

Extraction of Goals from Premier League Match Reports using Natural Language Processing Techniques

Student Project

presented by
Jochen Hülß & Matthias Rabus
Matriculation Number 1376749 & 1207834

submitted to the
Chair of Information Systems V
Prof. Dr. Christian Bizer
University Mannheim

Mai 2013

Contents

Chapter 1

Project Summary

1.1 Application domain and goals

Jochen. Einleitung, Uebertragbarkeit auf andere Text Mining Tasks. Steps to goal

1.2 Structure and size of the data set

Matze. siehe outline, punkt 1.2 Struktur input und struktur fuer web mining (satzweise getrennt)

1.3 Preprocessing

Matze: Wie wurde struktur aus vorherigem Abschnitt erreicht? Jochen: Gate

1.4 Actual Web Mining

Matze. Verbindung zur Vorlesung. Web mining techniken. Rapidminer prozess. classifier. Suche nach perfect settings. apply model

1.5 Evaluation

Jochen. Resultate von Precision und Recall der Classification. Diskussion. Positive und negative Einflsse auf ergebnisse. Analog fuer ungelabelte Daten

1.6 Conclusion

Wrap-up.