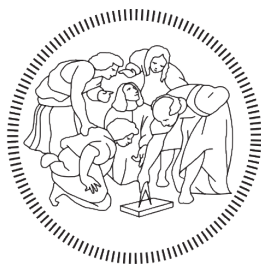


POLITECNICO DI MILANO  
Scuola di Ingegneria Industriale e dell'Informazione  
Corso di Laurea Magistrale in Ingegneria Informatica  
Dipartimento di Elettronica, Informazione e Bioingegneria



**POLITECNICO**  
**MILANO 1863**

Working title

Relatore: Prof.ssa Letizia TANCA  
Correlatori: Prof.ssa Maristella MATERA  
Prof. Riccardo MEDANA

Tesi di laurea di:  
Milica JOVANOVIĆ Matr. 000000  
Mirjam ŠKARICA Matr. 836505

Academic Year 2015–2016



*Some nice inspirational and aspirational quote. Some nice inspirational and aspirational quote.*

*Someone*



# Summary

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.



# Abstract

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.





# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Structure . . . . .	4
<b>2</b>	<b>State of the art</b>	<b>5</b>
2.1	Dummy section . . . . .	5
<b>3</b>	<b>Sentiment analysis workflow</b>	<b>7</b>
3.1	Prediction workflow . . . . .	8
3.2	Determining real sentiment workflow . . . . .	9
3.3	Evaluation workflow . . . . .	10
<b>4</b>	<b>Framework</b>	<b>11</b>
4.1	Design . . . . .	11
4.2	Implementation . . . . .	11
4.3	User interface . . . . .	11
<b>5</b>	<b>Results</b>	<b>13</b>
<b>6</b>	<b>Conclusion</b>	<b>15</b>
<b>7</b>	<b>Future work</b>	<b>17</b>
	<b>Bibliography</b>	<b>19</b>



# List of Figures

3.1	Sentiment analysis workflow . . . . .	7
3.2	Sentiment prediction workflow . . . . .	8
3.3	Determine real sentiment workflow . . . . .	9
3.4	Sentiment evaluation workflow . . . . .	10
4.1	Test caption . . . . .	12



# Chapter 1

## Introduction

Some text [1] Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is

there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

[2] This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

---

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

### 1.1 Structure

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

- In the chapter 2 blahblah
- In the chapter 3 Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.



## Chapter 2

# State of the art

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language. [3]

### 2.1 Dummy section

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words

should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

## Chapter 3

# Sentiment analysis workflow

This chapter describes the workflow used to analyze the sentiment of social media comments and their corresponding posts. In order to outline the workflow, a top down approach was taken where each subsequent section provides an ever more detailed insight into a particular step of the workflow. The big picture is shown in Figure 3.1 and consists of four parts:

1. Obtaining data
2. Sentiment prediction using an API
3. Determining real sentiment of data
4. Evaluation of that API's performance

First part is the simplest one and as such it doesn't merit a more detailed re-counting other than mentioning that we were provided with a small sample dataset which, most relevantly, contained about 6000 comments.

In the sections that follow, each of the three remaining parts are broken down into conceptual steps describing the methodology used whilst not cluttering it with too many implementation details. Additionally, it is interesting to note that the first and third steps are done only once. This means that, for each new API, the workflow for sentiment analysis effectively consists of only steps 2 and 4, namely sentiment prediction and performance evaluation.

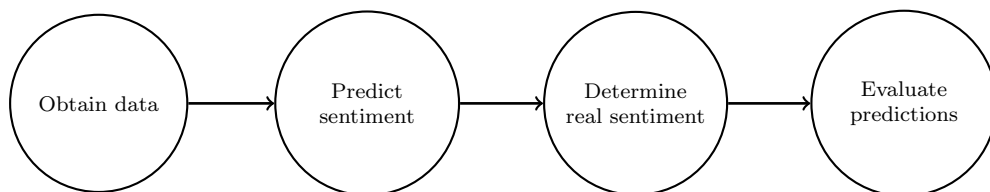


Figure 3.1: Sentiment analysis workflow

### 3.1 Prediction workflow

Figure 3.2 shows the main concepts that build up the workflow of our sentiment analysis. Since the term *workflow* can be a bit ambiguous, let us clarify exactly what we mean by it. In our case it is simply a python script named `automated_sentiment_analysis.py` that can be run manually, or scheduled to run on a server at desired times/intervals. Sections that follow will explain each step in more detail and will also provide motivation for some, perhaps not so obvious, choices.

