

DEVELOPING A NEURAL BAG-OF-WORDS MODEL FOR SENTIMENT ANALYSIS

Nijat Zeynalov

University of Tartu, The School of Economics and Business Administration

zenijat16631@sabah.edu.az

ABSTRACT: Sentiment Analysis is a fundamental task in Natural Language Processing (NLP). The goal of most sentiment tasks is to identify the overall sentiment polarity of the documents in question, i.e. is the sentiment of the document positive or negative? [3]. For our case, we use online user reviews from the most popular Azerbaijan movies. In order to perform this sentiment task, we use a mixture of baseline machine learning models and deep learning models to learn and predict the sentiment of binary reviews. This poses a supervised learning task. [2] For this purpose, we will use bag-of-words features with machine learning algorithms which is very simply and efficient, while having the ability to achieve very high accuracy. In this paper, we will develop a neural bag-of-words model, which collects high-level structural and semantic meaning of the words. The paper performs sentiment classification via neural bag-of-words approach.

KEYWORDS: sentiment analysis, bag-of-words model, text analysis, machine learning, deep learning, MLP

1. INTRODUCTION

The main issue in the analysis of sentiments is understanding how sentiments are conveyed in reviews and whether the expressions refer (acceptable) or negative (unacceptable) reviews toward the subject. [3] Analysis of sentiments is a laborious task [4] to conduct with computers and algorithms. It is difficult for machines to recognise certain patterns, or even unlikely, although it is easy for humans. For machines, there are certain intractable cases, such as: Can't grasp pronouns and what they refer to. We can not grasp the various definitions of words like irony or sarcasm, too. Evaluation of feelings involves a broad lexicon to be produced which takes a long time to search for the words needed. [8] In determining terms between two polarities, the Hardness: positive and negative. Sometimes we need to understand some language. Other problems face sentiment analysis in understanding reviews meaning and grammar for