CS 5604 Information Storage and Retrieval
Spring 2016
Interim Report 1

Text Classification

Team Members:

Hossameldin Shahin <hshahin@vt.edu>
Matthew Brockman <mattb95@vt.edu>
Michael Cantrell <mcantrell@vt.edu>

Project Advisor:

Prof. Edward A. Fox

02/12/2016
Virginia Tech, Blacksburg, Virginia 24061, USA

Abstract

Text of the Abstract.

Acknowledgements

I would like to express (whatever feelings I have) to:

- My supervisor
- My second supervisor
- Other researchers
- My family and friends

Contents

Abstract **iii**

Acknowledgements **v**

List of Figures **xi**

List of Tables **xiii**

1 Introduction **1**

1.1 Examples 1

1.2 How Does Sections, Subsections, and Subsections Look? 2

1.2.1 This is a Subsection 2

2 Literature Review **3**

2.1 Examples 3

2.2 How Does Sections, Subsections, and Subsections Look? 4

2.2.1 This is a Subsection 4

3 Requirements, Design, And Implementation 5

3.1 Examples 5

3.2 How Does Sections, Subsections, and Subsections Look? 6

3.2.1 This is a Subsection 6

4 User Manual 7

4.1 Examples 7

4.2 How Does Sections, Subsections, and Subsections Look? 8

4.2.1 This is a Subsection 8

5 Developer Manual 9

5.1 Examples 9

5.2 How Does Sections, Subsections, and Subsections Look? 10

5.2.1 This is a Subsection 10

6 Plan 11

7 Conclusion 15

7.1 Examples 15

7.2 How Does Sections, Subsections, and Subsections Look? 16

7.2.1 This is a Subsection 16

8 Future Work **17**

8.1 Examples 17

8.2 How Does Sections, Subsections, and Subsections Look? 18

8.2.1 This is a Subsection 18

A Appendix A name **19**

List of Figures

List of Tables

6.1 Weekly breakdown of work to be done. 12

Chapter 1

Introduction

Here is the introduction. The next chapter is chapter **??**.

a new paragraph

1.1 Examples

You can also have examples in your document such as in example 8.1.

Example 1.1 (An Example of an Example)
Here is an example with some math

$$0 = \exp(i\pi) + 1 \, . \tag{1.1}$$

You can adjust the colour and the line width in the `macros.tex` file.

1.2 How Does Sections, Subsections, and Subsections Look?

Well, like this

1.2.1 This is a Subsection

and this

This is a Subsubsection

and this.

A Paragraph You can also use paragraph titles which look like this.

A Subparagraph Moreover, you can also use subparagraph titles which look like this. They have a small indentation as opposed to the paragraph titles.

s it possible to add a sub-
subparagraph?

I think that a summary of this exciting chapter should be added.

Chapter 2

Literature Review

Here is the introduction. The next chapter is chapter **??**.

a new paragraph

2.1 Examples

You can also have examples in your document such as in example 8.1.

Example 2.1 (An Example of an Example)
Here is an example with some math

$$0 = \exp(i\pi) + 1 \, . \tag{2.1}$$

You can adjust the colour and the line width in the `macros.tex` file.

2.2 How Does Sections, Subsections, and Subsections Look?

Well, like this

2.2.1 This is a Subsection

and this

This is a Subsubsection

and this.

A Paragraph You can also use paragraph titles which look like this.

A Subparagraph Moreover, you can also use subparagraph titles which look like this. They have a small indentation as opposed to the paragraph titles.

s it possible to add a sub-
subparagraph?

I think that a summary of this exciting chapter should be added.

Chapter 3

Requirements, Design, And Implementation

Here is the introduction. The next chapter is chapter **??**.

a new paragraph

3.1 Examples

You can also have examples in your document such as in example 8.1.

Example 3.1 (An Example of an Example)
Here is an example with some math

$$0 = \exp(i\pi) + 1 \, . \tag{3.1}$$

You can adjust the colour and the line width in the `macros.tex` file.

3.2 How Does Sections, Subsections, and Subsections Look?

Well, like this

3.2.1 This is a Subsection

and this

This is a Subsubsection

and this.

A Paragraph You can also use paragraph titles which look like this.

A Subparagraph Moreover, you can also use subparagraph titles which look like this. They have a small indentation as opposed to the paragraph titles.

s it possible to add a sub-subparagraph?

I think that a summary of this exciting chapter should be added.