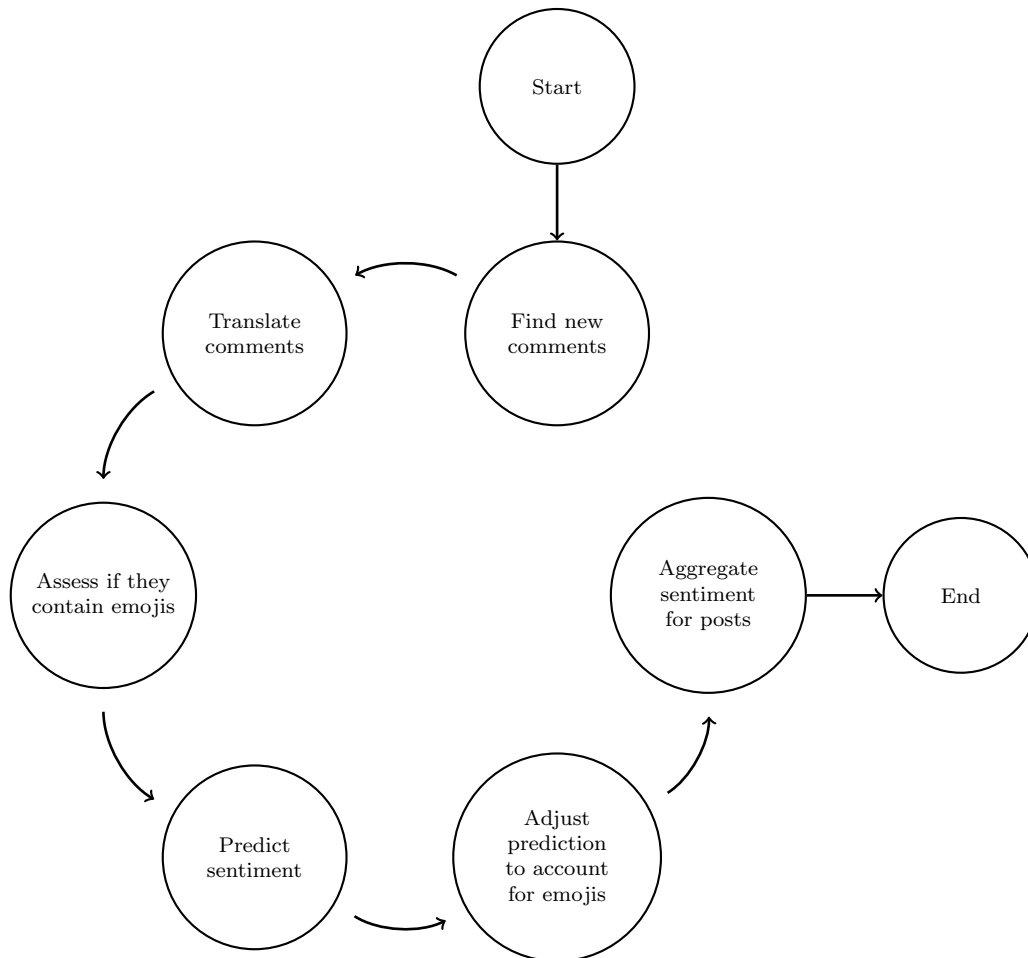


Figure 3.2: Sentiment prediction workflow

#### Find new comments

This part quite straight forward Once run, the script scans the database looking for comments that don't have a sentiment record and inserts one.

for the sake of completeness The inserted rows' sentiment columns default to a json shown in Listing 3.1. The reason for this particular choice of json and for using

the json format in the first place is discussed at length in Section 4.1.

```
{
  "sentiment_label": "",
  "sentiment_stats": {
    "positive": 0,
    "negative": 0
    "neutral" : 0
  }
}
```

Listing 3.1: Default sentiment json

## Translate comments

Translate comments, Mark comments containing emojis, Predict sentiment, Account for emojis, Calculate aggregate sentiment for posts by new we mean unanalyzed

...

...

...

...

## 3.2 Determining real sentiment workflow

How do we know our predictions are any good? real sentiment input can be done either by hand orby the REST API GUI or the REST API curl calls

...

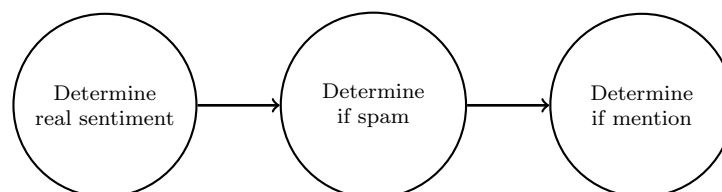


Figure 3.3: Determine real sentiment workflow

### 3.3 Evaluation workflow

performance evaluation of that particular API. how to evaluate? human input!

...

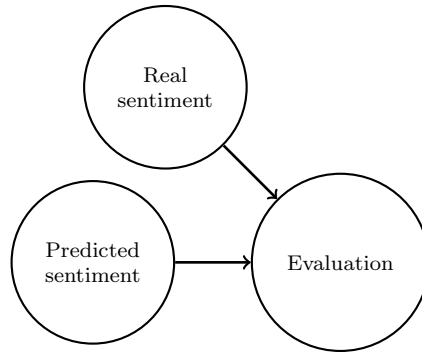


Figure 3.4: Sentiment evaluation workflow

## Chapter 4

# Framework

### 4.1 Design

Why json when it violated the 1NN rule? already in mysql, will eventually support json, and easily movable to nosql db, or even elastic search.

