World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:17, No:09, 2023

TransDrift: Modeling Word-Embedding Drift Using Transformer

Authors : Nishtha Madaan, Prateek Chaudhury, Nishant Kumar, Srikanta Bedathur

Abstract : In modern NLP applications, word embeddings are a crucial backbone that can be readily shared across a number of tasks. However, as the text distributions change and word semantics evolve over time, the downstream applications using the embeddings can suffer if the word representations do not conform to the data drift. Thus, maintaining word embeddings to be consistent with the underlying data distribution is a key problem. In this work, we tackle this problem and propose TransDrift, a transformer-based prediction model for word embeddings. Leveraging the flexibility of the transformer, our model accurately learns the dynamics of the embedding drift and predicts future embedding. In experiments, we compare with existing methods and show that our model makes significantly more accurate predictions of the word embedding than the baselines. Crucially, by applying the predicted embeddings as a backbone for downstream classification tasks, we show that our embeddings lead to superior performance compared to the previous methods.

Keywords : NLP applications, transformers, Word2vec, drift, word embeddings
Conference Title : ICSOC 2023 : International Conference on Service-Oriented Computing
Conference Location : Vancouver, Canada
Conference Dates : September 25-26, 2023