Chapter 15
Deep Learning Approaches for Affective Computing in Text

Ramón Zatarain Cabada
https://orcid.org/0000-0002-4524-3511
Tecnologico Nacional de Mexico, Culiacan, Mexico

María Lucía Barrón Estrada
https://orcid.org/0000-0002-3856-9361
Tecnologico Nacional de Mexico, Culiacan, Mexico

Víctor Manuel Bátiz Beltrán
https://orcid.org/0000-0003-4356-9793
Tecnologico Nacional de Mexico, Culiacan, Mexico

ABSTRACT

The field of natural language processing (NLP) is one of the first to be addressed since artificial intelligence emerged. NLP has made remarkable advances in recent years thanks to the development of new machine learning techniques, particularly novel deep learning methods such as LSTM networks and transformers. This chapter presents an overview of how deep learning techniques have been applied to NLP in the area of affective computing. The chapter examines traditional and novel deep learning architectures developed for natural language processing (NLP) tasks. These architectures comprise recurrent neural networks (RNNs), long short-term memory (LSTM) networks, and the cutting-edge transformers. Moreover, a methodology for NLP method training and fine-tuning is presented. The chapter also integrates Python code that demonstrates two NLP case studies specializing in the educational domain for text classification and sentiment analysis. In both cases, the transformer-based machine learning model (BERT) produced the best results.

DOI: 10.4018/979-8-3693-0502-7.ch015

INTRODUCTION

Natural Language Processing (NLP) is the study of how computers can recognize and alter human expressions in text. NLP integrates several fields, including computer science, linguistics, mathematics, Artificial Intelligence (AI), and psychology. NLP is concerned with a wide range of topics, including automatic text translation, multilingual user interfaces, and speech recognition, among others (Chollet, 2021). In relation with that, Machine Learning (ML) is one of the most popular techniques for implementing applications that use NLP. Broadly speaking, ML is a branch of AI that gives computers the ability to learn without being explicitly programmed to do so. ML algorithms allow the computer to learn from data without having to write any custom code specific to a particular problem. For example, an ML algorithm is capable of learning to classify and obtain the polarity (positive or negative) of an opinion given as input without the need to build a new algorithm that is specific to the solution of that problem.

Deep learning (DL) is a sort of machine learning that mimics how humans learn (cognitive learning process). DL algorithms, as opposed to traditional machine learning algorithms, are a hierarchical series of increasing complexity and abstraction (LeCun et al., 2015). This hierarchical learning is accomplished using artificial neural network methods, which are especially useful when there are a high number of samples describing an event. A model is a hierarchical depiction of the architecture of an artificial neural network. DL has made it possible to obtain important advances in various tasks that were previously carried out with ML techniques. It has revolutionized NLP by enabling the development of powerful models that can understand and generate human language with remarkable accuracy.

This chapter aims to provide an objective overview of the application of deep learning techniques in NLP. Two case studies are presented as the main contributions. The first case study focuses on text classification and aims to address the research question: what is the best deep learning classification algorithm to detect learning-centered emotions in a text dataset? The second study focuses on the particular task of sentiment analysis and seeks to answer the research question: what is the most effective deep learning classification algorithm to recognize negative/positive opinions in a text dataset?

This chapter is organized as follows. The chapter begins with a section named Affective Computing, that provides a basic understanding of the underlying concepts. Then discusses popular deep learning architectures designed specifically for NLP tasks. Recurrent Neural Networks (RNNs) are discussed, highlighting their ability to capture sequential dependencies and their applications in tasks such as text classification, language modeling, and machine translation. Long Short-Term Memory (LSTM) networks, a variant of RNNs, are explored for their effectiveness in handling long-range dependencies and mitigating the vanishing gradient problem. The innovative impact of transformers in NLP is thoroughly explored. These attention-based architectures have revolutionized the field, powering state-of-the-art models such as BERT. The chapter explains the basic principles of transformers and their applications in tasks such as text classification, and sentiment analysis. In the following section a methodology for training and tuning NLP task approaches is also presented. Based on this methodology, code in the Python language is presented to fully cover two NLP case studies. Both focused on the educational domain and using two different datasets to cover the tasks of text classification and sentiment analysis. Finally, the chapter finishes with some concluding remarks and suggestions for future work.

Related Content