Human Behavior Assessment using Ensemble Models

Abdullah Faiz Ur Rahman Khilji [1] Rituparna Khaund [2] Utkarsh Sinha [1]

[1]Computer Science and Engineering [2]Electrical Engineering
National Institute of Technology Silchar

Abstract

Behavioral analysis is a pertinent step in today's automated age.It is important to judge a statement on a variety of parameters before reaching a valid conclusion. In today's world of technology and automation, Natural language processing tools have benefited from growing access to data in order to analyze the context and scenario.A better understanding of human behaviors would empower a range of automated tools to provide users a customized experience. For precise analysis, behavior understanding is important.We have experimented with various machine learning techniques which are described as part of the ALTA 2020 shared task. In this work,we have enlisted our results and the challenges faced to solve the problem of the human behavior assessment.

Dataset

The labeled dataset `https://www.kaggle.com/c/alta-2020-challenge/data` for the ALTA 2020 shared task was provided by the organizers. The dataset included single, multiple, or no labels for a single sentence as the output label.(Mohammad et al., 2018)The train data contains a total of 200 instances of labeled data, whereas the test set contains 100 instances.For the purpose of experimentation, we have worked with both sets of data, with and without preprocessing.Preprocessing steps include removal of punctuation and stop words.

Experimental Setup

As the dataset was quite small , we employed Traditional Machine Learning techniques data craving deep learning methods.For the word embeddings, we have experimented with the XLNet(Yang et al., 2019) pre-trained embeddings and the freely available spaCy `https://spacy.io/` word embeddings. XLNet is a pretraining method based on generalized autoregressors, that learns bidirectional context information.The autoregressive nature overcomes the deficit of the BERT (Devlin et al., 2018) model. We have used the pretrained XLNet model as provided by spaCy and used the generated vectors for the downward classification tasks. We have used the`en_core_web_lg` as provided by spaCy. The sentence vectors generated by the model is used directly for the multiclassification step.The generated sentence vectors of each sentence are fixed to a length of 300. For reasons attributed to computational cost and efficiency, we have used polynomial features of degree 2 in our experiments.Polynomial features (James et al., 2013) can be termed as a feature engineering task, wherein new inputs are generated based on the current set of inputs. A pipeline of polynomial features of degree 2 was combined with Decision Tree Classfier.Decision Tree (Swain and Hauska, 1977) is a machine learning technique based on the supervised approach. It formulates the task as a graphical structure, wherein the features are represented as the internal nodes. The rules are represented by the tree branches. Finally, the outcome of the tree is given by the leaf.This ensemble model has experimented been on both XLNet and spaCy word embeddings. The model incorporating the use of XGBoost has also been used. XGBoost is a scalable algorithm frequently obtaining state-of-the-art results in many machine learning tasks with limited dataset size. The given algorithm is a combined model of decision trees,which uses copies of itself to improve the model performance and minimizes error. It is an efficient version of the well known stochastic gradient boosting algorithm. Various other approaches are employed and the obtained score is tabulated in Table 1.

Results and Analysis

As we can see from Table 1, the highest score of 0.1033 on the private dataset is using the XGBoost approach with pretrained spaCy embeddings. The highest score of 0.2200 on the public leaderboard is using a decision tree classifier with polynomial features of degree 2.

S. No.	Approach	Private Score	Public Score
1	XGBoost with spaCy pretrained embeddings	0.1033	0.1733
2	Polynomial features with degree 2 together with decision tree classifier, using pre-trained XLNet embeddings	0.1000	0.1600
3	Using polynomial features and decision tree regressors, with spaCy pretrained embeddings.	0.0593	0.1866
4	Decision tree with spacy embeddings	0.0533	0.2066
5	Polynomial features with degree 2 together with decision tree classifier, using pre-trained spaCy embeddings	0.0533	0.2200
6	Decision tree classifier along with polynomial features of degree 2 , incorporating removal of stop-words	0.0533	0.2033

Table 1:Techniques Employed with corresponding Public and Private Mean F-Score

From Table 2, we can see that the first three predictions go with the original analysis and the last three contradicts the original interpretation, we can also see that the actual output contains more than one class, our analysis engine can replicate the same,but since the textual description was so short, the system was not able to properly analyze and map it with the output.

S. No	Prediction	Text	Actual Behaviour	Predicted Behaviour
1	Correct	Actually be arsed with my sister sometimes, she controls the TV 90% of the time and when I watch one thing she gets in a huff	Normality	Normality
2	Correct	You ever just be really irritated with someone u love it's like god damn ur makin me angry but I love u so I forgive u but I'm angry	Capacity	Capacity
3	Correct	@SaraLuvvXXX : Whaaaat?!? Oh hell no. I was jealous because you got paid to f**k, but this is a whole new level. #anger #love #conflicted& Propriety	Propriety	Propriety
4	Incorrect	it makes me so f**king irate jesus. nobody is calling ppl who like hajime abusive stop with the strawmen lmao	Propriety	Normality
5	Incorrect	Goddamn headache.	Propriety	Capacity, Tenacity
6	Incorrect	I wanna kill you and destroy you. I want you died and I want Flint back. #emo #scene #f**k #die #hatered	Capacity, Tenacity	Propriety

Table 2:Model with the best Private Score

Discussion and Conclusion

In our work, we have worked with various deep learning algorithms and fusion techniques to study and investigate human behavior. We have also set up the analogy between the human sentiment analysis and behavior.As the dataset size was not so significant, the system is not trained on complex deep learning-based architectures. We can infer that a less complex framework can sometimes perform better than complex architecture, moreover, if the dataset size would be significantly more,then a more complex architecture could have been devised and incorporated. The semantic analysis could have been carried out using those datasets.Future works can involve a rule-based approach for the same problem statement. Such an approach would be able to provide much better results even on a smaller dataset. Various techniques could be used to improve on the dataset size, and a deep learning architecture can be developed to cater to the same.

Acknowledgment

The authors are thankful to the organizers of the ALTA 2020 shared task for organizing the event and Kaggle InClass for hosting the competition. Special thanks to Dr. Diego Mollá-Aliod for his support.

References

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. *An introduction to statistical learning*, volume 112. Springer.

Saif M. Mohammad, Felipe Bravo-Marquez, Mohammad Salameh, and Svetlana Kiritchenko. 2018. Semeval-2018 Task 1: Affect in tweets. In *Proceedings of International Workshop on Semantic Evaluation (SemEval-2018)*, New Orleans, LA, USA.

Philip H Swain and Hans Hauska. 1977. The decision tree classifier: Design and potential. *IEEE Transactions on Geoscience Electronics*, 15(3):142--147.

Zhilin Yang, Zihang Dai, Yiming Yang, Jaime Carbonell, Russ R Salakhutdinov, and Quoc V Le. 2019. Xlnet: Generalized autoregressive pretraining for language understanding. In *Advances in neural information processing systems*, pages 5753--5763.