COMP30027: Assignment 2 Reviews

Rohan Hitchcock (836598) and Patrick Randell (836026)

Report 1 review

This report approaches the task of predicting the ratings of restaurant reviews from review text by employing several classifiers popular in sentiment analysis, including Gaussian naïve Bayes, support vector machine, and logistic regression classifiers.

The analysis of selected models for the task is quite extensive. The report documents a thorough investigation into several models' hyperparameters, which resulted in a significant increases in prediction accuracy for the final models.

This report would benefit by including extended discussion of the motivation behind decisions, and a discussion of the literature surrounding the topics covered. Not including this obscures the assumptions the author(s) are making, making it difficult for a reader to assess whether they are reasonable or not. It also makes the direction of the report unclear, and excluding a discussion of related literature makes the report less accessible to readers' and further hampers their ability to assess the methodology.

This report would also benefit from improved formatting and presentation. Formatting problems such as missing spaces, significant variation in the style of tables, figures of raw terminal output containing extraneous information, and including an unreasonable number of significant figures are distracting to a reader and detract from the overall readability of the report. The names of general concepts should be used, rather than implementation-specific terminology (such as method names or file names), so readers can connect their existing knowledge to the concepts being discussed. Additionally, values are often included without explaining what they are (e.g. accuracy, F-score, etc.) which means the reader needs to guess.

Report 2 review

This report presented an approach to text sentiment analysis which is based on sophisticated feature generation and selection. The report details several different methods of feature generation, and assesses the performance of a number models on the generated features, eventually adopting a homogeneous ensemble approach.

The overall approach discussed in this report is creative and interesting. In particular, the feature generation and selection techniques are well considered, and appear to be based on a robust statistical understanding of natural language processing. This is reflected in the strong performance of the final classifier. Decisions are often justified using experimental data, which provides clear motivation for final methodology, and also allows readers better assess these decisions independently.

This report would benefit from more explanation of the theory behind the techniques. While the efficacy of the techniques is clear from the experimental data presented, it is not clear why this is the case. In particular, it would be good to include a discussion of why different classifiers had different performance with the selected features, and why the generated features result in strong model performance (even if this discussion is somewhat speculative). Along these lines, this report would be improved by including a more comprehensive introduction to the feature generation techniques employed. In the reviewers' opinion, the assumption that readers are "familiar with the basic theory of TF-IDF and n-gram vectorised representation of text" is not reasonable, and failing to include a detailed discussion of

this makes the remainder of the report less accessible. Although interesting, the introductory discussion of natural language processing and polarity classification in general is not strictly relevant to the remainder of the report, and it would be better to prioritise a discussion of the theory required to appreciate the techniques employed. In addition, including more references to literature would help guide a reader looking to explore these topics further.