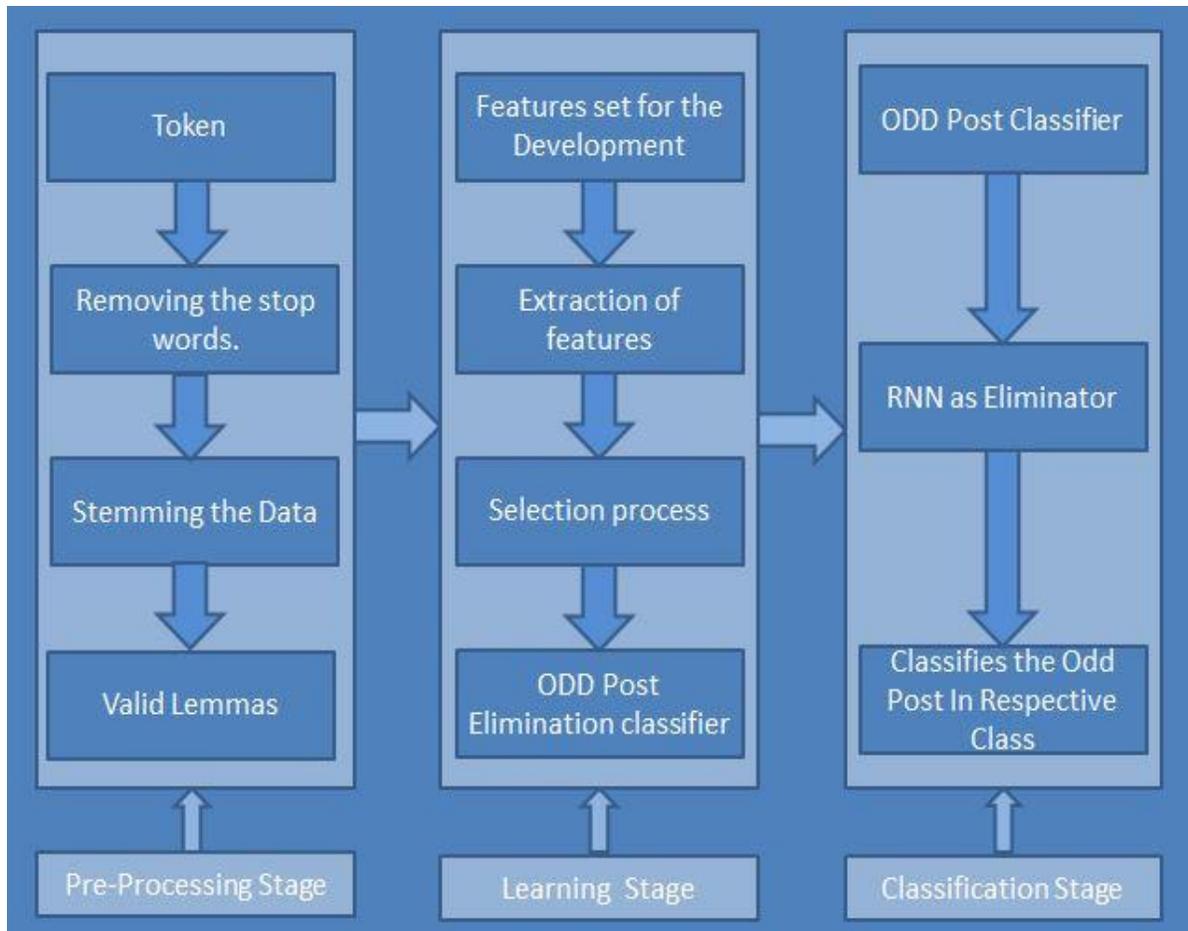
To eliminate the odd posts the data set of the tweets is introduced, in the preprocessing phase where the tokens are developed, removing the stop words is carried out, stemming of the data set is also done, and lemmatization for the valid lemmas is carried out. After the pre-processing of the tweets-dataset is lead to the feature extraction process, in this phase the features were extracted from the data, after extraction of features the selection of the process is carried out to eliminate the odd post using the classifier. The last stage of the processing is to classify the data and to eliminate the odd posts from the data. To classify the odd post classifier is used, the odd post classifier model is developed using RNN, and this technique is used to eliminate the odd posts to obtain high accuracy.
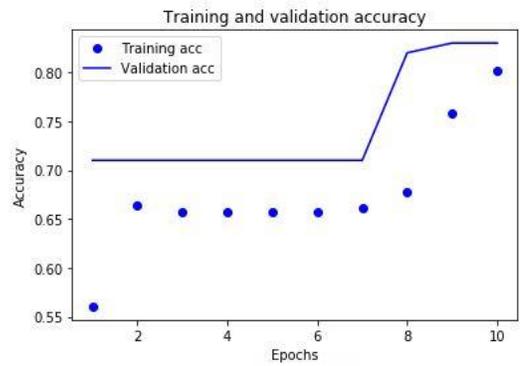
8. **Results:**

