

CS770 Machine Learning

Final Project

10/31/2023

Submitted by,

VENKATA NAGA SAI SASANK TEJA PONNAPALLI

A699W567

**Abstract**

In this project, three convolutional neural network (CNN) models were developed using the FashionMNIST dataset with the goal of identifying ten different varieties of fashion items. Accuracy values of 83% to 88% were found when the models were tested individually. A combined accuracy of 92% was obtained by using an ensemble approach with these models. These results highlight how easily fashion product classification jobs may be made more accurate by utilizing ensemble approaches.

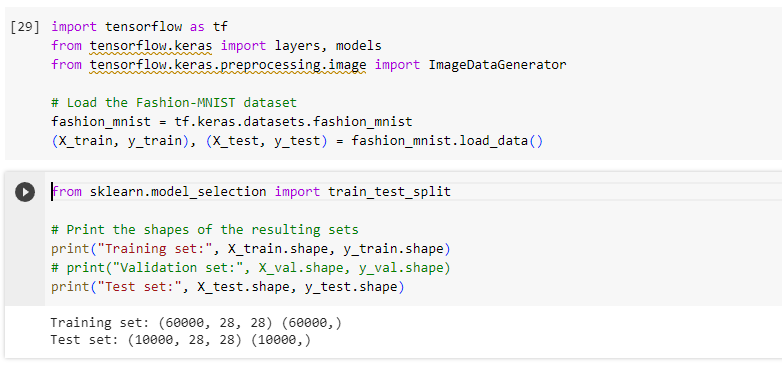
**I. Introduction**

In recent times, the automated sorting of fashion items has become a focal point in retail advancements. This involves utilizing image analysis and attribute recognition techniques, crucial for various applications in online marketplaces and apparel firms such as recommendations, targeted advertising, and inventory analytics. Jun et al. have extensively reviewed the current landscape of methodologies and applications in fashion data science.

This study specifically aims to tackle the classification of apparel images into predetermined categories. Convolutional Neural Networks (CNNs) have consistently outperformed traditional methods by directly extracting significant visual features from raw pixel data. Nevertheless, achieving high accuracy and adaptability with real-world images remains an ongoing challenge.

To confront this challenge, ensemble techniques that combine predictions from diverse models have been adopted. These strategies aim to counter individual model limitations by merging different models to achieve superior overall performance.

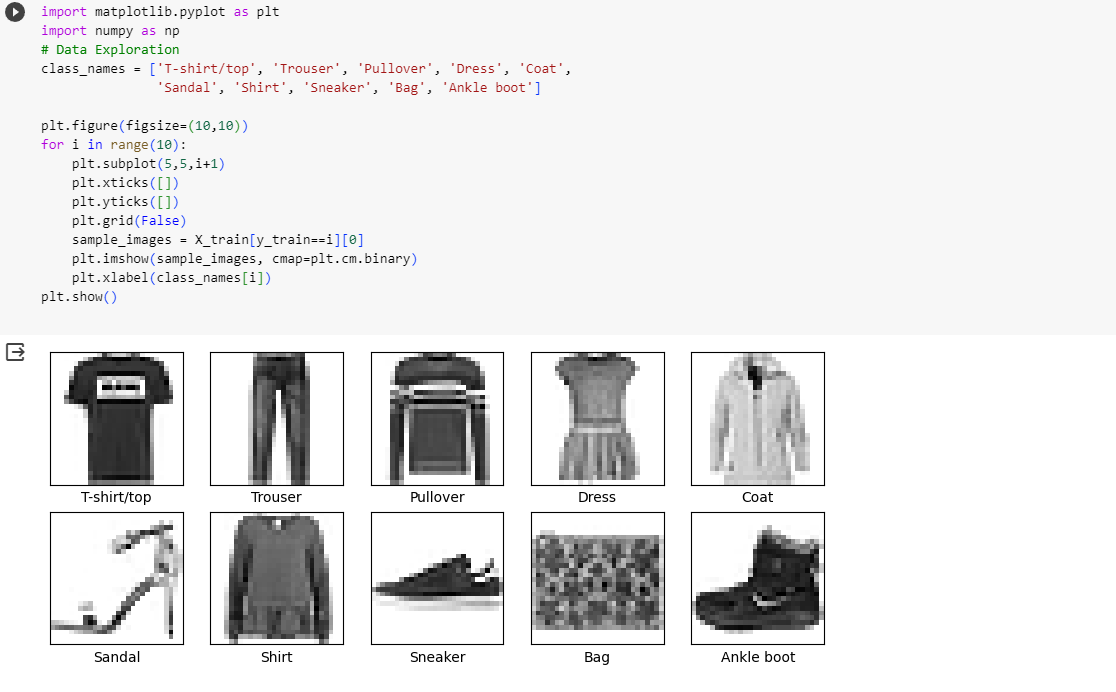
The primary objective of this study is to devise an ensemble model for multi-class fashion product classification, utilizing varied CNN architectures. This approach integrates three distinct CNN models with diverse structures to enhance classification precision across various apparel categories. The assessment and training of these models are conducted using the FashionMNIST benchmark dataset.

