Conquering Fashion MNIST with CNNs using Computer Vision

Team-Celestia

Kashik Ranjan Jha [MIT Manipal]

SUBMISSION DATE:15/07/23

ABSTRACT

The Fashion-MNIST clothing classification problem is a new standard dataset used in computer vision and deep learning The Dataset consists of a training set of 60,000 examples and a test set of 10,000 examples. Each example is a 28x28 grayscale image, associated with a label from 10 classes. CNNs are often preferred for this dataset due to their ability to capture spatial relationships and extract relevant features from images.

INTRODUCTION

Fashion-MNIST to serve as a direct drop-in replacement for the original MNIST dataset for benchmarking machine learning algorithms.It was designed to provide a more challenging task with greater complexity, requiring models to learn and differentiate between visually similar clothing items.The dataset’s easy availability and relevance makes it very suitable to use in real world applications.

NEED FOR THE PROBLEM

As technology has progressed, more powerful models have overshadowed previous datasets like the original MNIST. However, the Fashion MNIST dataset has emerged as a benchmark for contemporary classification models. It enables analysis of fashion trends,product recognition, recommendation systems, and visual search. By focusing on fashion items, the dataset captures the complexities and variations present in real-world fashion images, making it a valuable resource for researchers and practitioners working in the field of computer vision and fashion industry applications.

Consequently, it is crucial to advance classification models and problem-solving approaches to effectively cater to the dynamic and evolving technological landscape.

APPROACH

I build a CNN model using PyTorch for achieveing high accuracy.The model contains 9 layers having multiple convolutional layers,Activation,Max pool Layers,etc.Each of the layer and the values I have slowly checked and tweaked in order to get the best possible result.Our Neural Net has following layers:

- Convolutional Layer : This layer has 2 filters ranging from 16 to 32 which basically extract the meaningful feature from the input data
- 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 are used to activate different layers of the neural network
- Max Pooling : This layer helps in capturing important information and reducing computational complexity in subsequent layers.

RESULTS

Using PyTorch Convolutional Neural Network with Parameters :

- Learning Rate=0.001
- Epoch=12
- Adam optimizer
- CrossEntropyLoss
- Batch size=64

It yielded an accuracy of 91.55% and loss 0.0044 and Inference time of 1.006 s of on Test Set.
And on further optimization by Intel API for PyTorch(IPEX), accuracy increased to 91.61% with a loss of 0.0045 and Inference time of 0.914 s
Optimization by IPEX of model can lead decrease of upto 0.3 s in Inference Time of Model.

REFERENCES

Following are the references taken-:

- https://www.kaggle.com/datasets/zalando-research/fashionmnist
- https://towardsdatascience.com/pytorch-layer-dimensions-what-sizes-should-they-be-and-why-4265a41e01fd
- https://pytorch.org/tutorials/

ACKNOWLEDGEMENT

I would like to express my sincere appreciation to Intel Corporation for providing me with this incredible opportunity. I am also grateful to Mr. Mohan Nikam for his invaluable guidance and the mentor from my college, Prof. JayaShree, for their unwavering support and valuable time.

LINK TO SOLUTION

The link for the GitHub repository for the solution is given below:
https://github.com/KashikRJ/TEAMCelestia_MIT_FashionMNIST.git