Aspect Based Sentiment Analysis on Financial Data using Transferred Learning Approach using Pre-Trained BERT and Regressor Model

Ashish Salunkhe[1], Shubham Mhaske[2]

[1]*Pimpri Chinchwad College of Engineering and Research, Pune, India*
[2]*Pimpri Chinchwad College of Engineering and Research, Pune, India*

Abstract - *In this paper, we present a transferred learning approach for aspect classification and a regression approach for sentiment prediction on financial data provided by Financial Opinion Mining and Question Answering Open Challenge held at WWW 2018 Lyon, France. The transferred learning approach leverages the use of BERT and different regression approaches are used, with Linear Support Vector Regressor giving best results. Also, a comparative study of different existing techniques is done to provide a gist of recent advancements in this work. The performance is evaluated using performance metrics - precision, recall and F1-score for aspect classification and MSE and R Squared (R^2) metrics for sentiment prediction.*

***Key Words*: data mining, text mining, transferred learning, classification, regression, neural networks, predictive sentiment analysis, financial Sentiment analysis**

1. INTRODUCTION

Sentiment Analysis and Text Classification together have always been an important research area in Natural Language Processing. Work on sentiment analysis has received attention in academia as well as industry to analyze valuable insights from customer reviews over a specific product or a service offered. Sentiments about a certain product may differ based on the entity it is correlated with. For instance, 'the phone in red color looks good, but it is priced high.' The sentiment associated with the phone color is positive and that of its price is negative. Thus, aspect-based sentiment analysis aims to identify the polarity towards an entity based on its correlated aspects. This would enable the evaluation of sentiments based on its aspects up close. The field of financial sentiment analysis is relatively less explored. Exploring this domain to analyze the sentiments based on the aspects of unstructured text documents. Thus, based on the positive or negative sentiments users can obtain insights about possible investment opportunities and financial situations of a specific company. Future estimates about the existing market, investment, and analysis of the stability and instability of the financial entities can be done through sentiment prediction and aspect classification. Aspect-based sentiment analysis on financial data is less explored since it lacks the availability of financial sentiment data set. Current approaches in aspect-based sentiment analysis include the use of deep learning models [1], transfer learning approach [2]. The transfer learning approach has exhibited promising results with improvements in the F1 score for classification and MSE for regression tasks. Thus, the use of transfer learning with the advent of BERT [3], XLNet [4] has a definitive scope for improved results.

2. RELATED WORK

The work on aspect based sentiment analysis (ABSA) started with rule-based methods and progressed to the most recent Deep Learning methods. The task of ABSA is divided into aspect extraction and aspect sentiment classification [5]. Aspect extraction can be seen as special case of general information extraction problem. Sequential methods based on Conditional Random Field (CRF) which uses features such as POS tags, tokens, syntactic dependency, lemmas, ner, etc. gives state of the art performance in information extraction [6]. Hu and Liu [7] proposed association rule based method which finds frequent nouns and noun phrases using POS tagger. Further research with same approach has been done in the following years. Wenya Wang et al. [8] used framework consisting of Recurrent Neural Network based on dependency tree of each sentence and CRF for aspect and opinion extraction. MS Mubarok et al. [9] used Na¨ıve Bayes classifier for sentiment classification which showed excellent performance. M Al-Smadi et al. [10] compared performance of RNN and SVM for sentiment classification in which SVM outperformed RNN. ABSA was the one of the task in SemEval 2014, 2015 where most of the participated teams used rule-based approach, supervised learning methods such as SVM, Naive Bayes classifier for the sub-task of aspect sentiment classification. [11]. In recent deep learning approaches for ABSA, Duyu Tang et al. [12] used target dependent LSTM model which performed well as compared to SVM. Thien Hai Nguyen et al. [13] proposed extended RNN which uses target dependent binary phrase dependency tree constructed by combining the constituent and dependency trees of a sentence outperformed RNN and AdaRNN based models. The work on financial sentiment analysis is still in it's infancy. But promising work has been done in past year. Work by Xiliu Man; Tong Luo; Jianwu Lin [14] provides an in-depth survey on financial sentiment analysis. Their work provides comprehensive study of existing approaches including data source, lexicon-based approach, traditional machine learning approach and recent deep learning approach such as word embedding, CNN, RNN, LSTM and attention mechanism. The work by Jangid, Singhal, Shah and Zimmermann [1] displays use of multi-channel CNN for sentiment analysis and a RNN Bidirectional LSTM to extract aspect from a given headline or microblog. The work by Costa and da Silva [15] presented use of Linear Support Vector Classifier and Linear Support Vector Regressor as the solution to FiQA 2018 task 1. Shijia E., Li Yang et al. use the Attention based LSTM model [16] for aspect classification and sentiment score prediction. Yang, Rosenfeld, Makutonin have employed high-level semantic representations and methods of inductive transfer learning [2] and experimented with extensions of recently developed domain adaptation methods and target task fine-tuning.