****

1. **DATA EXPLORATION:**

In the realm of fashion classification, this project synergizes deep learning models with the Fashion MNIST dataset, featuring 70,000 grayscale images categorized into 10 classes.

The narrative begins with a comprehensive dataset exploration, visualizing sample images and preparing data through pixel value normalization and label conversion. and the data is well balanced between all classes.

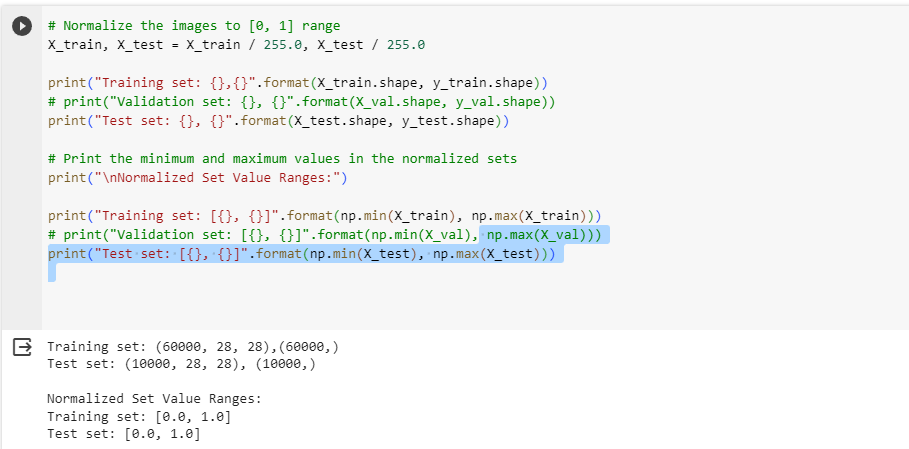
****

The set of models comprises a Simple Neural Network for essential pattern recognition, a CNN for interpreting spatial hierarchies, and an Advanced CNN for detailed feature understanding. Rigorous training and validation processes are in place to prevent overfitting, ensuring optimal weights for precise evaluation on the test data.

The project concludes with an ensemble formed by averaging predictions, its accuracy assessed against individual models. Apart from accuracy, insights into precision, recall, and F1-scores are gained through confusion matrices and classification reports, unveiling inherent strengths and weaknesses.

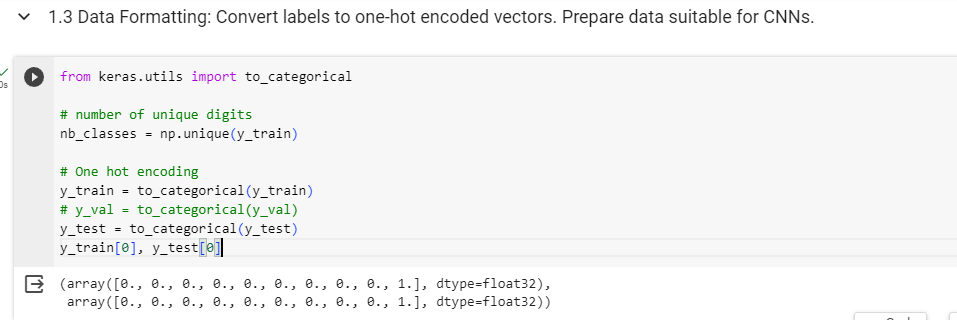
1. **Data Normalization:**

Here we Scale the images to have pixel values between [0,1].