Chapter 4

User Manual

Here is the introduction. The next chapter is chapter **??**.

a new paragraph

4.1 Examples

You can also have examples in your document such as in example 8.1.

Example 4.1 (An Example of an Example)
Here is an example with some math

$$0 = \exp(i\pi) + 1 \, . \tag{4.1}$$

You can adjust the colour and the line width in the `macros.tex` file.

4.2 How Does Sections, Subsections, and Subsections Look?

Well, like this

4.2.1 This is a Subsection

and this

This is a Subsubsection

and this.

A Paragraph You can also use paragraph titles which look like this.

A Subparagraph Moreover, you can also use subparagraph titles which look like this. They have a small indentation as opposed to the paragraph titles.

s it possible to add a sub-
subparagraph?

I think that a summary of this exciting chapter should be added.

Chapter 5

Developer Manual

Here is the introduction. The next chapter is chapter **??**.

a new paragraph

5.1 Examples

You can also have examples in your document such as in example 8.1.

Example 5.1 (An Example of an Example)
Here is an example with some math

$$0 = \exp(i\pi) + 1 \, . \tag{5.1}$$

You can adjust the colour and the line width in the `macros.tex` file.

5.2 How Does Sections, Subsections, and Subsections Look?

Well, like this

5.2.1 This is a Subsection

and this

This is a Subsubsection

and this.

A Paragraph You can also use paragraph titles which look like this.

A Subparagraph Moreover, you can also use subparagraph titles which look like this. They have a small indentation as opposed to the paragraph titles.

s it possible to add a sub-
subparagraph?

I think that a summary of this exciting chapter should be added.

Chapter 6

Plan

Please see table 6.1 for the weekly breakdown of work.

Table 6.1: Weekly breakdown of work to be done.

Weeks	End Date	Tasks
Week 1	22 Jan.	Understanding the classification task
Week 2	29 Jan.	• Understanding the classification task • Read about Hadoop streaming using python
Week 3	5 Feb.	Start online tutorials about Hadoop and Apache Spark.
Week 4	12 Feb.	Continue online tutorials about Hadoop and Apache Spark.
Week 5	19 Feb.	**Phase 1 will include only tweets small data set:** • Understanding the classification task • Read about Hadoop streaming using python
Week 6	26 Feb.	• Prepare training data using Solr as described in section • Build classifier using Apache Spark
Week 7	4 March	• Build classifier using Apache Spark • Get output data format from Solr team
Week 8	11 March	Optimize classifier performance

Weeks	End Date	Tasks
Week 9-15	TBD	**Phase 2 will include tweets and web pages:** • Contact collection management team to get the noise reduction techniques and apply it on the web pages. • Repeat the same process done on Phase 1 on web pages. We have ideas to further optimize the classifier performance: • Search for known approachs to select the most representative data sample for each collection. • Integrate hyper parameters optimization library [**bergstra2013hyperopt**] output with Spark.

Chapter 7

Conclusion

Here is the introduction. The next chapter is chapter **??**.

a new paragraph

7.1 Examples

You can also have examples in your document such as in example 8.1.

Example 7.1 (An Example of an Example)
Here is an example with some math

$$0 = \exp(i\pi) + 1 \; . \tag{7.1}$$

You can adjust the colour and the line width in the `macros.tex` file.

7.2 How Does Sections, Subsections, and Subsections Look?

Well, like this

7.2.1 This is a Subsection

and this

This is a Subsubsection

and this.

A Paragraph You can also use paragraph titles which look like this.

A Subparagraph Moreover, you can also use subparagraph titles which look like this. They have a small indentation as opposed to the paragraph titles.

s it possible to add a sub-
subparagraph?

I think that a summary of this exciting chapter should be added.

Chapter 8

Future Work

Here is the introduction. The next chapter is chapter **??**.

a new paragraph

8.1 Examples

You can also have examples in your document such as in example 8.1.

Example 8.1 (An Example of an Example)
Here is an example with some math

$$0 = \exp(i\pi) + 1 \; . \tag{8.1}$$

You can adjust the colour and the line width in the `macros.tex` file.

8.2 How Does Sections, Subsections, and Subsections Look?

Well, like this

8.2.1 This is a Subsection

and this

This is a Subsubsection

and this.

A Paragraph You can also use paragraph titles which look like this.

A Subparagraph Moreover, you can also use subparagraph titles which look like this. They have a small indentation as opposed to the paragraph titles.

s it possible to add a sub-
subparagraph?

I think that a summary of this exciting chapter should be added.

Appendix A

Appendix A name

Here is the first appendix