

Contents

1 Introduction

- 1.1 Motivation
- 1.2 Related Work
 - 1.2.1 NER systems
 - 1.2.2 NER techniques

2 Named Entity Data

- 2.1 BIO scheme
- 2.2 Named entities distribution

3 CRF

- 3.1 Overview
- 3.2 Algorithms for CRF
- 3.3 CRF toolkits. MaRMoT

4 Feature Description

- 4.1 Word-level features
- 4.2 Cluster features
- 4.3 List lookup features
 - 4.3.1 Wikipedia

5 Experiments

- 5.1 Evaluation metrics
- 5.2 Word-level features results
- 5.3 External features results
- 5.4 Modifying the BIO schema
- 5.5 Optimization
- 5.6 Error analysis

6 Conclusion and Future Work

7 Attachment

Bibliography

List of Figures

List of Tables