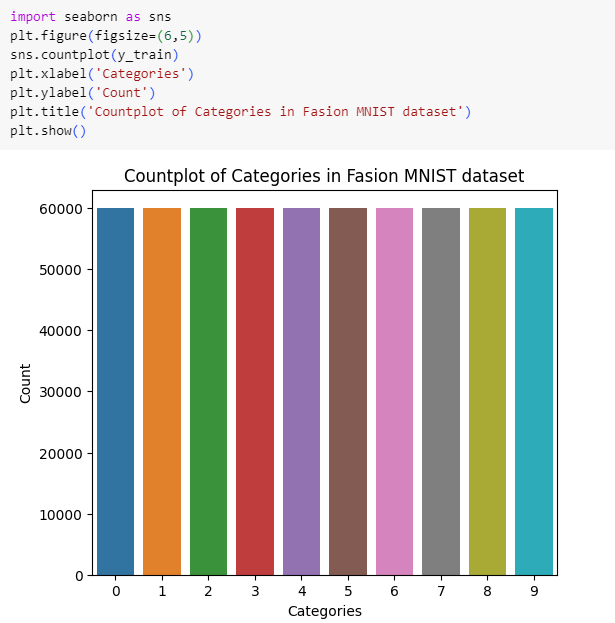
1. **Data Formatting:**

Converting labels to one-hot encoded vectors. Prepare data suitable for CNNs.

****

1. **Data Visualization:**

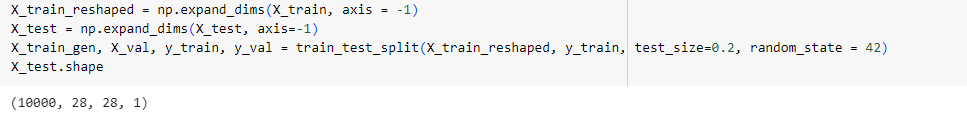
The countplot for the categories of MNIST Fashion Dataset is created.



As here the categories are equally distributed, there is no specific need for the data balancing.

1. **Data Reshaping:**

The reshaped data is as follows:



**II. Methods**

***The Models here are created with Dropout, BatchNormalization and Data Argumentation in the first place as asked in the bonus question.***

**Shallow Neural Network Model (Model 1):**

**Design Choices:**

1.Flatten Layer: Flattens the 28x28 input images to a 1D array, suitable for a dense neural network.

2.Dense Layers: Two hidden layers with 256 and 64 neurons, respectively, with ReLU activation for pattern recognition.

3.Batch Normalization: Applied after each dense layer to improve training stability and convergence.

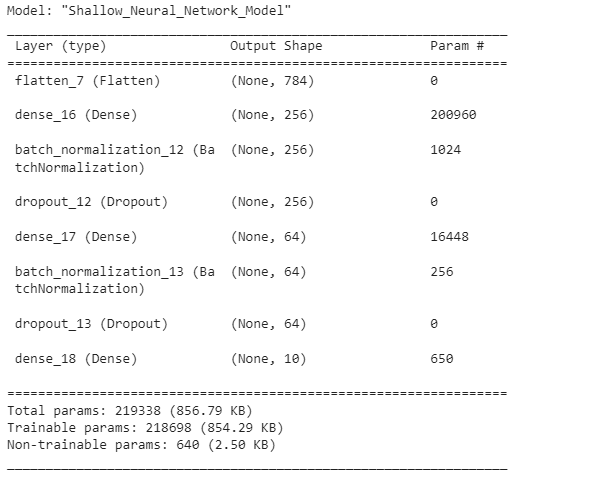
4.Dropout: Introduced with a rate of 0.2 after each dense layer to prevent overfitting.