3. METHODOLOGY

In this section, we elaborate the approach we have used. Targets and aspects related to the sentence and snippets are provided. The task is to detect the target aspects and predict the sentiment score based on the target aspect for the given text instance. The approach has two parts sentiment model and aspect model.

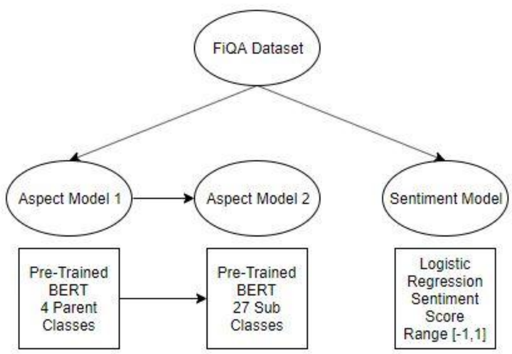

Fig -1: Methodology

3.1 Tackling Class Imbalance using SMOTE

SMOTE was first proposed by Nitesh Chawla, Kevin Bowyer, Lawrence Hall and Kegelmeyer [17]. SMOTE stands for Synthetic Minority Oversampling Technique. It is a statistical technique to overcome the issue of class imbalance by increasing the minority samples to balance the population of classes. Also, it doesn't affect the count of majority classes. We use SMOTE to address the class imbalance for level 1 aspect classification.

Table -1: SMOTE Analysis

	Corporate	Stock	Economy	Market	Total
Original dataset (equivalent to SMOTE percentage = 0)	460 (40%)	647 (56%)	7 (0.6%)	40 (3.4%)	1154
SMOTE percentage = 100	460 (35%)	647 (50%)	165 (13%)	29 (2.2%)	1301

3.2 Aspect Model

We largely use the methodology and architecture used in the BERT [3] paper and experiment with different methods of model fine-tuning, and hyper-parameter tuning. The aspect classification task is divided into two sub tasks. We divide the parent and child level aspects. On division, our first task classifies the sentences according to the parent level classes. Similarly, we classify the sentence according to the child level classes. There are 4 parent classes and 27 child / sub classes. We fine-tuned the BERT model for parent-level aspect classification and passed the same model for sub-level aspect classification.

Table -2: Distribution of aspects in training dataset

Aspect Level 1	Aspect Level 2	Count
Corporate	Reputation	10
	Company Communication	8
	Appointment	37
	Financial	26
	Regulatory	18
	Sales	92
	M&A	76
	Legal	28
	Dividend Policy	26
	Risks	57
	Rumors	33
	Strategy	49
Stock	Options	12
	IPO	8
	Signal	26
	Coverage	45
	Fundamentals	13
	Insider Activity	5
	Price Action	437
	Buyside	5
	Technical Analysis	98
Economy	Trade	2
	Central Banks	5
Market	Currency	2
	Conditions	3
	Market	24
	Volatility	11

3.3 Sentiment Model

We used the baseline machine learning models for sentiment prediction. First, the sentiment score is scaled to [0,1]. We use regression models - Linear Support Vector Regressor, Decision Tree and RNN. Word vectors are passed as input and sentiment score ranging between [0,1] is generated which is scaled to [- 1,1] as used in [1].

