SRI LANKA INSTITUTE OF INFORMATION TECHNOLOGY
LIBRARY

Opinion Mining and Sentiment Analysis, to Improve Legal Information Efficiency Analysis in Sri Lankan Legal cases

Research Thesis Document

SLIIT
COMPUTING | BUSINESS | ENGINEERING

M.Sc. in Information Technology

Sri Lanka Institute of Information Technology

December, 2014

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Research Supervisor: Dr. Koliya Pulasinghe

Approved for MSc. Research Project:

MSc. Research Project Co-ordinator, SLIIT

Approved for MSc:

MSc. Programme Co-ordinator, SLIIT

Declaration of Originality

This is to certify that the work is entirely my own and not of any other person, unless explicitly acknowledged (including citation of published and unpublished sources). The work has not previously been submitted in any form to the Sri Lanka Institute of Information Technology or to any other institution for assessment for any other purpose.

Signed : - _____

Date : - _____

Research Student: D.A.N. Jayaweera

[PGM-IT12-0417 / MS11905332]

Acknowledgements

This Research work has been possible with the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

My deepest gratitude is to my supervisor, Dr. Koliya Pulasinghe as I have been amazingly fortunate to have an advisor who gave me the freedom to explore on my own and at the same time the guidance to recover when my steps faltered and patience and support helped me overcome many crisis situations and finish this research work. I would like to give my deepest gratitude to Dr. Koliya Pulasinghe and Dr. Samantha Thelijjagoda who taught me the subject "*Research Methodology*", which guided me to understand what research, is, and how to conduct a research.

I would like to show my gratitude to all my colleagues who gave me a moral support in many areas. Last but not the least; I owe my deepest gratitude to my family members to the tremendous support they gave me to make this mission success.

Table of Content

Abstract iii

Acknowledgements iv

Table of Content v

List of Figures viii

List of Tables ix

Chapter 1 1

1. Introduction 1

1.1 Study Background 2

1.2 Study Area 2

1.2.1 Problem Domain 2

1.3 Aim and Objectives 4

1.3.1 Main Objective 4

1.3.2 Sub Objectives 4

Chapter 2 5

2. Literature Review 5

2.1 Introduction 5

2.1.1 What is Opinion Mining? 6

2.1.2 What is Sentiment Analysis? 7

2.1.3 Different Level of Sentiment Analysis 7

2.1.4 Key Components of Opinions 8

2.1.5 Types of Opining 8

2.1.6 Data Preparation 12

2.1.7 Review Analysis 12

2.1.8 Sentiment Classification 12

2.2 Data Mining 14

2.3 Text Mining 15

2.4 Vector space 15

Chapter 3 17

3 Data and Methods 17

3.1 Introduction 17
3.1.1 Descriptive Research *17*
3.1.2 An Exploratory Study *17*
3.2 Usage of Data 18
3.2.1 Secondary Data *18*
3.2.2 Primary Data *18*
3.3 Data Collection 18
3.3.1 Research Population *18*
3.3.2 Time Management *19*
3.3.3 Problems related to Sample *19*
3.3.4 Accuracy of data *19*
3.4 A specific methods 19
3.4.1 The Mining Process *20*
3.4.2 Term Generation *21*
3.4.3 The Research Process *22*
3.4.4 Requirement Gathering and Analyzing *26*
3.4.5 Design *26*
3.4.6 Development *26*
3.4.7 Testing *26*
3.4.8 Maintenance *26*
3.5 A specific method 26
3.5.1 Requirement Gathering and Analyzing *26*
3.5.2 Research requirements and data *26*
3.5.3 Design *28*
3.5.4 Implementation *29*
3.5.6 Testing *30*
3.5.7 Deliverables *30*

4 Results 31
4.1 Algorithm 32
4.2 Evaluation 34

5 Conclusions 37
References 38

Appendices **42**
Appendix A: List of Speech Tags Used 42
Appendix B: List of chunk tags used 44
Appendix C: Law Report Format 45

List of Figures

Figure 1: It's about finding out what people think 6
Figure 2 : Types of Opinion 9
Figure 3 : Task of Opinion Mining 11
Figure 4 : Sentiment Analysis model 12
Figure 5 : Opinion Mining Framework 13
Figure 6 : Overall Architecture of the Proposed System 19
Figure 7 : Architecture of the Mining Process 20
Figure 8 : A sample Part-of-Speech Tagged Text Block 21
Figure 9 : The Research Process 23
Figure 10 : Methods and techniques 24
Figure 11 : Waterfall model 25
Figure 12 : Law Report Format 27
Figure 13 : Analysis of Laws of Sri Lanka 28
Figure 14 : Usecase Diagram 29
Figure 15 : Class Draigram 29
Figure 16 : ER Model 30
Figure 17: Functionalities of the Opinion Mining Intelligent Services 31
Figure 18 : Algorithm 32
Figure 19 : Interface I 33
Figure 20 : Interface II 33
Figure 21 : Interface III 34
Figure 22 : Result of Evaluation for Precision 35

List of Tables

Table 1: Stop Words in the Context of Legal Research 22
Table 2 : Result of Evaluation for Precision 35
Table 3: List of Speech Tags Used 42
Table 4 : List of chunk tags used 44

Abstract

Opinion Mining and Sentiment Analysis, to Improve Legal Information Efficiency Analysis in Sri Lankan Legal cases

D.A.N. Jayaweera

MSc. in Information Technology

Supervisor: Dr. Koliya Pulasinghe

December 2014

Sri Lanka Information Technology field is now growing faster and the other industries try to combine their work with Information Technology to gain more profit and publicity. Sri Lanka law field is one of the most famous and oldest fields. Most of the industry field can combine with Information Technology and do new things as an example store data analyzing data. They need to analyze data because of they need to get correct decision and decision making process. Current Lawyers in Sri Lanka, use manual way to store client details, organize their work etc. Because of that they need lots of document writings, document reading. When the new case acquire to Lawyer, they need to read all history cases and they need to analyze those cases and get the final decision. Searching documents, creating new case files, organize the documents they need to do in efficiency manner. I proposed this system to how to they analyze the legal cases opinion and mining legal cases. Opinions are consequently important that whenever one needs to make a decision, one wants to hear others' opinions. This is true for both individuals and organizations. The technology of Opinion Mining thus has a great scope for practical applications. The concept of opinion is in the context of Sentiment Analysis, the main tasks of Sentiment Analysis, and the framework of opinion summarization. Along with them, two relevant and important concepts of subjectivity and emotion were also introduced, which are highly related to but not equivalent to opinion. Existing studies about them have mostly focused on their intersections with opinion. I create a system to analyze legal information in proper manner. The research is give design the application in the world context and introduce interactive customizable application for the purpose of analyze legal information. As the outcome of the research, it will present the statistics and survey outcomes of usage of the introducing application for legal information analyze in Sri Lanka and make available the developing.

Index Terms: Opinion Mining, Sentiment Analysis, Data Mining, Text Mining