Conquering Fashion MNIST with CNNs using Computer Vision

Team-1010
Manipal Institute of Technology
Akshat Garg-210953372
15th July 2023

Abstract

The Fashion MNIST problem is a computer vision classification problem in which images are supposed to be classified into multiple categories. The dataset contains of 70000 images of dimension 28X28 out of which 60000 are in the training set and 10000 are there in the test set . Simple neural networks can be used to classify but they compromise accuracy which can be increased using a convolutional neural network.

Introduction

The Fashion MNIST is a very popular dataset which is used to test the capability of a classification model, it is the more complex version of the MNIST dataset. The dataset's easy availability , simplicity and relevance makes it very suitable to use in real world applications.

NEED FOR THE PROBLEM

There are many datasets including the original MNIST dataset but due to the advancement of technology leading to even more powerful models the relevance of other models has slowly faded. This dataset is also called the benchmark dataset for the modern-day classification model. Using this model can be used to analyse the current fashion trends, making predictions and even designing personalised recommendation systems.

Hence the need is to advance the classification models and problems in-order to fit them properly to this ever-evolving technological world.

APPROACH

In the start I used a normal nn to check but found that it's not good in terms of the accuracy. The final approach I have used is to build a CNN model using TensorFlow. The model contains 20 layers having multiple convolutional layers, dense layers, batch normalisation layers, dropout layers etc. Each of the layer and the values I have slowly checked and tweaked in order to get the best result Following are some of them -:

- Convolutional Layer- This layer has many filters ranging from 32 to 512 which basically extract the meaningful feature from the input data
- Batch Normalisation Layer-This layer helps in improving convergence and stability of the model
- Dropout Layer- The dropout layer stops some of the neurons randomly to prevent overfitting of the training dataset into the model
- Flatten Layer – It changes the 2D feature into 1D vector for further layers
- Dense- Also known as the fully connected layer connects the previous layer to the next layer
- Activation Functions- Functions such as RELU and SoftMax are used to activate different layers of the neural network

RESULTS

Using three different ways has led to different results-:

- TensorFlow Neural Network
  A simpler model but sacrificed on accuracy has led to an accuracy of 82.96% on the test set
- TensorFlow Convolutional Neural Network
  A complex model with 20 layers takes some time but has led to an accuracy of 92.04% on the test set
- Pytorch Convolutional Neural Network
  This model is very similar to the tensorflow one but this yields an accuracy of 86.84% and when optimised with IPEX it leads to 86.89%

REFERENCES

Following are the references taken-:

- https://blog.tensorflow.org/2018/04/fashion-mnist-with-tfkeras.html
- https://www.tensorflow.org/tutorials/keras/classification
- https://medium.com/@aaysbt/fashion-mnist-data-training-using-pytorch-7f6ad71e96f4

ACKNOWLEDGEMENT

I would like to extend my gratitude towards Intel Corporation for giving me the opportunity. I would also like to thank Mr Mohan Nikam for his guidance and the college mentor Prof. Nisha P Shetty for their valuable time and support.

LINK TO SOLUTION

The link for the GitHub repository for the solution is given below.

https://github.com/akshatgarg333/Team1010_ManipalInstituteofTechnologyMAHE_FashionMNIST