Image Classification Challenge

Diem Le
ngocdiem.le@studenti.unipd.it

1. Introduction

Image classification is the most important part of digital image analysis. The objective is to use machine learning and deep learning algorithms to extract meaning from an image. Under supervised learning, image classification trains a machine learning model what each category "looks like" and then verifies the performance of the model by predicting a test dataset against its target. There are variety of machine learning techniques that can be applied to tackle image classification problem such as Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Decision Tree and Neural Networks (NN) or even Convolutional Neural Networks (CNN) - a class of deep learning neural networks.

This project will approach several machine learning algorithms to classify images of the Fashion-MNIST dataset of Zalando. The intention is to evaluate the performance of those methods on the same dataset and apply some regularization techniques to deal with over-fitting problem. Each method has its own advantages and disadvantages, in this dataset, SVM model gives the best result with 90% of accuracy among the basic techniques like KNN, Decision Tree, SGDClassifier and Naive Bayes Classifier. Meanwhile, CNN model with dropout regularization delivered better performance when addressing the over-fitting problem with 92.45% of accuracy on the validation dataset.

Throughout this project, Keras library is used for creating and training our Neural Network model because it is powerful and useful for beginners due to its simplicity. Beside Keras, Scikit-learn library is implemented for some basic supervised learning algorithms. Moreover, Matplotlib and Seaborn libraries are used for visualizing our results.

2. Dataset

This project used the dataset Fashion-MNIST from e-commerce company named Zalando. This dataset is meant to be a drop-in replacement for the MNIST dataset because of its challenges for evaluating machine learning algorithms. The Fashion-MNIST dataset consists of a training set of 50,000 images, a validation set of 10,000 images and a test set of 10,000 images. Each example is a 28x28 grayscale image, associated with a label from 10 classes. The labels of training and validation dataset of Fashion-MNIST are as follow: (Table 1)

T-shirt/Top: 0	Sandal: 5
Trouser: 1	Shirt: 6
Pullover: 2	Sneaker: 7
Dress: 3	Bag: 8
Coat: 4	Ankle boot: 9

Table 1. Labels

Data normalization is an important step which ensures that each input parameter has the same distribution. Moreover, this will makes algorithms converge faster while training a model. There are many ways to normalize data. For image input, the pixel values will be in the range between 0 and 255 with black and white respectively, they are rescaled to the range between 0 and 1 by dividing each input by 255.

In addition, many machine learning algorithms may not work with categorical or numerical data directly. Especially, when modeling multi-class classification problems using Neural Networks, it is better to represent each numerical output with a binary vector for each possible category and assigns a value of 1 to the feature of each sample that corresponds to its original category, this technique is called One-Hot encoding. Moreover, working with CNN model requires a 4D array as input. Therefore, reshaping its input to the original 28x28 grayscale image is mandatory.

3. Methodology

Neural networks have the ability to learn and model nonlinear and complex relationships. One of the most common issues occur while training neural networks is overfitting. Overfit model is the model which performs well on the training dataset but it is not as good at classifying data that it was not trained on. In this project, L2 regularization and dropout techniques are used to face with this problem.

Several neural network models from a basic 1-layer neural network to more complicated convolutional neural network are conducted to evaluate their performances on Fashion-MNIST dataset. Specifically, three main architectures are implemented included 1-layer neural network, 3-

layer neural network and simple convolutional neural network.

- 1-layer neural network: This model contains one hidden layer with 256 nodes and uses ReLu activation function. For the optimizer, Adam and Stochastic Gradient Descent (SGD) are considered to measure the convergence rate.
- 3-layer neural network: In addition to 1-layer neural network, two more hidden layers are added with 256 nodes respectively with ReLu activation function.
- For convolutional neural network, we used Sequential model with the following structure:
  - 3 layers of 2D convolutional layer (Conv2D) with kernel_size (3x3), ReLu activation function and the number of filters are 32, 64 and 128 respectively; 3 MaxPooling2D layers size (2x2);
  - Dense layer (a regular fully-connected NN layer) with 128 units and ReLu activation function;
  - Dense layer (final fully-connected NN layer) with 10 units (number of classes) and SoftMax activation function.

