Amirkabir University of Technology

Neural Networks

HW8

Transformer

BERT is a powerful NLP model that understands words in context from both directions. It uses a transformer architecture, is pre-trained on lots of text, and can be fine-tuned for specific tasks. BERT's contextual understanding makes it excel in various language tasks and has set new performance standards in NLP such as sentiment analysis. It's open-source and widely used in AI and research.

Transfer learning is a machine learning technique where a model trained on one task is adapted or fine-tuned to perform a different but related task. Instead of training a model from scratch, transfer learning leverages the knowledge and representations learned from one task and applies them to another, potentially saving time, computational resources, and data.

Q2)

In this section, we focus on using a pre-train transformer for sentiment analysis. A more detailed explanation of sentiment analysis:

Text Input: Sentiment analysis takes text as input. This text can be in the form of sentences, paragraphs, or entire documents, and it's typically extracted from sources like customer reviews, social media posts, news articles, or surveys.

Text Preprocessing: Before analysis, the text is often preprocessed to remove noise, such as special characters or punctuation, and to tokenize it into words or phrases.

Sentiment Classification: Sentiment analysis algorithms then classify the text into one or more predefined sentiment categories. The most common categories are:

Positive: Expressing a favorable or happy sentiment.

Negative: Expressing an unfavorable or unhappy sentiment.

Neutral: Lacking a clear positive or negative sentiment.

Here you should classify sentiment into 5 categories. Perform sentiment analysis using the BERT neural network and apply transfer learning during training. Use the HW06 dataset.

Q3)

Highlight two additional tasks that can be addressed using transfer learning models. Select two distinct problems within the specified domains (e.g., image analysis, text processing, or others) and utilize pre-trained transformers to produce the desired outputs from arbitrary inputs.