The result after applying the RNN, to the tweets dataset, the accuracy of the 78.68 percent is achieved. The results showed that high accuracy is obtained using the RNN (Evans, Jumper, Kirkpatrick, Sifre, Green, Qin., & Petersen. (2018)). The tweets were analyzed and results using RNN to identify the odd posts are observed and considered as a fast, and low-cost method to achieve high accuracy, this methodology may

use in upcoming works regarding decision making, and classifying the data (Hu, Zhao, Guo, Cheng, & Su (2019)).

This model may further implemented on the large scale data set as well.
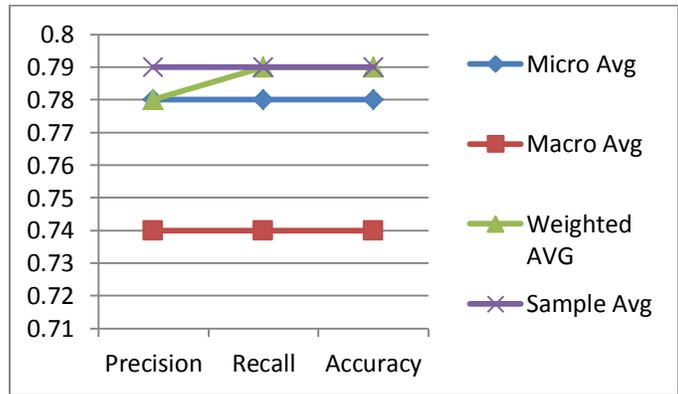
The model is trained to obtain accuracy using the dataset, the trained model improved the results, the training and validation accuracy is achieved. The trained model is applied to obtain the results using the RNN, to achieve high accuracy.

The RNN model is applied to the dataset to obtain the results and to improve the accuracy in identifying the odd posts. Moreover, in the figure here is the example of results that explains after applying the model to the data set the red sentences are odd posts and the green sentences present the normal posts.

