
FASHION ATTRIBUTES CLASSIFICATION AND SENTIMENT ANALYSIS: A DEEP LEARNING APPROACH

AI6127 Deep Natural Language Processing

Riemer van der Vliet
G2304212K

2024 - March - 12

1 Introduction

This report presents a comprehensive study on two major projects: the Fashion Attributes Classification Challenge and sentiment classification using the iMDB dataset. The former is an individual project aimed at identifying attribute labels depicted in fashion photographs using a dataset of 6000 images, while the latter explores sentiment classification through the application of Vision Transformers (ViT) and fine-tuning techniques using the Optuna framework.

1.1 Description of data

The Fashion Attributes Classification Challenge utilizes a dataset with 6000 images, divided into 5000 images for training and 1000 for validation. The dataset employs 26 attribute labels across 6 major categories, presenting a multi-label classification problem.

The sentiment classification project explores the use of the iMDB dataset, a benchmark dataset in natural language processing for sentiment analysis, through advanced vision-based models, demonstrating the adaptability of vision transformers beyond traditional image tasks.

1.2 Description of model

1.2.1 ViT

The ViT model represents a shift from conventional convolutional neural networks to transformers applied to sequences of image patches for classification tasks. Its application in the sentiment classification project underscores the versatility and robustness of transformers in handling diverse data types. Within this architecture some of the final encoder layers, which have architecture as seen in figure ?? have been unfrozen.

1.2.2 Weight regularization

Because we work with a very uneven dataset I have opted to regularize and normalize for the within class differences -; need plot of the differences in relative abundance.

1.2.3 MultiTaskHead

The MultiTaskHead module is designed to output predictions for multiple attribute labels simultaneously. This approach is crucial for the Fashion Attributes Classification Challenge, allowing for efficient and effective multi-label classification.

1.2.4 Optuna

Optuna is an optimization framework used to fine-tune hyperparameters for deep learning models. In our study, it played a significant role in optimizing the learning rate and the number of layers to unfreeze in the ViT model, leading to improved performance on the sentiment classification task.

During the training the metric for accuracy is decided on pergroupAverage * totalAccurate as to insentivise Optuna to take both of these into account

2 Results

The implementation of the ViT model with a custom MultiTaskHead and the application of Optuna for hyperparameter tuning yielded promising results in both projects. In the Fashion Attributes Classification Challenge, **the model achieved competitive accuracy**, demonstrating the effectiveness of transformers in complex multi-label classification tasks. Similarly, the sentiment classification project saw notable

3 Bibliography

no external sources have been used that need to be cited.