Ph.D. Dissertation
S.S.J University, Almora

Hindi Tweets Sentiment Analysis Using Transfer Learning

Department Of Computer Science

By ANUBHAV MEHRA

Acknowledgment

It is a genuine pleasure to express my deep sense of gratitude to my mentor and my Ph.D. supervisor Dr. Manoj Singh Bisht. His advice and approach have helped me to a very great extent to accomplish this task.

I would also like to extend my gratitude to the Department of Computer Science, S.S.J University, Almora and all its faculty members who have been of great help during the course of this project.

Finally I would like to thanks my head Dr. Ashish Mehta, Department of Computer Science, DSB Campus Nainital for his immense help and providing me with time to complete my dissertation work.

Contents

Introduction

Abstract

Sentiment analysis is a **natural language processing** technique to find if the sentiment of the text is positive, neutral or negative. Traditionally, to train a model for sentiment analysis require very dense neural networks to train on very huge datasets. But, here we have used a technique called **Transfer Learning** that stores a model which has learned some knowledge, that we can leverage in solving some other tasks based on the knowledge of the previous model. Here we are using a language model called **BERT(Bidirectional Encoder Representations from Transformers)**. BERT is a pertained model which learns using the learning techniques developed by Google. The BERT multilingual base model that we are using is pertained on the top 104 languages including Hindi. We then leverage the power of this model for the Sentiment analysis of the Hindi texts dataset that we've got. This allows us to achieve moderately high accuracy scores using a comparatively small dataset.

Transfer Learning is a Machine Learning method where a model that is trained for a certain task is utilized as the starting point for solving some other task i.e., to train a second model from the knowledge learned from the first model as well as the dataset. It is a very popular approach in natural language processing domain to solve problems such as getting the context of the text.

The inspiration for transfer learning comes from us - humans, ourselves - where in, we have an inherent ability to not learn everything from scratch. We transfer and leverage our knowledge from what we have learn in the past for tackling a wide variety of tasks.[1]

One such problem is **Sentiment Analysis**. Sentiment Analysis or emotion AI, is the process in natural language processing of subjective emotional analysis of the text. Primarily sentiment analysis finds if the emotional tone of a piece of writing is ***positive, neutral or negative***. **Table 1** lists some examples of what a sentiment analysis categorization may look like.

Table 1: Sentiment Analysis Categorization.

Text	Category
That restaurant has a great food	Positive
He is my brother's colleague	Neutral
Bollywood movies are not entertaining	Negative

As, we can see it is easily understood by a human brain what sentiments these pieces of writing represent. But, for a computer this is a very challenging problem.

References

[1] D. D. Sarkar. (Dec. 13, 2018). "Deep transfer learning for natural language processing — text classification with universal…," Medium, [Online]. Available: https://towardsdatascience.com/deep-transfer-learning-for-natural-language-processing-text-classification-with-universal-1a2c69e5baa9 (visited on 07/25/2021).