xIlkogretim Online - Elementary Education Online,2021; Vol 20 (Issue 5): pp.2702-2711
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2021.05.293

Mining Suggestions from imbalanced datasets of online reviews using SMOTE- Random Multimodel Deep Learning

Pooja kumari, Student of Mtech(computer science), patel college of science and technology Bhopal (MP), RGPV university, erpoojakumari.1992@gmail.com

Abstract—Suggestion mining is a relatively new area & is challenged by issues like the complexity in a task or manual formulation, the knowledge of sentence-level semantics, figurative sentences, handling long & complex words, context dependence, & also very imbalanced class distribution. Deep learning is an industry that can be highly competitive in machine learning. We use the Random Multimodel Deep Learning (RMDL) approach in this paper to address the problem of suggestion mining using the SemEval-2019 Task 9 data sets. Though its data sets are very imbalanced and unstructured, we have utilized SMOTE techniques to extract class imbalance problems. To solve the imbalanced dataset problem, SMOTE (synthetic Minority oversampling technique) is a widely used over-sampling tool. Experimental findings show that the advantages of SMOTE to manage complex data and imbalanced data set are superior to our current SMOTE-RMDL (SMO-RMDL) model of the existing research process.

Keywords—Suggestion Mining, Deep learning, CNN, RNN, DNN, RMDL, SMOTE.

I. INTRODUCTION

The process of identification and extraction of sentences of unstructured text comprising suggestion [1] can be described by suggestion mining (SM). In different social media channels, discussion boards, review websites, and blogs, suggestions in form of unstructured texts could be found. They are also conveyed in explicit and implied ways in the form of suggestions, tips, recommendations, notifications, acts as well as other types.

In an industrial environment, it can be helpful to identify or retrieve feedback from the text for product development, summarize consumers' views, provide recommendations, and help in the decision-making process[2]. It could enable normal users of online platforms to seek advice on general issues such as travel, fitness, food, shopping, education, and much more. Due to the vast amount of textual knowledge on several issues on the Internet, suggestion mining is a worthwhile activity for industry & academic researchers.

Almost all of the previous attempts to understand online responses have been restricted to the creation of tools for fields such as sentiment analysis (SA) or opinion mining(OM)[3]. The use of mines or suggestions for understanding can open up new areas for studying consumer behavior or gathering nuggets of knowledge that can be directly related to product creation or improvement [4], Improve customer service, and help to understand language complexities of consultancy [5]. SM is a relatively new area & is challenged by issues such as complexity in task & manual formulation, awareness of sentence-level semantics, figurative sentences, handling long & complex words, context dependence, & imbalanced class distribution. Similar problems can also be seen in dataset organizers share for SemEval work, as a real-world application with suggestions embedded in the unstructured textual content is accessible.

Deep learning (DL) is an industry in machine learning that can perform extremely well. DL approaches are better at processing Big Data relative to conventional machine learning techniques. In comparison, the in-depth methods of learning will automate learn a representation of features from raw data, then performance results. The deep structure, which includes many secret layers, is a remarkable aspect of deep learning. Traditional learning models, including SVM (Support Vector Machine) & KNN (K-Nearest

Neighbour), contain zero or only one hidden layer, on the other. These conventional model learning machines are also called shallow models.

SMOTE has been used to eliminate the imbalanced data set problem with an over-sampling technique. SMOTE is the technique that over-samples minority class data.SMOTE is a common over-sampling approach, which is intended to enhance random over-sampling but has not been thoroughly studying its conduct on high-dimensional data. SMOTE is a technique for over-sampling which generates synthetic minority samples. It's being used to get a synthetic class-balanced or almost class-balanced training set, which is then used for classification training [6].

SUGGESTION MINING

SM may be described as extracting suggestions from unstructured text, which denotes expressions of tips, recommendations, etc. Customers are usually conveyed in the form of online views on corporate objectssuch as brands, services, and products, blogs & forums, or social media sites. These views are usually positive & negative about a specified company but may include suggestions for the improvisation of the institution or tips for other customers. Traditional OM systems primarily concentrate on measuring the distribution of sentiment to an individual of interest automatically through SA methods. AnSM can broaden the capabilities of conventional OM systems, which can then be used for multiple effects. Such systems will allow the public and private sectors to draw their suggestions from different online platforms so that organizations can collect suggestions from much broader & more diverse sources than from conventional input box or online feedback formats. [7].

