INFO 6146 Tensor Flow & Keras (Group 3)

Problem Statement:

The Fashion Design dataset (Fashion Design - MNIST) is akin to MNIST but encompasses images of specific clothing items and accessories.

This project aims to categorize these images into distinct classes using three different approaches:

1- a single-layer Artificial Neural Network
2- double-layered Artificial Neural Network
3- a traditional supervised machine learning algorithm.
4-

Dataset Description:

Fashion-MNIST is a dataset of Zalando's article images:

- Total images: 70,000
- Training set: 60,000
- Test set: 10,000
- Image size (grayscale): 28x28 pixels
- Pixel intensity range: 0-255 **(normalize)**
- Classes (10): T-shirt/top, Trouser, Pullover, Dress, Coat, Sandal, Shirt, Sneaker, Bag, Ankle-boot

Tasks to be Completed:

Create a document to answer these questions – 0 – No attempt; 1 – Attempt, 2 – with 50% correctness, 3 - Completed

Marks: 30

1. How does the accuracy of classification compare among the single-layer Artificial Neural Network, double-layers Artificial Neural Network, and the traditional supervised machine learning algorithm for the MNIST fashion dataset?

For single-layer ANN the accuracy was around 80-85%. It’s simple, so it can’t learn complicated patterns well.

The double-layer ANN performs better, achieving around 90-95% accuracy because it has an extra layer that helps it understand the data better. we noticed a difference of about 5% when comparing single layer and double layer ANN's

For traditional Supervised ML methods, different algorithms like SVM and Random Forest can vary in accuracy. Some can be just as good as the double-layer ANN, while simpler methods might not do as well. Random forest (100 trees) showed 88%accuracy which is reasonable for the dataset. However, the nature of the calculations makes the model linear. SVM and Random Forest were much better than Linear Regression which was the worst of all demonstrating that every model isn't a good fit for every problem. Linear regression shows the lowest accuracy at 39% which is due to the multi dimensional nature of the data that cannot be fitted with straight lines.

2. What are the differences in terms of training time between the three approaches when classifying the MNIST fashion images?

For a single-layer ANN, it trains quickly, usually taking seconds to a couple of minutes.

For double-layer ANN, it takes a bit longer (usually a few more minutes) because it has more parameters to adjust. For this dataset it took around 2 mins and 6 seconds

As for traditional ML methods they varied a lot. SVM can be slow, especially with complex settings, while Random forest (100 trees) training time is fast (1min and 58 seconds) . However, with 5 trees it can achieve good accuracy in 7 seconds.

3. How do the three methods perform in terms of generalization to unseen data? Do any of them show signs of overfitting or underfitting?

Single-layer ANN often doesn't do well with new data because it's too simple. It may underfit, as indicated by lower accuracy and limited class-specific performance.

Double-layer ANN performs generally better at handling new data but can sometimes memorize the training data too well causing overfitting. In this task, the double layer was able to generalize better because of its ability to capture more complex features as compared to a single layer ANN.

The traditional ML methods varied based on the method; Random Forests handled new data well for example. It showed 88% accuracy on unseen data which is reasonable for a traditional ML method. The SVM demonstrates good generalization, achieving comparable accuracy to the ANN models on unseen data.

4. Can you analyze and compare the learning curves of the three approaches? How does the loss and accuracy change over epochs for each method?

Single-layer ANN showed good performance on training data but poorly on validation data. Learning curves show that the single-layer ANN plateaus early with limited learning capacity.

Double-layer ANN generally showed improvement over time but can start to perform worse on validation data when it overfits. The training loss steadily decreases showing the model is getting better with each iteration, the validation loss however plateaus after some iterations possibly suggesting some overfitting.

Traditional ML methods performance trends depend on the specific method. For example, random forest does not have learning curve as it follows entropy and information gain calculations at each step.

5. Are there notable differences in the model's performance when dealing with certain classes of fashion items? Do specific approaches excel in classifying certain types of items?

Some models, like double-layer ANNs, are good at recognizing certain items (like shoes), but struggle with similar items (like T-shirts and pullovers). Random Forest shows lower performance on the shirt class. However, every model does struggle with the shirt class (lowest true positives for all methods), specially SVM and linear regression

6. How sensitive are the three methods to hyperparameter tuning? Are there certain parameters that drastically affect their performance on the MNIST fashion dataset?

All three methods can benefit from tuning. The single-layer ANN are impacted by learning rate and batch size matter. The double-layer ANN are impacted by learning rate, number of hidden neurons, activation function and dropout rate.

For traditional ML methods such as Random Forest model; changing the number of trees from 2 to 5 improves the accuracy from 76% to 83% but when increasing the trees to 100 the accuracy increases to 88% and so there is a less increase in accuracy while the training time increases from 7 seconds for 5 trees to 1min and 58 seconds for 100 trees. For SVM; kernel type, regularization and C can impact its performance but may also increase training time significantly.

7. What are the limitations of each approach in terms of handling the complexity and nuances of fashion images compared to more advanced techniques?

The ANN's work fine on the fashion MNIST since its a grayscale image dataset but it would struggle if any colour channels were added to it since ANN's doesn't maintain the integrity of the image structure and flattens the whole image into a linear array, comparatively CNN maintains the integrity of the image structure and performs better on multi channel images and structures of that sort.

Similarly, traditional ML methods like Random forest consider multiple combination of trees that are split on different features (pixels), however it would not provide the same level of complexity as CNN for example as it does not consider spatial relations

between pixels. On the other hand, SVM is effective but computationally intensive and may not scale well to more complex, higher-dimensional datasets.

8. How do the single-layer and double-layered Artificial Neural Networks compare in terms of the depth of features they can extract from the images?

Single-layer ANN only learns simple patterns and shallow features, which limits its accuracy

Double-layer ANN can learn more complex, abstract features enabling it to recognize subtle distinctions between similar classes like Shirts and Pullovers. Consecutively every layer you add will learn more and more depth features

Traditional ML methods vary; some can capture non-linear patterns, but it depends on the specific algorithm.

9. Can you analyze the decision boundaries created by each approach? Do they handle complex decision boundaries in the dataset differently?

Single-layer ANN creates simple, straight-line boundaries that often don’t work well in complex problems.

Double-layer ANN can create more complex and non linear boundaries which helps with difficult classifications.

Traditional ML methods like the Random forest model; uses information gain and entropy calculations as the decision boundaries. Linear regression and single-layer ANN create simpler more linear boundaries leading to possible misclassification. While SVM also provides strong linear separations it may struggle with nonlinear boundaries without more advanced kernel functions.

10. What insights can you draw about the trade-offs between computational complexity, model performance, and ease of implementation when comparing these three approaches?

Single-layer ANN is easy to set up and fast, offers simplicity and fast training but sacrifices accuracy.

Double-layer ANN is a good balance but requires more setup It delivers better results and balances complexity and performance well, providing higher accuracy with moderate computational demands

Traditional ML methods vary a lot. Simpler methods are easier to use and understand but might miss complex details. For this specific dataset Random forest provides a reasonable accuracy with very short training time. However, the SVM, while accurate, it is computationally intensive, making it less practical for larger datasets despite ease of implementation for linear separations

Marks 20

Create a detailed Python notebook for each method (a total of 3 files) and share your implemented Python codes with comments in each step of the procedure.