4. EXPERIMENTS

4.1 Dataset

The FiQA task 1 dataset [18] contains information about aspect-based sentiment analysis information about posts and news headlines extracted from finance domain web pages like Wikinews, Stocktwits and Reddit. There are 435 annotated headlines and 675 annotated financial tweets provided with aspect and sentiment score provided to every target. An example of the dataset:


```
"55": {
  "sentence": "Tesco Abandons Video-Streaming Ambitions in Blinkbox Sale",
  "info": [
   {
    "snippets": "['Video-Streaming Ambitions']",
    "target": "Blinkbox",
    "sentiment_score": "-0.195",
    "aspects": "['Corporate/Stategy']"
   },
   {
    "snippets": "['Tesco Abandons Video-Streaming Ambitions ']",
    "target": "Tesco",
    "sentiment_score": "-0.335",
    "aspects": "['Corporate/Stategy']"
   }
     ]
   }
```


To label each sentence the aspect finance tree follows node levels describe each aspect: E.g.: Stock / Price Action / Bullish / Bull Position Where: Level 1 / Level 2 / Level 3 / Level 4 Here, Level 1 represent most generic financial aspect challenges and Level 4 represents most specific financial aspect categories [18]. Aspects can have be 6 levels. For this challenge, the classification/predication up to level 2 aspect is expected.

4.2 Data Preprocessing

Data is in the plain text format so it can no be directly fed to the model. Data has some components are not helpful for analysing the nature of the data.

• The data contains punctuation marks, special characters like “ ’ !; : # & ( ) * + / ¡ ¿ = []ˆ. We removed punctuation marks and special characters by using inbuilt python string functions.

• Data also contains numbers and white spaces which are removed by using inbuilt python string functions.

• Data contains URLs which are not useful. We removed URLs from the data by using ”re” package which provides functionality for Regular Expressions.

• Data also had capitalized words which are treated differently than same words in lowercase and hence all data need to be converted to single case.

• Some common words such as to, and, am, ok which are also called as Stop-words were removed.

• There are different forms of single words exists in data and they need to be grouped so that they can be referred as single word. This is called as lemmatization.

• Label Encoding: We perform one-hot encoding for both level 1 and level 2 aspects. So before feeding it to the model, it needs to be preprocessed for optimum results. We used following approaches to preprocess our data.

4.3 Fine-tuning BERT

BERT [3] which is a pre-trained language representation model fine-tunes on other tasks. We fine-tune the pre-trained BERT model for this task.

4.4 Bert Single for Target-Aspect Based Sentiment Analysis (TABSA)

Bert for single sentence classification tasks was first introduced by Chi Sun, Luyao Huang, Xipeng Qiu [19]. Based on their work, the number of target categories are nt and na aspect categories, so the TABSA combination is nt .na

5. RESULTS

In this section we present the results for sentiment analysis and aspect classification tasks of FiQA (2018). The metrics used to evaluate sentiment model were Mean Squared Error(MSE) for sentiment model, and F1 score for aspect model. We achieved these results using pre-trained BERT [3] for aspect model and Linear Support Vector Regressor for sentiment model.

F1 Score $= \frac{2 \cdot precision \cdot recall}{precision + recall}$

Mean Squared Error $MSE = \frac{1}{q}\sum_{i=n+1}^{n+q}(Y_i - \hat{Y}_i)^2$

Table - 3: Aspect Model

	Precision	Recall	F1-Score
Microblog Posts	0.5921	0.4732	0.4610
Headlines and Statements	0.4361	0.3812	0.4068

6. CONCLUSIONS

In this paper, we present a combination of transferred learning and baseline models to do aspect-based sentiment analysis on financial tweets and headlines. We plan to train these models on a larger dataset in the future to collect information about the aspect groups that have not been adequately studied due to the lack of sufficient training samples in the current dataset. In recent times, a set of deep learning models have shown state-of-the-art performance, and we also would choose to explore and study the effects of the ensemble on our approach. Use of other transferred learning approaches like XLNet [4] can be done to improve the results of performance metrics.

Table - 3: Sentiment Model

	MSE
Microblog Posts	0.357811
Headlines and Statements	0.134721

ACKNOWLEDGEMENT

We would like to thank Department of Computer Engineering, Pimpri Chinchwad College of Engineering and Research, Ravet, Pune for their valuable assistance in this literature survey. We would also like to extend our special thanks to Prof. Dr. Archana Chaugule, Head, Department of Computer Engineering, Pimpri Chinchwad College of Engineering and Research, for her encouragement and useful critiques for this research work.

REFERENCES

[1] H. Jangid, S. Singhal, R. R. Shah, and R. Zimmermann, "Aspectbased financial sentiment analysis using deep learning," in Companion Proceedings of the The Web Conference 2018. International World Wide Web Conferences Steering Committee, 2018, pp. 1961–1966.