II. DEEP LEARNING

DL[8,9] is the idea of the human brain that has many kinds of depiction, which have simplistic characteristics at the lower levels and abstractions at high levels. Humans hierarchically order their thoughts and principles. People will learn basic concepts first and write them to represent more complex concepts. The human mind is like a DNN, composed of several neuron layers serving as role detectors, which sense more abstract characteristics as their levels increase. This is simpler for computers to generalize knowledge in a more complex manner. The key value of DL is its condensed representation of a broader variety of functions than low networks used by the more common form of learning.

A. DEEP LEARNING TECHNIQUES

1. Convolutional Neural Networks (CNN)

CNN's are very much the same as common NNs, they comprise of neurons with learning weights and preconditions. Every neuron receives such inputs, generates a dot product & follows it with non-linearity optionally. A single differential score function is now represented in the entire network, from raw pixels on one end to class scores on the other. The later (completely connected) stage still has a missing purpose and we still apply all the tips and tricks we have learned to learn standard NNs. CNN uses the fact that the feedback consists of pictures to more sensitively limit the architecture.

2. Recurrent Neural Networks (RNN)

RNN is the basic algorithm for more popular and robust sequential data. Even Siri Apple uses profoundly slow RNN to process speech. RNN has a clear memory that recalls the memory input. This segment consists of the RNN principles and Process. RNN works, which are more common because of the super memory it remembers and forecasts future events. In time-series data, speech data, and other uses, RNN is commonly used. Their profound interpretability makes them more valuable. The names are made up of FFNN and RNN channel information.

3. DEEP NEURAL NETWORK (DNN)

DL is the subsurface of artificial intelligence machine learning which focuses on biologically-inspired algorithms and the brain's functioning to support intelligent machines. A simplified DNN is depicted as a (layered) system of neurons, with connection between neurons and other neurons, which mimic neurons in the brain. These neurons transmit a communication or signal to other neurons, based on the input they receive or form a complicated network with some feedback mechanism. The following diagram shows a DNN layered 'N'.

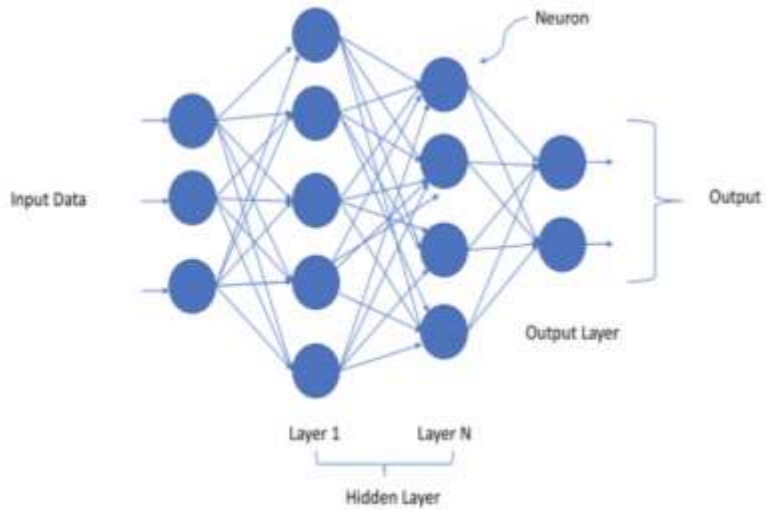

Fig 1: A DNN with N hidden layers

III. LITERATURE REVIEW

S. -H. Wu and Y. -K. Chen [2020] They have obtained data sets from Amazon.com which have been available for evaluating the utility of their reviews in 5 categories of items (Apple, Video Game, Clothing, Shoes & Jewelry, Sport & Outdoor or Prime Video) & have for 6 weeks follow the results. The tests are carried out on the data set so that zero & nonzero vote reviews are graded accordingly. We create a classification system to classify online reviews using the BERT model of DL. The results suggest that classification will produce a good result on a prediction of helpfulness. We test cross-domain prediction classifier & obtain promising results [10].