ODD POSTS	Normal Posts
@skyteam get your systems integrated. this experience is crap! upgrade experience sucks @delta @airfrance #diamondmembermisery	@azizu787 @airfrance that must be a truly awesome experience aziz, we're sure you've enjoyed the ride!
#airfrance = #airchance i am not taking any chances. this was my last flight with #airfrance	@delta i have a travel voucher from air france - can i redeem it on bookings from https://t.co/klgoqxh934 ?
@airfrance i will never fly @airfrance ... it was my first time and last time to book their airline.	@arianespace @esa @airfrance wow, that's awesome!
@airfrance is there a toll free number to call air france corporate service. i am in canada and the agent was not helpful in Canada	@airfrance it is my account
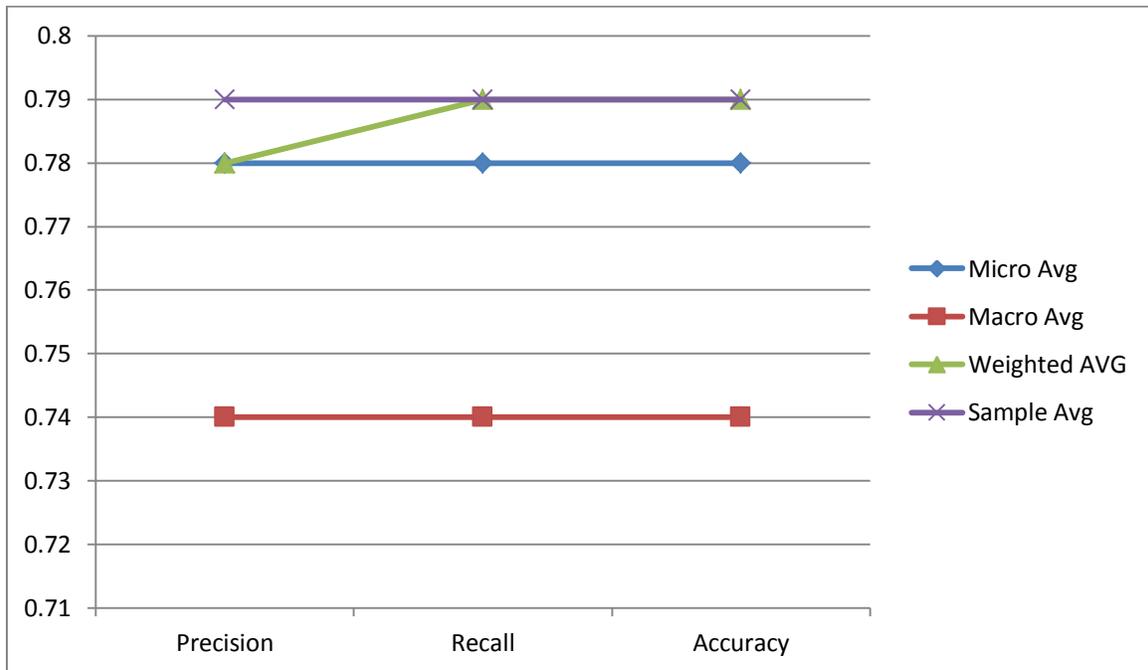
9. Finding:

After analyzing different techniques to eliminate the odd posts, different results were observed from the different techniques. To eliminate the odd posts from the small data set the tf-idf works properly (Zhang, L., Wang, S., & Liu, B. (2018)). Similarly, the support vector machine is used for the large data set and the RNN is used for the big data. Using the sentiment analysis and extracting the features from the machine

learning approach the mentioned approaches are the far better solutions that obtain high results to eliminate the odd posts (Hu, Zhao, Guo, Cheng, & Su. (2019)).
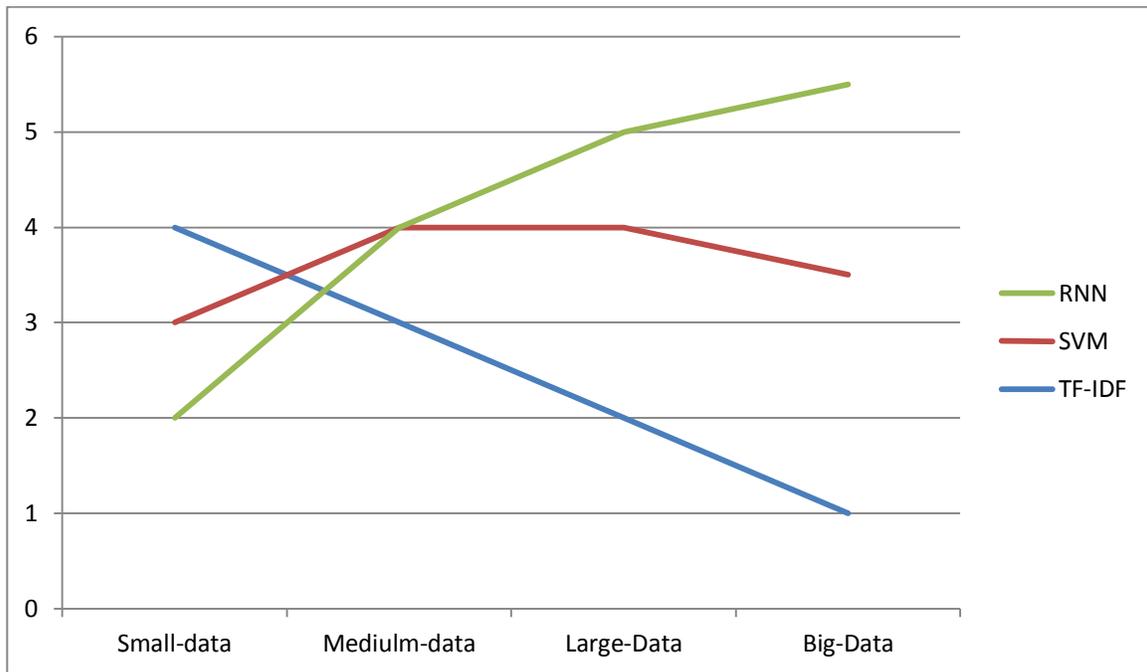
RNN obtained high-end results after applying on the tweets to classify the data by eliminating the tweets as odd posts. This may lead to the other classification issues, and may use for the further problems regarding classification, and optimization of the system.