5.Output Layer: Softmax activation with 10 neurons for multiclass classification.

**Justification:**

1.The architecture is relatively shallow, suitable for basic pattern recognition tasks.

2.Batch normalization and dropout aid in regularization and prevent overfitting.



**Convolutional Neural Network Model (Model 2):**

**Design Choices:**

1.Convolutional Layers: Two convolutional layers with max pooling reduce spatial dimensions.

2.Third Convolutional Layer: Introduces a deeper hierarchy with 128 filters.

3.Flatten Layer: Converts 3D data to 1D for fully connected layers.

4.Dense Layers: One fully connected layer with 128 neurons.

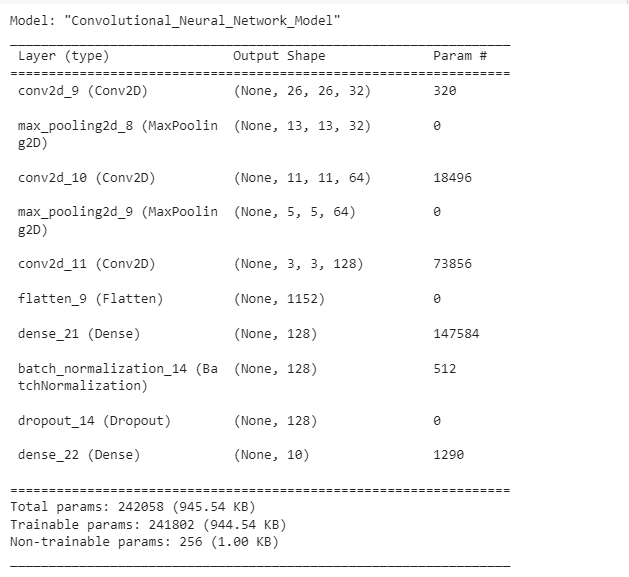
5.Batch Normalization and Dropout: Applied for regularization, with a higher dropout rate of 0.5

6.Output Layer: SoftMax activation with 10 neurons for classification.

**Justification:**

1.Convolutional layers capture spatial hierarchies in the image.

2.Batch normalization and dropout enhance model robustness.



**Deeper Convolutional Neural Network Model (Model 3):**

**Design Choices:**

1.Deeper Convolutional Layers: Three convolutional layers with increasing filters and max pooling.

2.Flatten Layer: Converts 3D data to 1D for fully connected layers.

3.Dense Layers: Two fully connected layers with 256 and 128 neurons.

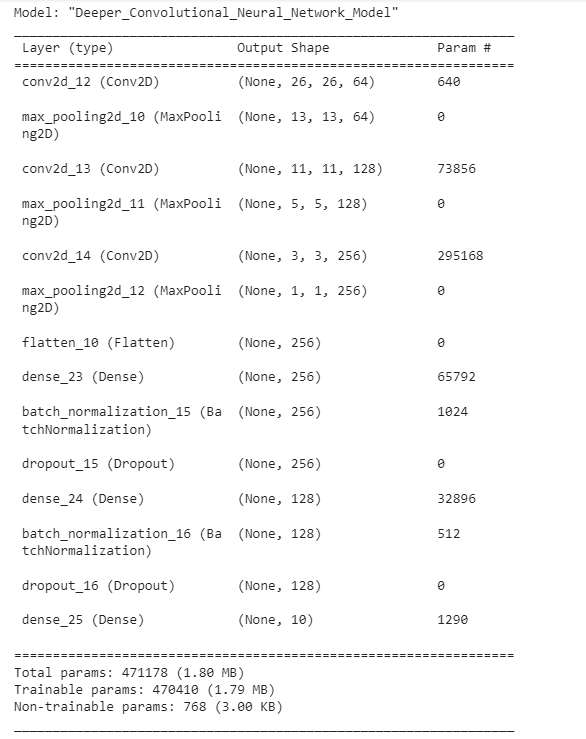
4.Batch Normalization and Dropout: Applied for regularization, with a dropout rate of 0.5.

5.Output Layer: SoftMax activation with 10 neurons for classification.