Xu, X. (2020) This study finds that consumer emotion or related hotel feedback varies according to consumers' expectations of online managerial responses. In their feedback/complaints to providers, among consumers who expect MRs, extremely unhappy consumers show extraordinarily negative emotions. Only those that praise a service provider are very positive or emotional in feedback for customers with high ratings. The two user groups – consumer guidance or criticism – have different review behaviors. We describe attributes in a positive, negative, or irrelevant relationship between comprehensive explanations or customer emotion into exciting, annoyance, or dual attributes. Consumers expecting MRs would turn more double-sided attributes into frustration. When consumers expect MM, the user concentrates on them in reviews or with clear positive or negative perceptions of those attributes. This pattern decreases the diversity of languages in comments or categorizes highly positive or negative emotional issues [11].

F. Liu et al. [2019]In this text, we propose a new method for suggestion mining dependent on DL. Cross-domain concerns & problems created by unstructured & highly imbalanced data structures are the key challenges of suggestion mining. We propose to apply RMDL to address these difficulties, merging three distinct DL architectures (DNNs, RNNs & CNNs) and choosing an ideal hyper-parameter to boost model robustness and versatility. Our experimental findings on Task 9 SemEval-2019 indicate that the majority of the current suggestion mining methods of our RMDL are better than they are now. [12].

X. Jin et al. [2019] This paper identifies deep multi-model RIS (DM-RIS) forexternal defects in which quick & robust spatially restricted Gaussian mixing models are presented for the segmentation proposal. Second, spatial information among pixels is integrated into a better Gaussian random field (MRF) model for precise & fast edge segmentation of defects. In particular, direct parameter learning in EM algorithm is futureNumerous samples of slight lighting, inequality of reflection, external noise, rust or meadow dirt are fed into Faster RCNN for the elimination of non-defective so that DM-RIS is naturally resilient for different lighting, angle, context &gaining systems. Lastly, there is a real defect in the joint hit area. The results of studies have shown that the approach proposed works well with an accuracy of 96.74 percent, 94.13 percent, 95.18 percent overlap, or 0.485 s/frame efficiencies on an average. [13].

J. Tao & L. Zhou [2019]The proposed method of forecasting the closure of a company involves many objects including the integration of DL and study of time series techniques, extraction by using a hybrid classification method of knowledge embedded in online evaluation processes, or implementation of a triple-word model toa text representation. The findings achieved from Yelp's online reviews indicate higher success in business closure forecasts. Online reviews, therefore, offer clear signals for the forecast of a business closure. The test results are consistent and show the generalizability of the proposed approach. An additional experiment is carried out using the data from TripAdvisors.com obtained. The results of this research have significant management effects for decision-making on business and investment. [14].

T. Dholpuria [2018]This essay provides anevaluation of dissimilar learning algorithms with an overview of sentiments from film reviews. The area of the evaluation sentiment is closely linked to the processing of natural language or text mining. The SA, called OM, is the sphere of a look at human assessments as thoughts to understand if the character is 'glad,' 'unhappy,' 'angry,' etc. The main purpose of the paper is to explain DL model analysis using CNN for the supervised ML classification system (Naïve Bayes, SVM, Logistic Regression, KNN & Ensemble Methods). The CNN classification system change is viewed as a comparative performance study [15].

Zhang, J., et al. [2018] The goal is to help managers prioritize new functions from the perspective of raising user ratings by mining online feedback in the next release. First, in each analysis with the LDA, we extract software features from user feedback & decide how likely they are. The simple truth rating for each function is then calculated by linear regression, going to assume that the rating of the software feature is a convex combination of all features weighted by their probabilities of distribution over the study. Eventually, we formalize the priority feature refinement as an optimization problem, which maximizes the software functionality rating of user groups under the budgetary restrictions. The solution proposed will use the topic model to semi-monitor features from user feedback together & assess the weight of each feature in the software functionality rating of each user. The evaluation of the ground truth of all characteristics reveals how reviewers judge the characteristics. Finally, we offer an illustration of the core concept of our model [16].

