Chapter 1

Natural Language Processing via Deep Learning

Bárbara Stéphanie Neves Oliveira, Luís Gustavo C. do Rêgo, Lucas Peres, Ticiana L. Coelho da Silva, José Antônio F. de Macêdo

Abstract

Humans need to communicate. Out of this basic need combined with the Web, a vast amount of text has been generated on a daily basis. Given the presence of a lot of information allocated in different resources, it becomes vital to enable machines to understand spoken and written texts. This chapter presents how Deep Learning techniques can solve Natural Language Processing (NLP) tasks (eg, Text Classification and Sentence Summarization), aiming to benefit from the computational power currently available and the low need for feature engineering when using these models. Initially, some essential concepts about NLP and Deep Learning are presented. Then, different pre-processing and textual Representation techniques are explained to be used as input in Deep Learning models. Finally, it is shown how to apply the knowledge acquired in real applications of NLP.

Summary

Human beings need communication. From the combination of this basic need
With the Web, a large amount of text has been generated daily. Given the presence of a lot of information allocated in different media, it becomes vital to allow
machines understand spoken and written texts. This chapter presents how techniques of Deep Learning can be used to solve Natural Language Processing (NLP) tasks, such as Sentence Classification and Summarization,
aiming to benefit from the computational power currently available and the low need for feature engineering when using these models. Initially, some important concepts about NLP and Deep Learning are presented. Right away,
Different pre-processing and textual representation techniques are explained in order to be used as input in Deep Learning models. Finally, it is shown
how to apply the knowledge acquired in real PLN applications.

1.1. Introduction

Communication, as a basic need of the human condition, together with existence of the Web allows a vast amount of written and spoken texts to be generated daily. Given the textual content present in social media, applications chat, emails, product reviews, news articles, research papers and ebooks, the existence of automatic text processing has become vital in order to offer assistance or make decisions for various daily tasks.

The ability to understand text or audio in natural language by a machine It is a problem that has been investigated for a long time [Chollet 2021]. The first attempts to build Natural Language Processing (NLP) systems Natural Language Processing or NLP) were done through the intrinsic analysis of languages that are naturally shaped by a process of evolution (hence the term "Natural"). Modern PLN involves not only the ability to understand a language but also allows, automatically, the extraction of information through through tasks such as Text Classification, Named Entity Recognition (NER), Word Meaning Disambiguation, and Part-of-Speech (POS) tagging.

Deep Learning models (also known as Deep Learning or Deep Neural Networks) learn at various levels representation of data complexity/abstraction in an increasing manner. Several factors show why these models have been widely used in NLP tasks: (i) require little feature engineering; (ii) produce vector representations that capture similarities of linguistic units (words, for example) present in texts, allowing NLP systems to have a kind of knowledge dependence; (iii) allow unsupervised or semi-supervised learning, which is important when you have a large volume of data and no labels; (iv) learn several levels of representation, allowing the lowest level to generally be shared between different tasks; and (v) naturally deal with the recursion of the human language, being able to capture information sequentially.

The use of Deep Learning models in PLN began with the investigation of the language understanding capacity of Recurrent Neural Networks (RNNs, Recurrent Neural Networks) and LSTM (Longest Shortest Term Memory) networks [Hochreiter and Schmidhuber 1997]. These two architectures dominated PLN in a general from 2015 to 2017, as they process texts of variable length. Bidirectional LSTM models, in particular, have defined the state of the art in many tasks. important topics, from Text Summarization to Automatic Translation. However, around In 2017 and 2018, a new architecture emerged to “replace” RNNs: Transformer [Vaswani et al. 2017], which allowed PLN to make considerable progress in a short period of time.

The purpose of this chapter is not to expose all of these applications and architectures. comprehensively. Instead, the focus is on how to practically apply existing textual representations obtained through Deep Learning techniques to solve PLN problems. This chapter also addresses the different text processing steps performed before training a neural network for an NLP task. To the