

# Assignment two: Detection opinion spam

Mike Vink

October 23, 2020

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Dataset construction</b>	<b>1</b>
2.1	Selection of data . . . . .	1
2.2	Preprocessing of the textual data . . . . .	2
2.3	Construction of derived features . . . . .	2
<b>3</b>		<b>2</b>
<b>4</b>	<b>Deception classification results</b>	<b>2</b>
<b>5</b>	<b>Conclusion</b>	<b>2</b>
<b>6</b>	<b>References</b>	<b>2</b>

## 1 Introduction

	qwoiefjoij oiwwqofij
hello im data	

## 2 Dataset construction

### 2.1 Selection of data

The dataset used in this paper is derived from the publicly available dataset of *negative* deceptive opinion spam contributed by [1], containing 400 gold standard

deceptive negative reviews of 20 popular Chicago hotels, and 400 truthfully negative reviews. This data consists of text files that were used to construct our dataset, no cleaning of the data was necessary in this case.

## **2.2 Preprocessing of the textual data**

## **2.3 Construction of derived features**

## **3**

## **4 Deception classification results**

## **5 Conclusion**

## **6 References**

### **References**

- [1] Myle Ott, Claire Cardie, and Jeffrey T Hancock. Negative deceptive opinion spam. In *Proceedings of the 2013 conference of the north american chapter of the association for computational linguistics: human language technologies*, pages 497–501, 2013.