Different choices of optimization algorithm can bring the different performance for the model. In this case, Stochastic gradient descent and Adaptive moment estimation (Adam) are chosen to evaluate the effectiveness on the dataset. As expected, Adam performs the best on average because it converges faster. After 80 epochs, we can clearly see that our model does not generalize well from our training data to unseen data. It means that the problem we have to encountered is the over-fitting. (Figure 1). Therefore, Adam optimizer is chosen to be the optimizer parameter of neural network model in this project.

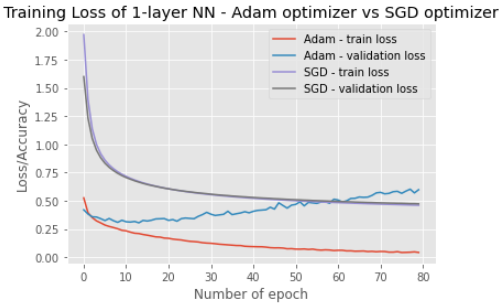

Figure 1. Training loss of 1-layer neural network

Dropout is a regularization technique for preventing over-fitting. Basically, dropout is a method where randomly selected neurons are dropped during training. Figure 3 shows that the validation loss has been significantly

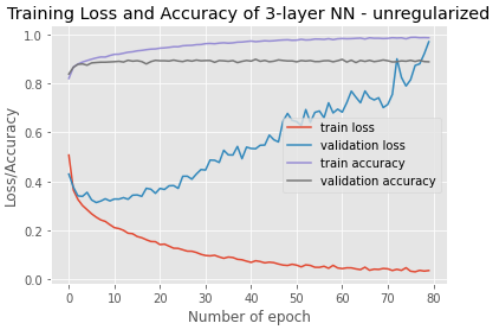

Figure 2. Unregularized 3-layer neural network

decreased compare to the loss without regularization of 3-layer neural network (Figure 2). Moreover, in CNN model,

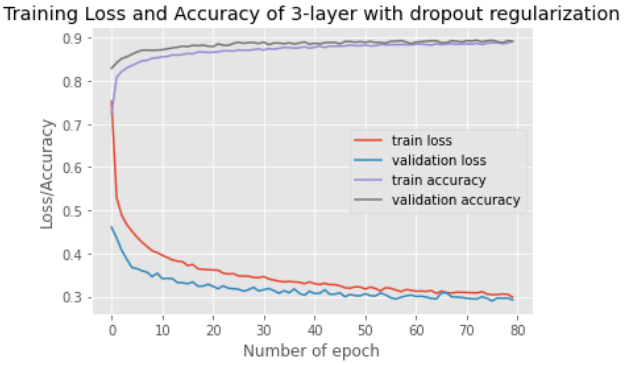

Figure 3. Regularized 3-layer neural network with dropout

dropout is also used to fight against over-fitting. With the drop probability of 0.2 for the first layer and 0.5 for other layers, the outcome is the same as in the regular model, which shows the effectiveness of dropout regularization to overcome over-fitting.

Another method for regularization is used in this project is L2 regularization, which can be implemented by augmenting the error function with the squared magnitude of all weights in the neural network. The value of weight decay while regularized the model is 0.001.

4. Experiments

4.1. Machine learning techniques

Some machine learning techniques are used to tackle the classification task. Table 2 shows the accuracy score of each method.

Based on table 2, SVM gives the best result with the highest accuracy score (0.8858). Therefore, we choose SVM as the baseline and the aim is to improve the performance of SVM by using optimizing hyperparameter method called Grid Search. From sklearn's GridSearchCV, we define our grid of parameters include the number of C and the

Method	Accuracy
K-Nearest Neighbors	0.8458
Decision tree	0.7871
SGD Classifier	0.8414
Naive Bayes classification	0.5826
SVM	**0.8858**

Table 2. Accuracy score of machine learning methods

type of kernel. After tunning, the result shows a slight improvement with the accuracy score when C = 10 and kernel = 'rbf' is 0.90.

