ESTIMATING SEARCH RELEVANCE USING MODERN DEEP NEURAL NETWORKS

Hanish Pallapothu
Stevens Institute of Technology
hpallap1@stevens.edu

ABSTRACT

Most of the websites today have search functionality to allow users to search for various documents such as products, user posts, text articles etc. It is important to suggest semantically relevant documents to user. The goal of this project to predict the relevance score given product information and search term. The relevance score describes how relevant a product is to the given search term. The idea is to use state of the neural network models to predict the relevance score. I have trained various neural network models including Bidirectional LSTM, One-dimensional Convolution, Attention Mechanism, Transformer based models such as BERT and Sentence BERT to solve the problem. I have evaluated each of the model using RMSE as a metric. While many models perform well, most of the modern neural networks has many parameters and take significant amount of time predicting the relevance. It is also important to consider time take to predict the relevance as a constraint while evaluating models.

1 Introduction

Search relevancy is an important measure we can recommend users with products for a given search term. Currently, human raters evaluate the impact of potential changes to their search algorithms, which is a slow and subjective process. A company could have thousands to millions of products. If there are n products and m search terms, we need to find relevancy for around $n*m$ combinations of products and search terms. It is not feasible to do this task manually. Often times we need to find the relevance score from multiple resources like product title, description and attributes of the product, it makes the problem even harder. Many researchers proposed various deep neural networks to find word embeddings that help find semantically similar documents. If we could build a model to predict the relevance score given product information and search term, it would save lots of time and decrease manual cost. Further, training such models will help us to build semantic search engine. Lets suppose whenever user enters a search term in the search bar, we can use the trained model to predict the relevance score between the the entered search term and all the products in the website. We will display the top 10 or 20 products with higher relevancy score to the user. while building such search engines is not the scope of this project, we can use trained models to build such systems. Building such systems will help us display semantically similar products to the user.

2 Background Work

Traditionally, Lexical similarity is used to find the relevant documents for the given search term. Keyword matching is used to find the relevant documents where the model look for literal matches of the query words or variants of them, without understanding the overall meaning of the query. To overcome this problem, Semantic similarity is used to find the semantically relevant documents. It is a technique in which a search query aims to not only find keywords but to determine the intent and contextual meaning of the words a person is using for search. The theory of semantic similarity goes as far back as 2003, and a paper [5] written by R. Guha et al., of IBM, Stanford, and W3C, where they introduced Semantics for Text data. Word Embeddings were initially estimated using techniques such as Latent semantic analysis to analyze relationships between a set of documents and the terms they contain by producing a set of concepts related to the documents and terms. However, these techniques faced problems such as synonymy and polysemy. Advances in Deep Neural Networks were helpful in overcoming such problems by finding semantically valid word embeddings. Recently transformer-based models such as BERT, GPT2, XLNet to name a few have gained popularity as they outperform all existing models

3 Experimental Design

I will use the data provided by Home Depot on the Kaggle website [6]. The data set contains a number of products and real customer search terms from Home Depot's website. Every product has a title, description, and certain attributes that describe it. There are 74607 data points in the train data set. The data also has a ground truth score of how relevant the search term is for a given product. To create the ground truth labels, Home Depot has crowd sourced the search/product pairs to multiple human raters. The relevance is a number between one to three. For example, a search for "AA battery" would be considered highly relevant to a pack of size AA batteries with relevance equal to three, mildly relevant to a cordless drill battery with a relevance equal to 2, and not relevant to a snow shovel with relevance equal to 1. Each pair was evaluated by at least three human raters. The provided relevance scores are the average value of the ratings. I will use the title, description, and also associated attributes of the product to predict the relevance of a given search term.

4 Section 3

5 Section 4

References