Summary

You begin with an introduction of the whole paper, with a general overview of the method. This then moves into a shallow analysis of the dataset, including sentiment counts and an example instance. The method defines the ways in which the data were pre-processed, vectorized, selected to train a set of models. The models and their hyperparameters are defined, followed by the evaluation metrics used. The results section contains weighted metrics for the different models (and for some hyperparameter configurations). The analysis section explores two research avenues: comparing evaluation train/test split generation methods and comparing the classifiers and their hyperparameters. Finally, the best model is listed and further areas for research listed in the conclusion.

What was done well

The introduction is exhaustive and does a good job of summarising what's to come in the text. The writing is concise and well explained throughout the report. In your analysis, I like how you relate the mechanisms by which the SVM and logistic regression operate to explain the similar accuracies/F1 scores. This is something I did not pick up on. Your conclusion includes contextually appropriate areas of improvement on the research.

Potential areas for growth

In data pre-processing, you say that you perform both stemming and lemmatization. This is redundant as both methods aim to determine the root of a word. Lemmatization relies on a predefined corpus of terms, so lemmatizing a stemmed word will not always return the correct lemma. I believe the intended audience for this report are knowledgeable machine learning conference attendees. Therefore, the explanations of how the classifiers work may not be necessary. Finally, I'd recommend making use of visuals where possible. For example, when discussing tuning C in the SVM, consider using a learning curve to present accuracy results.