[2] S. Yang, J. Rosenfeld, and J. Makutonin, "Financial aspect-based sentiment analysis using deep representations," 2018.

[3] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," arXiv preprint arXiv: 1810.04805, 2018.

[4] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. Salakhutdinov, and Q. V. Le, "Xlnet: Generalized autoregressive pretraining for language understanding," arXiv preprint arXiv: 1906.08237, 2019.

[5] B. Liu, "Sentiment analysis and opinion mining," Synthesis lectures on human language technologies, vol. 5, no. 1, pp. 1–167, 2012.

[6] T. Brychc´ın, M. Konkol, and J. Steinberger, "Uwb: Machine learning approach to aspect-based sentiment analysis," in Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), 2014, pp. 817–822.

[7] M. Hu and B. Liu, "Mining and summarizing customer reviews," in Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2004, pp. 168–177.

[8] W. Wang, S. J. Pan, D. Dahlmeier, and X. Xiao, "Recursive neural conditional random fields for aspect-based sentiment analysis," arXiv preprint arXiv:1603.06679, 2016.

[9] M. S. Mubarok, Adiwijaya, and M. D. Aldhi, "Aspect-based sentiment analysis to review products using na¨ıve bayes," in AIP Conference Proceedings, vol. 1867, no. 1. AIP Publishing, 2017, p. 020060.

[10] M. Al-Smadi, O. Qawasmeh, M. Al-Ayyoub, Y. Jararweh, and B. Gupta, "Deep recurrent neural network vs. support vector machine for aspectbased sentiment analysis of arabic hotels' reviews," Journal of computational science, vol. 27, pp. 386–393, 2018.

[11] M. Pontiki, D. Galanis, H. Papageorgiou, S. Manandhar, and I. Androutsopoulos, "SemEval-2015 task 12: Aspect based sentiment analysis," in Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015). Denver, Colorado: Association for Computational Linguistics, Jun. 2015, pp. 486–495. [Online]. Available: https://www.aclweb.org/anthology/S15-2082

[12] D. Tang, B. Qin, X. Feng, and T. Liu, "Target-dependent sentiment classification with long short term memory," arXiv preprint arXiv: 1512.01100, 2015.

[13] T. H. Nguyen and K. Shirai, "PhraseRNN: Phrase recursive neural network for aspect-based sentiment analysis," in Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon, Portugal: Association for Computational

Linguistics, Sep. 2015, pp. 2509–2514. [Online]. Available: https://www.aclweb.org/anthology/D15-1298

[14] X. Man, T. Luo, and J. Lin, "Financial sentiment analysis (fsa): A survey," in 2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS), May 2019, pp. 617–622.

[15] D. de Franc¸a Costa and N. F. F. da Silva, "Inf-ufg at fiqa 2018 task 1: Predicting sentiments and aspects on financial tweets and news headlines," in Companion Proceedings of the The Web Conference 2018, ser. WWW '18. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee, 2018, pp. 1967–1971. [Online]. Available: https://doi.org/10.1145/3184558.3191828

[16] S. E., L. Yang, M. Zhang, and Y. Xiang, "Aspect-based financial sentiment analysis with deep neural networks," in Companion Proceedings of the The Web Conference 2018, ser. WWW '18. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee, 2018, pp. 1951–1954. [Online]. Available: https://doi.org/10.1145/3184558.3191825

[17] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," Journal of artificial intelligence research, vol. 16, pp. 321–357, 2002.

[18] M. Maia, S. Handschuh, A. Freitas, B. Davis, R. McDermott, M. Zarrouk, and A. Balahur, "Www'18 open challenge: Financial opinion mining and question answering," in Companion Proceedings of the The Web Conference 2018, ser. WWW '18. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee, 2018, pp. 1941–1942. [Online]. Available: https://doi.org/10.1145/3184558.3192301

[19] C. Sun, L. Huang, and X. Qiu, "Utilizing BERT for aspectbased sentiment analysis via constructing auxiliary sentence," in Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). Minneapolis, Minnesota: Association for Computational Linguistics, Jun. 2019, pp. 380–385. [Online]. Available: https://www.aclweb.org/anthology/N19- 1035