IV. RESEARCH METHODOLOGY

A. PROBLEM FORMULATION

The problems of SM presented in SemEval 2019 are described as a problem in binary classification. In the form of a labeled sentence dataset D, the task is to learn a grading/prediction feature which can predict l label for a sentence s, where l ∈ {Suggestion, non-Suggestion

B. PROPOSED METHODOLOGY

The existing research work depends on the neural network called Random Multimodel Deep Learning (RMDL) which combines three different DL architectures (DNNs, RNNs, and CNNs). For this research work, we will develop a new model of Neural network where the oversampling technique maybe used to solve the problem of an imbalanced dataset.

C. PROPOSED TECHNIQUES

This paper proposes using a DL, the RMDL method, or the SMOTE combining three basic architectures of deep learning for SM. RMDL: combines 3 basic DL architectures DNNs, RNNs & CNNs, which have an over-sampling approach hybridization with SMOTE, which generates synthetic minority samples. It can be better than simple over-sampling and is commonly used.

RMDL

To categorize and identify complex data, including images, documents, or videos, RMDL [17] is first studied. Three separate networks – DNNs, RNNs, and CNNs – are combined. The three DL architectures have different input layer characteristics. The method searches for the best structure to solve the problem automatically over randomly generated hyperparameters. Therefore the model learned becomes better and more detailed than conventional methods of deep learning. Since RMDL manages complex data, it is particularly suitable for mining suggestion.

RMDL typically includes a random model con-sisting for d DNN classifiers, c CNN classifiers, or r RNN classifier where d + c + r = n for binary classification problems likeSM. Randomly or independently, several layers or nodes for all n models are generated. the outcome is supported by a majority. The datapoint is labeled with Di = (xi, yi) where the input text is xi or yi the input text{0,1} is the corresponding label. For example, in mining tasks suggestion, 1 & 0 indicate whether or not the text piece includes suggestions. forecast of Di obtained by jth models is yˆi,j as described.

$$\hat{y}_{i,j} = arg\,max_k\,[softmax(y^*_{i,j})]. \qquad (1)$$

The final RMDL estimate is then determined by the equation below.

$$M(\hat{y}_{i,1}, \dots, \hat{y}_{i,1}) = \left\lfloor \frac{1}{2} + \frac{\sum_{j=1}^{n} \hat{y}_{i,j} - \frac{1}{2}}{n} \right\rfloor \quad (2)$$

One can easily verify that M($\hat{y}_{i,1}$, ..., $\hat{y}_{i,1}$) returns a majority
vote of { $\hat{y}_{i,j}$}$_j$.

SMOTE

SMOTE [19]is a technique of oversampling that produces synthetic minority class data examples. Due to its capacity to solve certain problem aspects associated with random re-sampling, it was demonstrated to be better as a technique of over-sampling than random over-sampling. Chawla introduced the technique (2003). Over 100 new versions have been released since its release. Technical extensions and variants were developed to improve algorithmic efficiency under various circumstances.