10. Conclusion:

In this work using machine learning techniques for the sentiment analysis was observed. Sentiment analysis using unsupervised supervised and semi-supervised techniques are analyzed. Sentiment analysis has different types but using it with machine learning algorithms improves the overall results of the problem (Ma, Peng, & Cambria. (2018, April)). The overall methods to eliminate the odd posts have analyzed the importance of the odd posts in the current era with an exponential increase of the data is explained (Banerjee, Ling, Chen, Hasan, Langlotz,, Moradzadeh, & Farri. (2019)). The social media applications are the major source of this huge data, and to remain such platforms secure to overcome odd posts issue, the requirement was to construct an approach to eliminate the odd posts. So to eliminate the posts machine learning approaches including the TF-IDF, SVM, and the RNN is used to obtain high results (Banerjee,

Ling, Chen, Hasan, Langlotz,, Moradzadeh, & Farri. (2019)). RNN produces high accuracy results, on the tweets data set in eliminating the odd posts from the tweets.

References:

1. Rădulescu, C., Dinsoreanu, M., & Potolea, R. (2014, September). Identification of spam comments using natural language processing techniques. In 2014 IEEE 10th International Conference on Intelligent Computer Communication and Processing (ICCP) (pp. 29-35). IEEE.
2. Ruihai Dong, Markus Schaal and Barry Smyth, "Topic extraction from online reviews for classification and recommandation". In Proccedings of the Twenty-Third International Joint Conference on Artificial Intelligence, 2013, pp. 1310-1317.
3. Bhatt, N. P., & Saini, J. R. (2016). Identification and Analysis of Spam Words from Facebook Spam Comments. International Journal of Computer Applications, 975, 8887.
4. Green, T., & Spezzano, F. (2017, May). Spam users identification in wikipedia via editing behavior. In Eleventh International AAAI Conference on Web and Social Media.
5. Mahatab, U., & Jabeen, F. (2018, November). Spam Tags Detection and Protection Using Tags' Relationship Based Anti-Spam Approach. In 2018 IEEE 21st International Multi-Topic Conference (INMIC) (pp. 1-7). IEEE.
6. Muhammad, I., Qamar, U., & Noureen, R. (2018, November). A Systematic Review of Time Series Based Spam Identification Techniques. In Proceedings of the Future Technologies Conference (pp. 435-443). Springer, Cham.
7. Mateen, M., Iqbal, M. A., Aleem, M., & Islam, M. A. (2017, January). A hybrid approach for spam detection for Twitter. In 2017 14th International Bhurban Conference on Applied Sciences and Technology (IBCAST) (pp. 466-471). IEEE.
8. Jain, V., Kapoor, R., Gulyani, S., & Dubey, A. K. (2019). Categorization of spam images and identification of controversial images on mobile phones using machine learning and predictive learning. Journal of Discrete Mathematical Sciences and Cryptography, 22(2), 293-307.
9. Matthias, S. E. U. L., & Vogeley, V. (2019). U.S. Patent Application No. 10/212,175.
10. Luna, S., & Pennock, M. J. (2018). Social media applications and emergency management: A literature review and research agenda. *International journal of disaster risk reduction*, *28*, 565-577.
11. Kasemsap, K. (2019). Professional and business applications of social media platforms. In *Social Entrepreneurship: Concepts, Methodologies, Tools, and Applications* (pp. 824-847). IGI Global.
12. Zhang, T., & Ge, S. S. (2019, March). An Improved TF-IDF Algorithm Based on Class Discriminative Strength for Text Categorization on Desensitized Data. In *Proceedings of the 2019 3rd International Conference on Innovation in Artificial Intelligence* (pp. 39-44). ACM.
13. Arroyo-Fernández, I., Méndez-Cruz, C. F., Sierra, G., Torres-Moreno, J. M., & Sidorov, G. (2019). Unsupervised sentence representations as word information series: Revisiting TF–IDF. *Computer Speech & Language*, *56*, 107-129.(Arroyo-Fernández, Méndez-Cruz, Sierra, Torres-Moreno, & Sidorov, (2019)107-129)
14. González, I., & Mateos, A. (2019, September). Multilayer Identification: Combining N-Grams, TF-IDF and Monge-Elkan in Massive Real Time Processing. In *International Conference on Modeling Decisions for Artificial Intelligence* (pp. 213-223). Springer, Cham.
15. Yousef, M. (2019). K–Means Based One-Class SVM Classifier. In *Database and Expert Systems Applications: DEXA 2019 International Workshops BIOKDD, IWCFS, MLKgraphs and TIR, Linz, Austria, August 26–29, 2019, Proceedings* (p. 45). Springer.
16. Badal, L., & Franzén, S. (2019). A Comparative Analysis of RNN and SVM: Electricity Price Forecasting in Energy Management Systems.
17. Chathavalappil, N., & Satyanarayana, S. V. (2019). Schemes to avoid entanglement sudden death of decohering two qubit system. *The European Physical Journal D*, *73*(2), 36.
18. Mariño, J., Alborodo, R. N., Fredlund, L. Å., & Herranz, Á. (2019). Synthesis of verifiable concurrent Java components from formal models. *Software & Systems Modeling*, *18*(1), 71-105.
19. Khosravi, A., Koury, R. N. N., Machado, L., & Pabon, J. J. G. (2018). Prediction of hourly solar radiation in Abu Musa Island using machine learning algorithms. *Journal of Cleaner Production*, *176*, 63-75.