### 4.2 Implementation

### 4.3 User interface

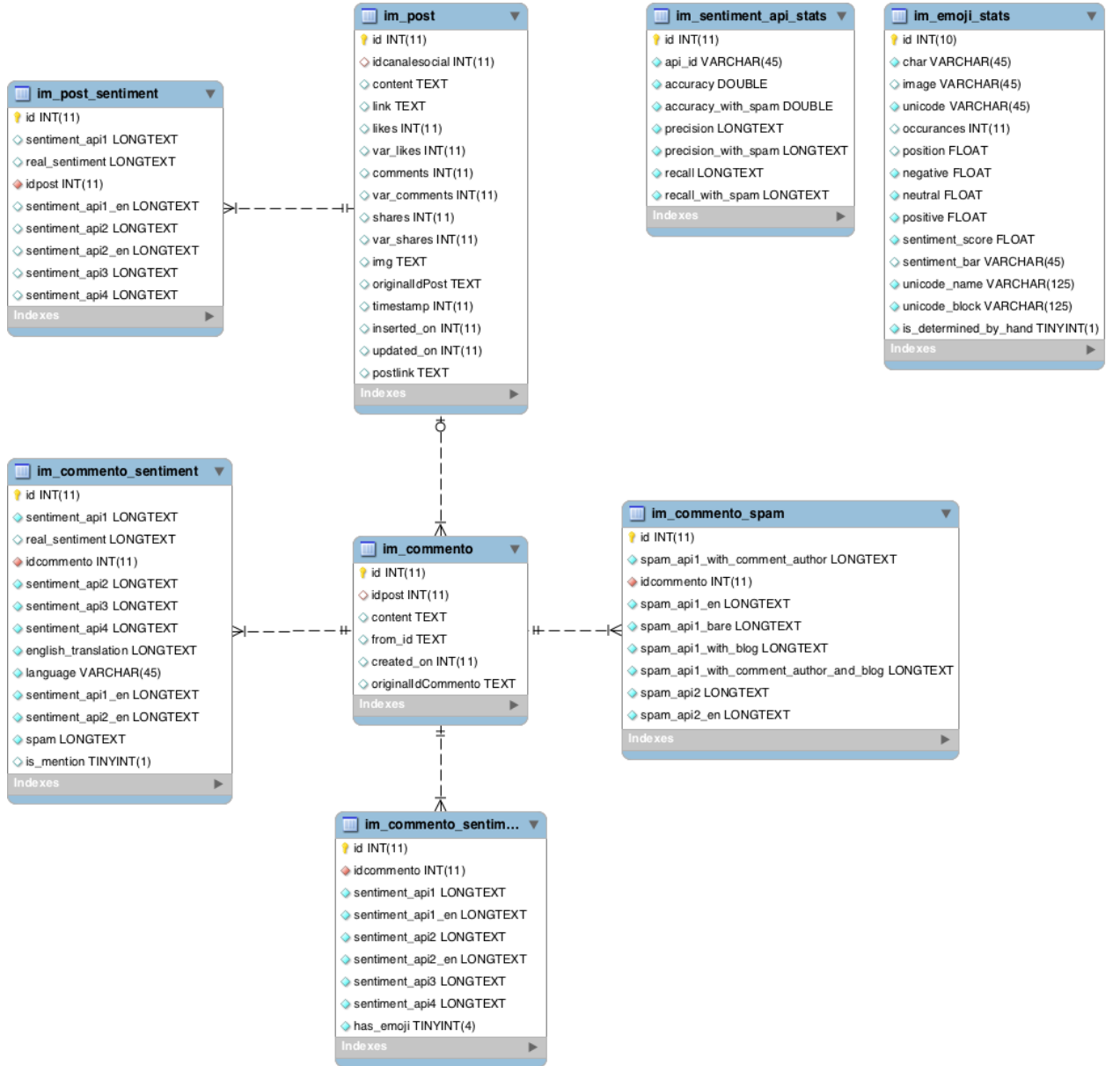


Figure 4.1: TTest caption



## Chapter 5

## Results



## Chapter 6

## Conclusion



## Chapter 7

### Future work



# Bibliography

- [1] M. Masse, *REST API design rulebook*. O'Reilly Media, Inc., 2011.
- [2] F. Corvetta and V. Rizzo, “Progettazione e sviluppo di mashup mobile context aware: il progetto camus,” Master’s thesis, Politecnico di Milano, 2015.
- [3] M. Odell and R. Russell, “The soundex coding system,” *US Patents*, vol. 1261167, 1918.