Summary

Your paper walks me through your research clearly. Beginning with a shallow exploration of the dataset (sentiment distribution and set lengths), then moving onto the methodology. This includes steps taken to clean the data of unwanted features, then describing tokenization and vectorization methods. You then move onto describing the models you consider. Then, evaluations metrics (mostly accuracies and F1-score-related metrics) are listed per model. Next, a series of research questions are posed, and the results analysed to determine findings. Finally, simple conclusion summarizes the paper and proposes further research.

What was done well

Thorough and exhaustive explanation of the method used throughout the assignments. I knew which classifiers, features and hyperparameters were used for testing and the final models. The research questions you pose cover most of the different steps in model creation (feature selection, training data construction, model construction, parameter tuning). The conclusion includes a paragraph devoted to identifying further avenues for research from these results. This included mentioning features that merited extraction, or other tests that could have been performed.

Potential areas for growth

First, I'd recommend removing the names in-text citations from the brackets if the author's work is being directly referenced. For example, "According to Dietterich (1997), …". However, this depends on the citation style being used. Next, the "hypothesis" headers throughout the text appear to be research questions, not hypotheses. This can be disorienting to the reader (since no hypothesis is given) and therefore subtracts from the analysis under the header. Finally, Twitter handles do not appear in every tweet in the dataset and are used to address other users within the platform (not the original poster). It would be good to keep them in their own feature category and see if some handles mentioned correlate to a specific sentiment.