20. Petrucci, G., & Dragoni, M. (2018, June). The IRMUDOSA System at ESWC-2018 Challenge on Semantic Sentiment Analysis. In *Semantic Web Evaluation Challenge* (pp. 167-185). Springer, Cham.

21. Recupero, D. R., Dragoni, M., Buscaldi, D., Alam, M., Cambria, E., Kessler, F. B., & STLab, I. C. (2018). Proceedings of EMSASW2018-4th Workshop on Sentic Computing, Sentiment Analysis, Opinion Mining, and Emotion Detection.

22. Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), e1253.

23. Tang, D., Qin, B., & Liu, T. (2015, September). Document modeling with gated recurrent neural network for sentiment classification. In Proceedings of the 2015 conference on empirical methods in natural language processing (pp. 1422-1432).

24. Wang, L., Zhang, W., He, X., & Zha, H. (2018, July). Supervised reinforcement learning with recurrent neural network for dynamic treatment recommendation. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2447-2456). ACM.

25. Kusupati, A., Singh, M., Bhatia, K., Kumar, A., Jain, P., & Varma, M. (2018). Fastgrnn: A fast, accurate, stable and tiny kilobyte sized gated recurrent neural network. In Advances in Neural Information Processing Systems (pp. 9017-9028).

26. Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., ... & Petersen, S. (2018). De novo structure prediction with deeplearning based scoring. Annu Rev Biochem, 77(363-382), 6.

27. Hu, M., Zhao, S., Guo, H., Cheng, R., & Su, Z. (2019). Learning to Detect Opinion Snippet for Aspect-Based Sentiment Analysis. *arXiv preprint arXiv:1909.11297*.

28. Ma, Y., Peng, H., & Cambria, E. (2018, April). Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive LSTM. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

29. Banerjee, I., Ling, Y., Chen, M. C., Hasan, S. A., Langlotz, C. P., Moradzadeh, N., ... & Farri, O. (2019). Comparative effectiveness of convolutional neural network (CNN) and recurrent neural network (RNN) architectures for radiology text report classification. Artificial intelligence in medicine, 97, 79-88.