

R222 Assignment

Kaho Sato
King's College
ks789@cam.ac.uk

1 Introduction

There is an increasing interest in Native Language Identification (NLI). Native Language Identification involves detecting the native language (L1) of the author of the text. This is a sub-task of author profiling, where the goal is to predict traits of the author such as gender, age or psychometric traits (Estival et al., 2007). Author profiling has been used to combat the criminal activity (Abbasi and Chen, 2005) to gain an insight to the suspects by analysing the online mediums used for communication among them.

2 Related Work

3 Background

4 Dataset

5 Architecture

6 Results

7 Conclusion

References

- Ahmed Abbasi and Hsinchun Chen. Applying authorship analysis to arabic web content. In *International Conference on Intelligence and Security Informatics*, pages 183–197. Springer, 2005.
- Dominique Estival, Tanja Gaustad, Son Bao Pham, Will Radford, and Ben Hutchinson. Author profiling for english emails. In *Proceedings of the 10th Conference of the Pacific Association for Computational Linguistics (PACLING07)*, pages 263–272, 2007.