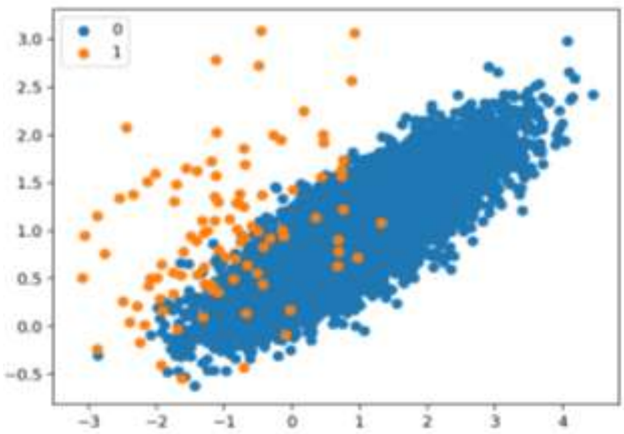

Figure 2: SMOTE for Imbalance Classification

In an over-sampling strategy, we propose to over-sample the minority class by producing synthetic" examples rather than an over-sampling substitute. This method is based on a technology which was effective in the recognition of handwritten character (Ha &Bunke, 1997). By conducting such operations on real data, they provided extra training data. Operations such as turning and skewing have become normal forms of disrupting training data. By using "function space" rather than "data space," we create synthetic instances in a less application-specific method. By taking every minority class sample & adding synthetic examples through line segments, the minority class is over-sampled by entering each neighbor of the K minority class. Neighbors from the closest neighborhoods are selected randomly based on the number of over-samples needed. Our implementation now has five nearest neighbors. [18].

V. RESULTS ILLUSTRATIONS

A. Data Description

SemEval-2019 Task9 contains the data sets that we use to execute our experiments. These two sets of data derived from different fields. Subtask A receives input from user voice platform app developers. Subtask B generates Trip Advisor website hotel reviews. Each point of data is labeled with 1 or 0, which is a suggestion or non-suggestion. In particular, one sentence belongs to one of these four categories in our data set [2].