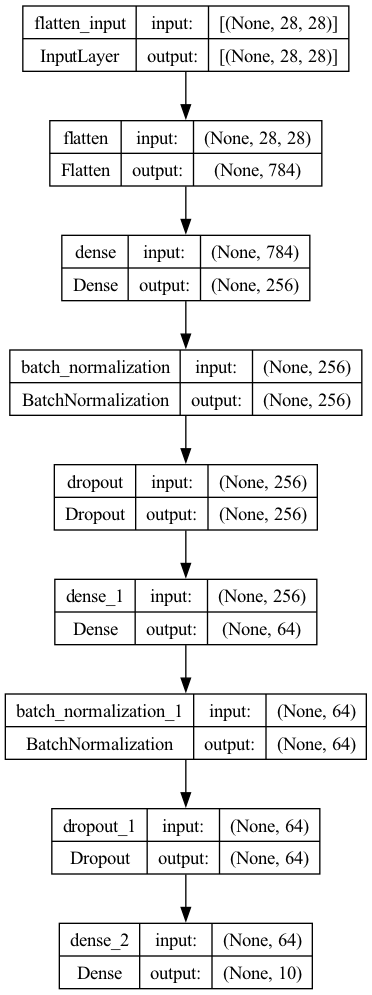
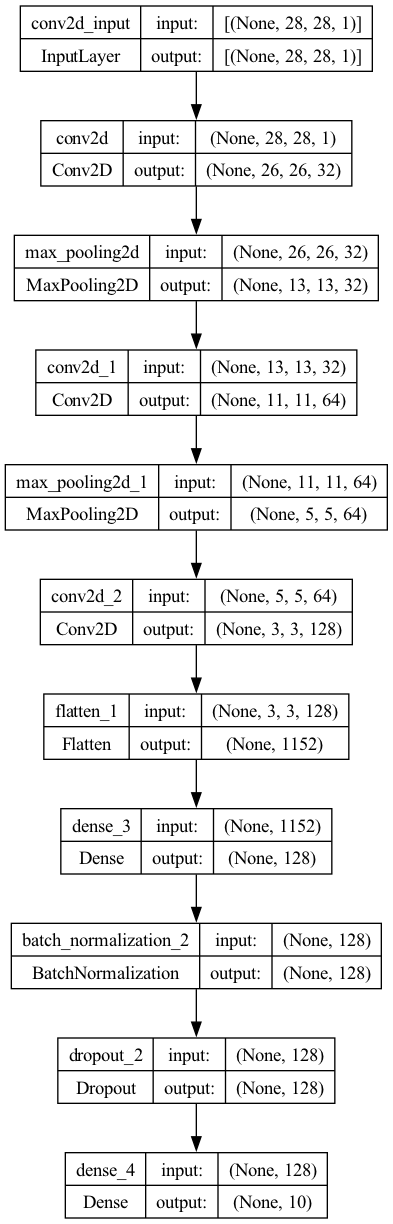
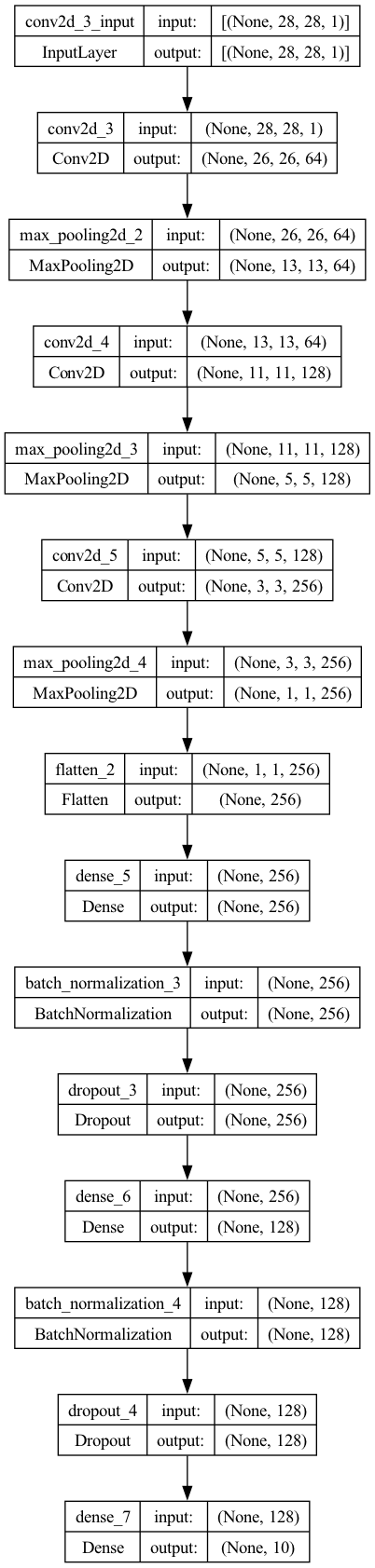
**Justification:**

