Check-in Report: SemEval-2020 Task 7 - Assessing Humor in Edited News Headlines

Anonymous ACL submission

Abstract

This paper provides a report of the progress made since the proposal of this task was submitted.

1 Literature Review

Further study has been carried out on a few underlying theories and hypotheses that provide useful insight into the nature of humour. These theories have been adopted over the years as baselines to understanding the intricacies of humour.

The Script Semantic Theory of Humor (SSTH) (Raskin, 1979) performs a formal semantic analysis to identify the semantic property of humorous text. The main hypothesis is that humorous text contains the overlap of two or more opposite semantic scripts that are also compatible.

The theory of verbal humour, as introduced by (Attardo, 2017), made similar assertions to the SSTH, providing further knowledge resources. Namely, the logical Mechanism, describing incongruity resolution in the script opposition, the setting of the joke, the target, the narrative and the language, including its morphological and syntactical makeup.

(Ahuja et al., 2018) proposes a framework based on the success of previous theories such as SSTH and GVTH for classification of humor. They have identified three major characteristics that tend to persist across all types of jokes, which are the mode of delivery, theme and central topic. From their results we can see that SVM Classifier has a better accuracy than Naïve Bayes and Logistic regression for humor classification.

2 Methods

As previously proposed, we plan to employ both statistical machine learning and deep learning approaches to solve the tasks.

From our review of previous submissions, we identified the use of both Linear Regression and Support Vector Regression algorithms for solving task 1. We have started our process of implementing these two algorithms as part of the statistical machine learning approaches that we intend to implement for the project.

Linear Regression is a linear model that assumes a linear relation between the input variables and the output variable. Support Vector Regression algorithm tries to find the best fit line which is the hyperplane that has the maximum number of points in order to determine the output value.

The vectorization method that has been utilized as part of preprocessing the data is TFIDFVectorizer, abbreviated as Term Frequency Inverse Document Frequency.is an algorithm that transforms text to vector representations which are used as an input to the estimator.

For the machine learning methods feature columns are ‘original’ and ‘edit’ of the newspaper headlines, and the output is the comparison between ‘meanGrade’ and the predicted mean grade from running the previously mentioned regression algorithms on the dataset. We shall use the prediction outputs from this stage for task 2 classification task. In terms of a deep learning approach, we have decided on using BERT given its tremendous success by the leading teams in solving the tasks. Thus, we shall be able to perform our comparison in terms of result for both the approaches to further analyze the indications from the results.

3 Evaluation

As stated in the proposal, evaluation will be carried out using two metrics. For the first task; the Root Mean Squared Error metric (RMSE) would be adopted to measure the difference between the values predicted by our proposed model and the

actual result provided. The RSME metric comes as a built-in method in the Sklearn library of Python. Furthermore, Accuracy would be adopted in evaluating the result generated for the second Classification task.

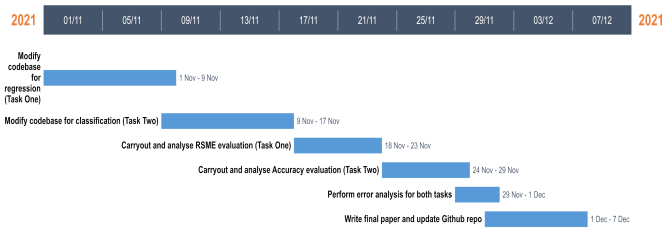
4 Result

In terms of results, we have identified two codebases in which Linear Regression and Support Vector Regression algorithms were adopted for solving task 1. The current challenge experienced in running the codes is solely machine-dependent. We would employ the use of a virtual machine to produce an output of the code.

5 Timeline

Title	T/M	Start	End		%
Modify codebase for regression (Task One)	T	01/11/2021	09/11/2021	7 days	%
Modify codebase for classification (Task Two)	T	09/11/2021	17/11/2021	7 days	%
Carryout and analyse RSME evaluation (Task One)	T	18/11/2021	23/11/2021	4 days	%
Carryout and analyse Accuracy evaluation (Task Two)	T	24/11/2021	29/11/2021	4 days	%
Perform error analysis for both tasks	T	29/11/2021	01/12/2021	3 days	%
Write final paper and update Github repo	T	01/12/2021	07/12/2021	5 days	%

Project Completion Timeline



6 Repository

https://github.com/UOFA-INTRO-NLP-F21/f2021-proj-udituen

References

Vikram Ahuja, Taradheesh Bali, and Navjyoti Singh. 2018. What makes us laugh? investigations into automatic humor classification. In *Proceedings of the Second Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media*, pages 1–9.

Salvatore Attardo. 2017. The general theory of verbal humor. In *The Routledge handbook of language and humor*, pages 126–142. Routledge.

Victor Raskin. 1979. Semantic mechanisms of humor. In *Annual Meeting of the Berkeley Linguistics Society*, volume 5, pages 325–335.