Good News: A Machine Learning Project

James Hardwick , **Nathan Miller** and **Sebastian Annett**

University of the West of England
Bristol, England
{james3.hardwick, sebastian2.annett, nathan2.miller}@live.uwe.ac.uk

Abstract

Reading the news for many, especially recently, has not been an enjoyable task. Articles of COVID, wars, and celebrity deaths have made many previous readers of the news reluctant to keep up to date with more recent, lighter world events. In this project, the possibility of classifying news into 'good' news and 'bad' news will be explored.

1 Introduction (10%)

Describe the problem you are working on, why it's important, and an overview of your results.

Reading the news for many, especially recently, has not been an enjoyable task. Articles of COVID, wars, and celebrity deaths have made many previous readers of the news reluctant to keep up to date with more recent, lighter world events. In this project, the possibility of classifying news into 'good' news and 'bad' news will be explored.

2 Related Work (10%)

Discuss published work that relates to your project. How is your approach similar or different from others?

3 Data (10%)

Describe the data you are working with for your project. What type of data is it? Where did it come from? How much data are you working with? Did you have to do any preprocessing, filtering, or other kinds of treatment to use this data in your project?

Locating data for news article sentiment analysis proved tricky, however a dataset of posts from *Twitter* ('Tweets') was located [**?**], labelled with whether the Tweet was interpreted as good or bad news. This dataset simply contained a post *id* and a *label*. Unfortunately, this was not much help on it's own, so to obtain the original Tweets, a developer account for *Twitter* was needed, in order to scrape original text from the Tweet ids. Once original texts had been obtained, sentiment analysis could begin.

4 Methods (30%)

Discuss your approach for solving the problem that you set up in the introduction. Why is your approach the right thing to do? Did you consider alternative approaches? You should demonstrate that you have applied ideas and skills built up during the module to tackling your problem of choice. It may be helpful to include figures, diagrams, or tables to describe your method or compare it with other methods.

5 Experimentation (30%)

Discuss the experiments that you performed to demonstrate that your approach solves the problem. The exact experiments will vary depending on the project, but you might compare with previously published methods, perform an ablation study to determine the impact of various components of your system, experiment with different hyperparameters or architectural choices, use visualisation techniques to gain insights into how your model works, discuss common failure modes of your model, etc. You should include graphs, tables, or other figures to illustrate your experimental results.

6 Conclusion (5%)

Summarize your key results - what have you learned? Suggest ideas for future extensions or new applications of your ideas.

7 Writing and Formatting (5%)

Is your paper clearly written and nicely formatted?