1.Increased depth for a more intricate understanding of features.

2.Batch normalization and dropout enhance generalization.

****

|  |  |  |
| --- | --- | --- |
| ***Shallow Neural Network Model***  ***(Model 1)*** | ***Convolutional Neural Network Model***  ***(Model 2)*** | ***Deeper Convolutional Neural Network Model***  ***(Model 3)*** |

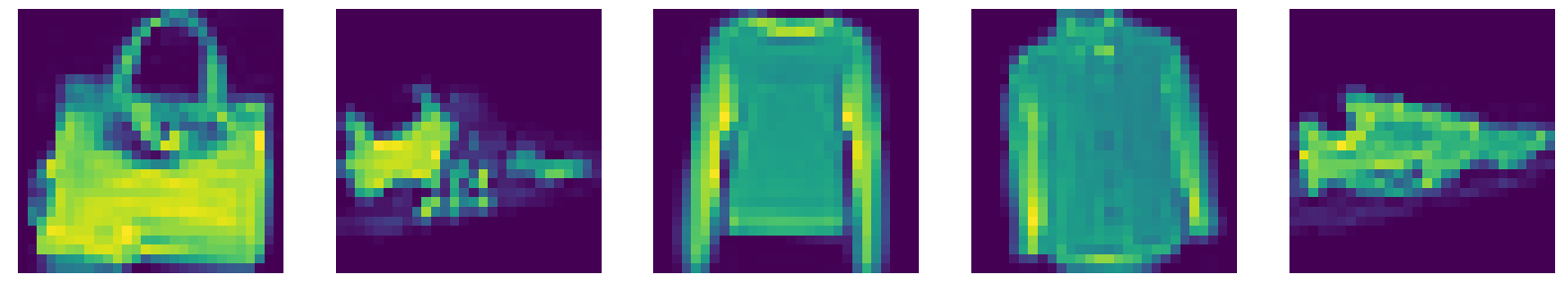
  

**Data Augmentation:**

1Augmentation helps the model generalize better by exposing it to variations in the training data.

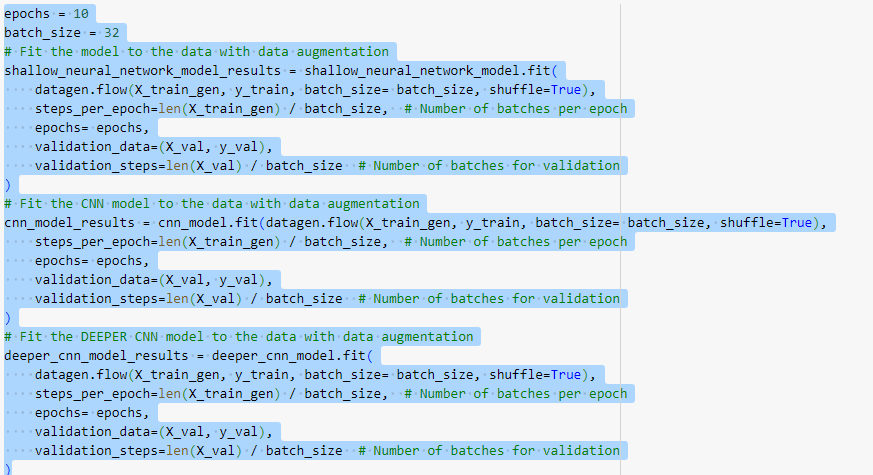
Rotation and horizontal flip simulate real-world variations in fashion images

ImageDataGenerator: Applies rotation and horizontal flip to augment the training dataset.