B. Base results

Figure 3: After stop word removal & Training result of DNN

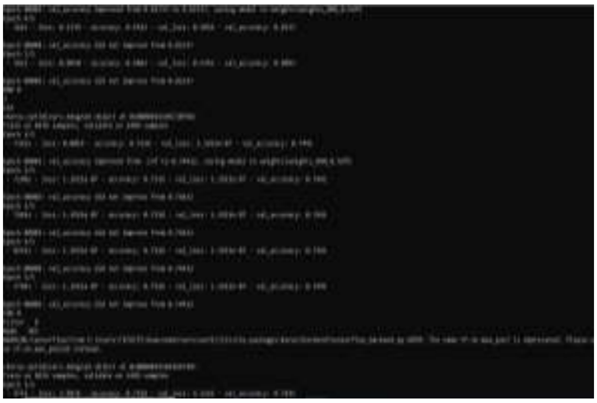

Figure 4: Training result of RNN

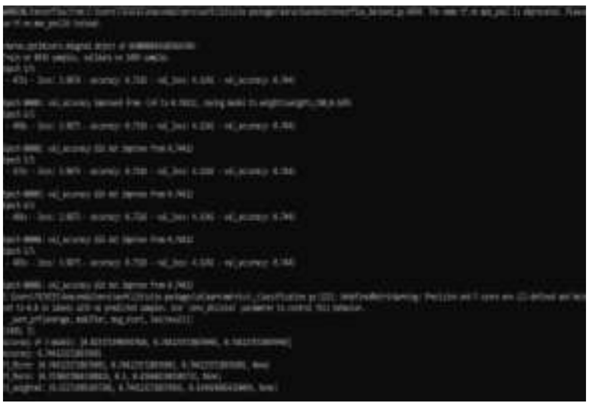

Figure 5: Training result of CNN

C. Propose Results (using Smote for data imbalance + RMDL)

Figure 6: After stop word removal usingSMOTE+RMDL

Figure 7: Training result of DNN with SMOTE+RMDL

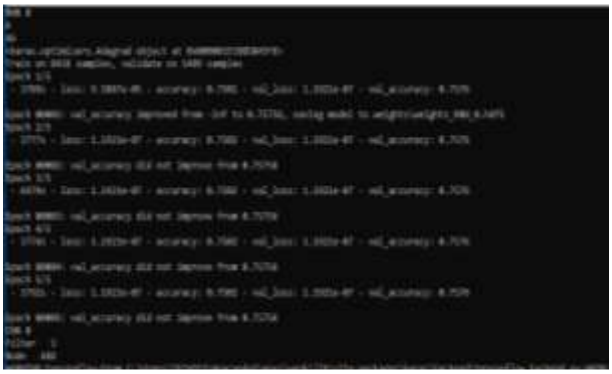

Figure 8: Training result of RNN with SMOTE+RMDL

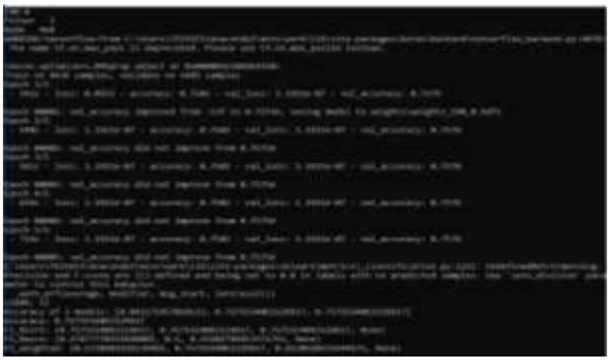

Figure 9: Training result of CNN with SMOTE+RMDL

Table 1: Loss & accuracy comparison for exiting & proposed (Smote +RMDL) with DNN, RNN & CNN

Epoch	DNN				RNN				CNN			
	Exiting		Propose		Exiting		Propose		Exiting		Propose	
	Loss	Accuracy	Loss	Accuracy	Loss	Accuracy	Loss	Accuracy	Loss	Accuracy	Loss	Accuracy
1	0.4833	0.7784	0.4221	0.8440	0. 0019	0.7526	9.5887e-05	0.7501	3.9876	0.7526	0.0555	0.7502
2	0.3154	0.8727	0.2680	0.9291	1.1921e-07	0.7526	1.1921e-07	0.7502	3.9875	0.7526	1.1921e-07	0.7502
3	0.2110	0.9221	0.2093	0.9580	1.1921e-07	0.7526	1.1921e-07	0.7502	3.9875	0.7526	1.1921e-07	0.7502
4	0.1339	0.9542	0.1834	0.9734	1.1921e-07	0.7526	1.1921e-07	0.7502	3.9875	0.7526	1.1921e-07	0.7502
5	0.0950	0.9681	0.1605	0.9785	1.1921e-07	0.7526	1.1921e-07	0.7502	3.9875	0.7526	1.1921e-07	0.7502

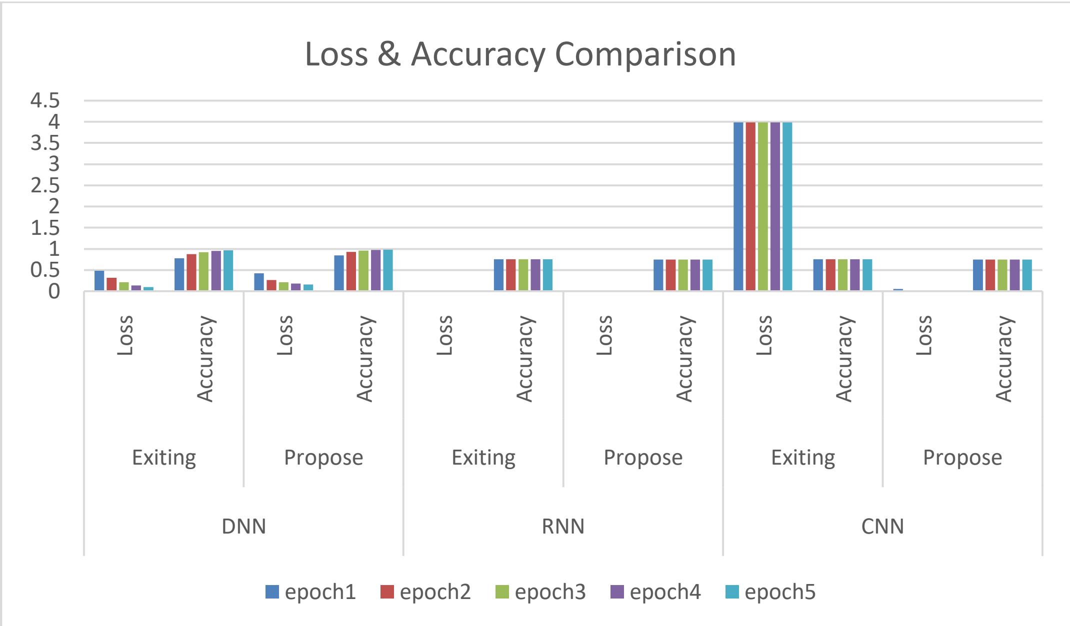

Figure 10: Loss & Accuracy Comparison graph of Exiting& Propose

Table 2: Val_loss&Val_accuracycomparison for exiting & proposed (SMOTE +RMDL) with DNN, RNN & CNN

