Abstract—

I. Introduction

Understanding the sentiment of the text is a vital component for many different industries. For instance, providing insight into the sentiment of a movie script could provide hidden knowledge towards the population it should direct. Selecting the right types of movies for consumers is key to businesses such as Netflix and other providers of video content. A movie containing mainly positive or negative content can suggest the opinion of a consumer towards certain types of sentiment. The recommendation of movies to a user can be altered based on their preferences.

Past approaches focused on the formation of the input shapes rather than modifying the model to the task. Modeling language is not as simple when more complex representations are introduced to the problem. Noise is contained in the language no matter the form. Short or long shapes the texts are going to contain noise that cannot be avoided. Deep models need an ability to adjust to the weights applied throughout the networks layers in order to learn deeper features throughout the network. Not all representations should adjust the model weights significantly unless the input warrants it. The PTCN is implemented assuming that most of the noise in language samples is primarily junk data that helps transfer information more smoothly when humans talk. In text classification tasks the importance of transfering information in the same manner is not relevant. The texts are processed and automatically classifies the sentiment of the text, either positive or negative utilizing a custom hybrid deep neural network algorithm. The adaptable convolutional and long-short-term-memory neural networks provide a novel solution to adjusting the initialization and regulation of the kernels. The highly sparse nature of the language utilized within texts is difficult to adjust to; however, the adaptability of the network to the structure of the inputs helps identify latent patterns throughout the samples. The adaptable nature of the algorithm tackles high variance in the input's dimensions by allowing adjustment to the weights of the inputs. Adjusting the convolutional layers to extract features under controlled but adaptable parameters produces an accurate prediction of sentiment from the movie script texts and the tweets. The model is able to be improved to competitive results compared to prior research classifying the sentiment of the texts utilizing the practice of hyper-tuning and adaptable kernel intializers.

Promoting a network to only recognize significant features that help class samples of highly sparsity texts based on their sentiment provides a valuable advantage. Tokenized word vector spaces are perfect representations of language as they represent the purest form to shape an input for deep neural networks. Both conditions allow the PTCN to capture the most relevant features for classification.

The experiments express the ability of the PTCN in recognizing sentiment from both short and long text inputs. The long texts a corpus of movie scripts and the short texts are a corpus of tweets. Being able to perform well on both samples expresses the PTCN cpabilites to adjust to inputs. The PTCN even performed well between the change in the use of the langauge from corpus to corpus. Unlike past approaches the PTCN does not require heavy pre-processing mantaining the inputs purest representation of the language. The combination of differnt types of neural layers provides the successful solution to text classification.