4.2. Neural network methods

In the neural network methods, as discussed above, 3 models and some regularization techniques are combined to evaluate the effectiveness on facing with over-fitting. After training 80 epochs, CNN model with dropout regularization gives the best performance among 6 models. Table 3 below shows the accuracy score of all training models.

Method	Accuracy
1-layer neural network (NN)	0.8906
3-layer unregularized NN	0.8886
Regularized 3-layer NN with dropout	0.8916
Regularized 3-layer NN with L2	0.8856
Unregularized CNN	0.9181
Regularized CNN with dropout	**0.9245**

Table 3. Accuracy score of neural network methods

The figure 4 and 5 show the accuracy and loss of all neu ral network models.

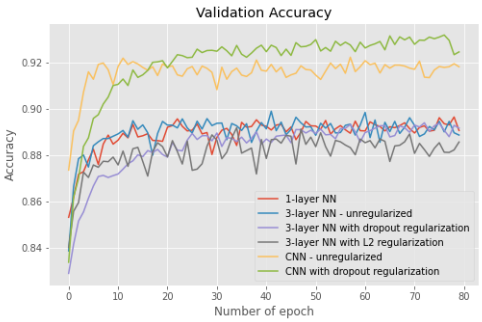

Figure 4. The accuracy between models

It can be seen that regularized CNN model with dropout regularization has the highest accuracy score (0.9245) and the lowest loss with the validation dataset (Figure 5). Moreover, the accuracy of 3-layer unregularized neural network is smaller than the score of 1-layer model, and from the figure 5, the loss of 3-layer model is bigger than 1-layer as well. Sometimes, a deep neural network is not always better

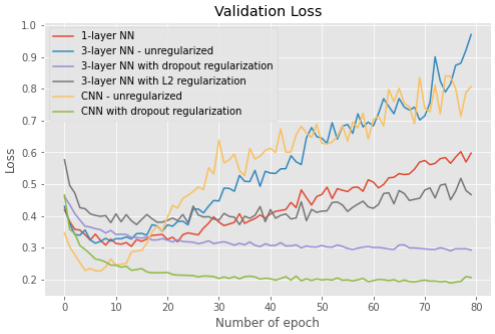

Figure 5. The loss between models

than a shallow one, it depends on the number of parameters and the size of training data, because limited training data is more prone to cause over-fitting with complicated model. In addition, regularized model with dropout shows higher accuracy score and lower loss than L2 regularization, which prove that dropout is more effective for complex networks. Despite the slight improvement of the accuracy score between unregularized CNN and regularized CNN, the loss is significantly reduced, which can be seen on figure 5.

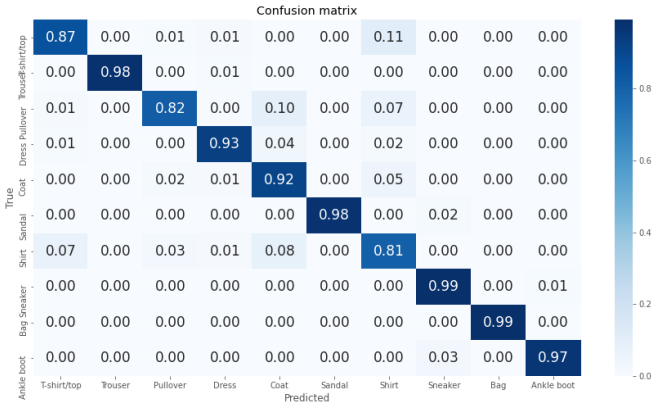

Figure 6. Confusion matrix of regularized CNN with dropout

The figure 6 is the confusion matrix of regularized CNN model with dropout regularization. While shirt and pullover have the worst performance (below 0.85), sneaker and bag are classified correctly with approximation 100%.

When evaluating the accuracy on the test set, although the model classifies quite accurately the images from test images dataset but the accuracy score is very low (0.1022). Because of the randomness labeling images makes the distribution of data in the test set is very different from the one of the training/validation set.