Epoch	DNN				RNN				CNN			
	Exiting		Propose		Exiting		Propose		Exiting		Propose	
	Val loss	Valaccuracy	Valloss	Valaccuracy	Valloss	Val accuracy	Val loss	Valaccuracy	Valloss	Valaccuracy	Valloss	Valaccuracy
1	0.3994	0. 8173	0.4820	0.8301	1.1921e-07	0.7441	1.1921e-07	0.7576	4.1242	0.7441	1.1921e-07	0.7576

2	0.40 09	0. 8120	0.745 4	0.842 2	1.1921e -07	0.7441	1.192 1e-07	0.75 76	4.12 42	0.744 1	1.192 1e-07	0.7576
3	0.43 40	0. 8234	1.011 2	0.832 8	1.1921e -07	0.7441	1.192 1e-07	0.75 76	4.12 42	0.744 1	1.192 1e-07	0.7576
4	0.50 56	0. 8133	1.350 3	0.835 5	1.1921e -07	0.7441	1.192 1e-07	0.75 76	4.12 42	0.744 1	1.192 1e-07	0.7576
5	0.57 63	0. 8093	1.738 8	0.813 3	1.1921e -07	0.7441	1.192 1e-07	0.75 76	4.12 42	0.744 1	1.192 1e-07	0.7576

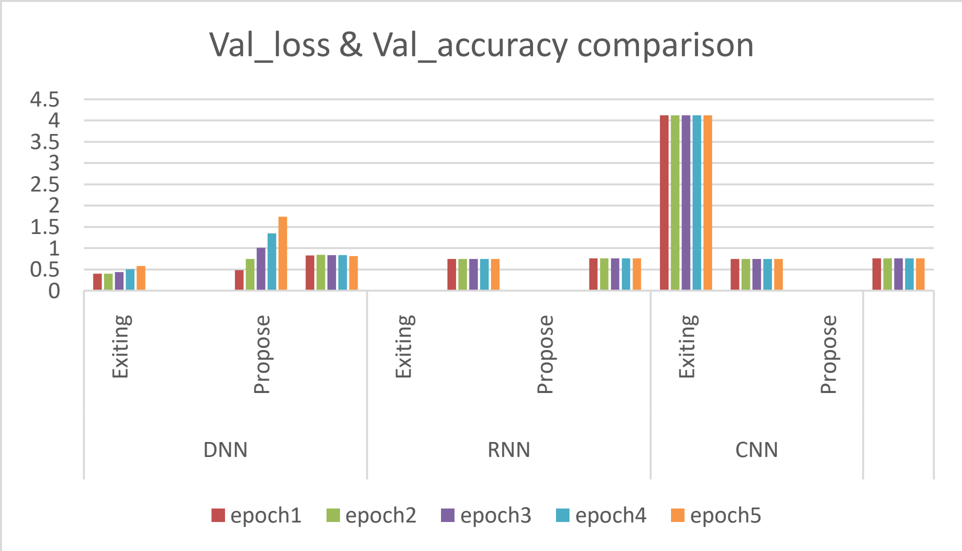

Figure 10: Val_loss&Val_accuracycomparison graphof Exiting& Propose

VI. CONCLUSION

In this paper we present our methodology & system description for the 2019 Task 9 sub-tasks A & B: online analysis and forum suggestion mining. We have applied Random Multimodel Deep Learning (RMDL) hybridized with SMOTE (SMO-RMDL) oversampling technique method to solve the suggestion mining problem where a dataset is highly imbalanced. Through the above results, we can conclude that SMO-RMDL has outperformed the previous technique where RMDL was independently applied on the same dataset but due to the class imbalance problem the results were not efficient and accurate.

