Sentiment Classification on Product Reviews

AUTHOR 1
AUTHOR 2
AUTHOR 3

May 10, 2019

Abstract

Contents

1 Introduction

This section should briefly introduce the project, **for example**

- What is the project about?
- What is the goal of the project?
- How the project was done in your group? For example, how was the work allocated to each individual group member?

2 Pre-processing and feature generation

This section discusses

- The steps you used to pre-process the product reviews
- What are the feature set(s)? and how were they generated?
- For each feature set, what are the corresponding corpus statistics which include
  - The total number of documents
  - The size of the vocabulary
  - The average number of tokens of the product reviews
  - The minimum and maximum number of tokens of the product reviews
  - The standard deviation of the size of the product reviews

  Or any other statistics that you think is necessary to be documented [1, 3, 2].

2.1 Pre-processing

2.2 Feature generation

3 Models

This section discusses the classifiers that your group used in the project, particularly those used in the comparison. The discussion should include

- A description of each of the model. For example,
  - the basic idea of each model,
  - how the model is constructed and works,
  - what are the assumptions made by each model if there is any.

- The difference between the models
  - the advantage(s) and disadvantage(s) of the models
- Or any other aspects of the models that you think should be discussed.

Please replace the title of each subsection below with the name of the model your group used in the development.

3.1 Model 1

3.2 Model 2

3.3 Model 3

3.4 Discussion of model difference(s)

4 Experiment setups

This section discusses how the experiments were setup. For example

- What are the parameter settings for the models, if there is any?
- How was the cross-validate setup, if you used cross-validation?
- If you used training-validation approach on the provided training dataset, how your setup the split?
- How was the accuracy computed?

5 Experimental results

This section should discuss the comparison of different models with different feature set(s) (under difference settings if necessary). You could report your results in table(s) and discuss the results.

6 Conclusion

This section concludes the report with, for example,

- What is the optimal classifier? What is the best set of features used by that classifier?
- Any lessons learned from this project in regard to Sentiment Classification?
- Any suggestions for the future work?

References

[1] Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., Dean, J.: Distributed representations of words and phrases and their compositionally. In: Advances in Neural Information Processing Systems 26, pp. 3111–3119 (2013)

[2] Miller, G.A.: WordNet: a lexical database for English. Communications of the ACM **38**(11), 39–41 (1995)

[3] Pennington, J., Socher, R., Manning, C.: GloVe: Global vectors for word representation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, pp. 1532–1543 (2014)