

Using Machine Learning and Deep Learning Methods to Detect Cyberbullying in Messages

Student Name: Jack Leyland

Supervisor Name: Dr Noura Al Moubayed

Submitted as part of the degree of BSc Computer Science (with Year Abroad) to the
Board of Examiners in the Department of Computer Sciences, Durham University

Abstract — Half a page. Write this at the end.

Keywords —

I INTRODUCTION - 2/3 PAGES

- What is the project about (Bullying, Cyberbullying)
- Context of project
- Research question and Objectives
- What was achieved?

II RELATED WORK - 2-4 PAGES

- Related work, state of the art.
- RELATE TO MY PROJECT.
- Table about the datasets. Vocab size, length, number of examples etc.

III SOLUTION - 4/7 PAGES

- Tools used and stages of life cycle
- Cleaning data
- Feature extraction (TF, TF-IDF, Bigrams/Trigrams)
- (+ different embeddings, GloVe, ELMo, training my own)
- Machine Learning models (Naive Bayes, Log.Reg, Gradient Boosted Classifier)
- Deep Learning models (LSTM, CNN)
- F1 score *as a loss function*

- Testing (training graphs)
- Any extra things I try (ensemble, one-class classification)

IV RESULTS - 2/3 PAGES

- Measures I decided to use and why. Differences between F1, MicroF1, WeightedF1, why accuracy isn't good.
- Training curves
- Table with best architecture (with F1 score and test accuracy) for each dataset. Model info.
- Other interesting settings that I tried with results.

V EVALUATION - 2/3 PAGES

- Strengths and Weakness of solution
- Lessons learnt
- Not just evaluating model, it's about evaluating results effectively.

VI CONCLUSIONS - 1 (OR 2) PAGES

- What I have achieved overall
- A little bit about future work

References