In the future, we will try to explore the RMDL by making new changed and updations in this technique for more accurate results in different domains.

REFERENCES

[1] Sapna Negi, Maarten de Rijke, and Paul Buitelaar. 2018. Open domain suggestion mining: Problem definition and datasets. arXiv preprint arXiv:1806.02179.

[2] Valentin Jijkoun, Wouter Weerkamp, Maarten de Rijke, Paul Ackermans, and Gijs Geleijnse. 2010. Mining user experiences from online forums: an exploration. In Proceedings of the NAACL HLT 2010 Workshop on Computational Linguistics in a World of Social Media, pages 17–18.

[3] Walaa Medhat, Ahmed Hassan, and Hoda Korashy. 2014. Sentiment analysis algorithms and applications: A survey. Ain Shams engineering journal, 5(4):1093–1113.

[4] Caroline Brun and Caroline Hagege. 2013. Suggestion mining: Detecting suggestions for improvement in users' comments. Research in Computing Science, 70(79.7179):5379–62.

[5] Alfan Farizki Wicaksono and Sung-Hyon Myaeng. 2013. Automatic extraction of advice-revealing sentences foradvice mining from online forums. In Proceedings of the seventh international conference on Knowledge capture, pages 97–104. ACM.

[6] https://arrow.tudublin.ie/cgi/viewcontent.cgi?article=1217&context=scschcomdis

[7] https://competitions.codalab.org/competitions/19955

[8] J.Pamina, J.Beschi Raja," Survey On Deep Learning Algorithms", International Journal of Emerging Technology and Innovative Engineering Volume 5, Issue 1, January 2019 (ISSN: 2394 – 6598)

[9] V. Pream Sudha," A SURVEY ON DEEP LEARNING TECHNIQUES, APPLICATIONS AND CHALLENGES", International Journal of Advance Research In Science And Engineering, ISSN-2319-8354, IJARSE, Vol. No.4, Issue 03, March 2015

[10] S. -H. Wu and Y. -K. Chen, "Cross-Domain Helpfulness Prediction of Online Consumer Reviews by Deep Learning Model," 2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI), Las Vegas, NV, USA, 2020, pp. 412-418, DOI: 10.1109/IRI49571.2020.00069.

[11] Xu, X. (2020). Examining consumer emotion and behavior in online reviews of hotels when expecting a managerial response. International Journal of Hospitality Management, 89, 102559. doi:10.1016/j.ijhm.2020.102559

[12] F. Liu, L. Wang, X. Zhu, and D. Wang, "Suggestion Mining from Online Reviews using random Multimodel Deep Learning," 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), Boca Raton, FL, USA, 2019, pp. 667-672, DOI: 10.1109/ICMLA.2019.00121.

[13] X. Jin et al., "DM-RIS: Deep Multimodel Rail Inspection System With Improved MRF-GMM and CNN," in IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 4, pp. 1051-1065, April 2020, DOI: 10.1109/TIM.2019.2909940.

[14] J. Tao and L. Zhou, "Can Online Consumer Reviews Signal Restaurant Closure: A Deep Learning-Based Time-Series Analysis," in IEEE Transactions on Engineering Management, DOI: 10.1109/TEM.2020.3016329.

[15] T. Dholpuria, Y. K. Rana and C. Agrawal, "A Sentiment analysis approach through deep learning for a movie review," 2018 8th International Conference on Communication Systems and Network Technologies (CSNT), Bhopal, India, 2018, pp. 173-181, DOI: 10.1109/CSNT.2018.8820260.

[16] Zhang, J., Wang, Y., & Xie, T. (2018). Software feature refinement prioritization based on online user review mining. Information and Software Technology. doi:10.1016/j.infsof.2018.12.002

[17] Feng Liu," Suggestion Mining from Online Reviews using Random Multimodel Deep Learning", 2019 18th IEEE International Conference on Machine Learning and Applications (ICMLA)

[18] Nitesh V. Chawla et al., "SMOTE: Synthetic Minority Over-sampling Technique", Journal of Artificial Intelligence Research 16 (2002) 321–357.