ABSTRACT

Convolutional neural networks are known for their excellent performance in computer vision, achieving results in the state of the art. However, recent work has shown that these networks can also work for natural language processing. In this case, the basic idea is to concatenate the vector representations of words into a single block and use it as an image. However, despite the good results, the problem of convolution networks is the large numbers of design decisions that need to be made. These models require the definition of many hyperparameters, including the type of word embeddings, which consists of the vector representation of the data, the activation function that introduces nonlinearity to the model, the size of the filter that applies convolution to the data, the number of feature maps which are responsible for identifying the attributes, the pooling method used in data reduction, in addition to the regularization constant and the dropout rate, which are responsible for avoiding overfitting of the network. In existing works, convolutional neural network architectures capable of overcoming the performance of traditional machine learning models were presented and these can compete with more complex models. However, it has not been explored how the different possibilities of hyperparameters can affect the performance of this type of network. In this dissertation, the objective is to create an efficient sentiment analysis classifier using convolutional neural networks by analyzing the impact of their hyperparameters on model performance. The interest in analyzing sentiment comes from the advent of social media and the technological advances that flood the Internet with opinions. The results achieved were obtained with the use of GPU and show that the different configurations exceed the reference models in the majority of them with gains of up to 18% and have similar performance to the models in the state of the art with gains of up to 2% in some cases.

Keywords: Convolutional Neural Network. Sentiment Analysis. Hyperparameters. Natural Language Processing.