基於 Transformer 及遞歸神經網路的自然語言情感分析方法

摘要

本計畫的目的是以遞歸神經網路(Recurrent Neural Network)實現 Transformer 注意力機制(Attention)並結合的過程。本研究來源以 BERT (Bidirectional Encoder Representation from Transformers) 、以 RNN 實現單頭注意力(Single Headed- Attention)及 Transformer 注意力機制為主軸進行研究。本研究計畫所使用的結合方法是基於 RNN 建構自然語言推理模型，並透過改造後的長短期記憶模型(LSTM)實現單頭注意力機制改良 Transformer 自我注意力機制弱於捕獲文本中的局部依賴問題，並取代 Transformer 中解碼器的注意力機制，進行自然語言情感分析任務。進一步以能進行平行處理的 Transformer 為基礎並融合改造後的 LSTM，改良現有的自然語言情感分析任務。本研究所獲得的研究成果可以輔助並應用於情感分析、社交網路分析及疾病傳播預測。

A Transformer and Recurrent Neural Network based Sentiment Analysis Method

Abstract

The dominant sequence transduction models are based on complex recurrent neural networks or Transformer models that include an encoder and a decoder. In recent years, language models are rising with attention mechanism proposed in Transformer before. The best performed language model also connect encoder and decoder with attention layers. The source of this research is based on BERT (Bidirectional Encoder Representation from Transformers), SHA-RNN (Single Headed-Attention

with RNN), and attention mechanism from Transformer as the main axis. In previous work, Transformer works well whenever long range sequence is available, but they cannot be used to capture short dependency from a sequence. In this research, we present a general language model to improve the defect of capturing short-range sequence in Transformer. Our model replaces the attention layer in decoder from Transformer with SHA-RNN because doing so introduced many short-term dependencies between the source and the target sentence which make the optimization problem easier. The result of this research is developed for improving specific tasks well by applying to sentiment analysis, social media analysis, and disease spreading prediction.