

Deep Neural Networks in Text Classification using Active Learning

Neural Networks (NNs) and Natural language processing (NLP) are getting more advanced and more useful in industries and human life. One of the subsets in Deep Learning is active learning (AL) which the model is able to query user operators of humans while the learning process is going on. Natural language processing (NLP) is a process to extract the desired goal from the texts created by translations, speech, auto-captioning and searching. The classification of the texts are the main role in processing natural language. In order to get this goal, the Neural Networks text classification implementation in artificial intelligence (AI) methods, has a key role in increasing the performance of a model with the same data size or even less. So, the required annotation tries will keep the performance of the model when we do not have that much data. In this short story, the text classification processes will be reviewed and will be focused on deep neural networks (DNNs) usages. Then it will be discussed about the two main reasons which used to in case of the adoption hindrance:

- (a) the Neural Networks inability to make available reliable measures of ambiguity on which the most frequently deployed query strategies
- (b) the training DNNs difficulties when we have small data.

We are building up a taxonomy of query techniques in order to analyze the first approach, which differentiates between data-based, model-based, and prediction-based instance collection and investigates their prevalence in recent investigations.

In addition, we study recent NN-based developments in NLP in the sense of (D)NNs, survey the existing state-of-the-art at the convergence of Artificial intelligence, text classification, and DNNs, such as sentiment analysis or language models, and apply recent developments in NLP to Artificial intelligence.

In the end, for text classification, we review the latest work in Artificial Intelligence, link the corresponding question methods to the categorization, and detail similarities and drawbacks. As a consequence, we identify gaps in the existing science and current accessible research problems.