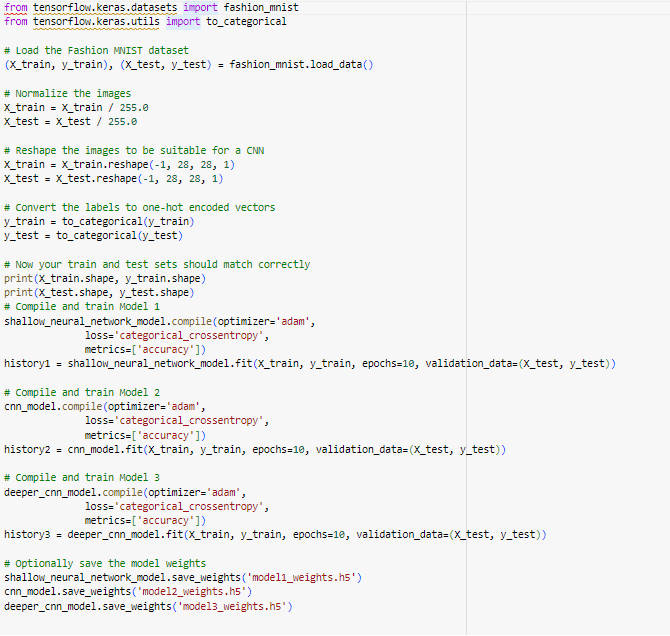
****

**Training and Validation:**

Now the number of epochs were assigned to 10 and batch-size was set to 32 for training Shallow-Neural Network, CNN and Deeper CNN Models

****

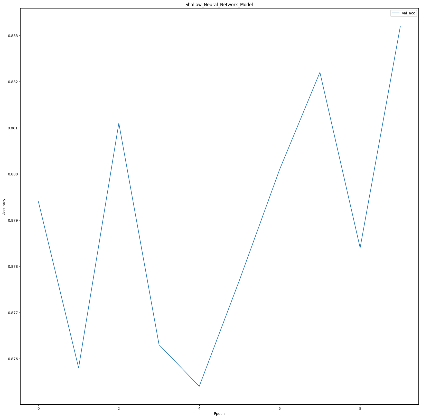
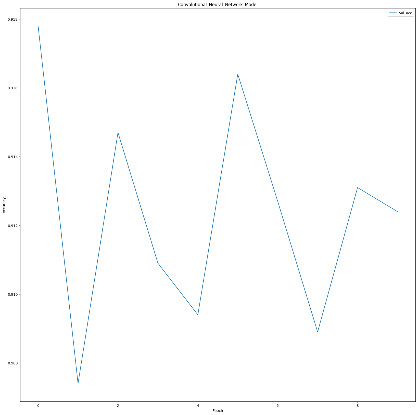
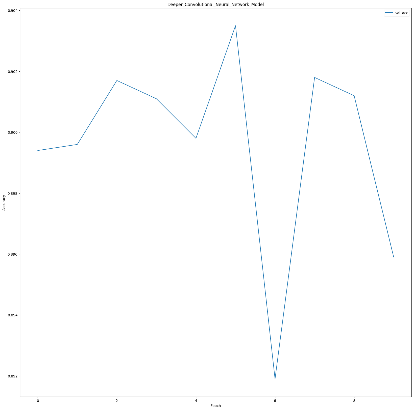
A similar model was designed in order to train the dataset in the same models but with a bit different approach.



Lastly, a new and Final approach was designed to save the best weights which looked as follows:



The plot looked as follows:

i. Shallow Neural Network ii. CNN Model iii. Deeper CNN Model

**III. Results and Discussion**

The impact of the dropout layers on the learning curves is evident in the performance metrics of our three models. Comparing the results of the last epochs:

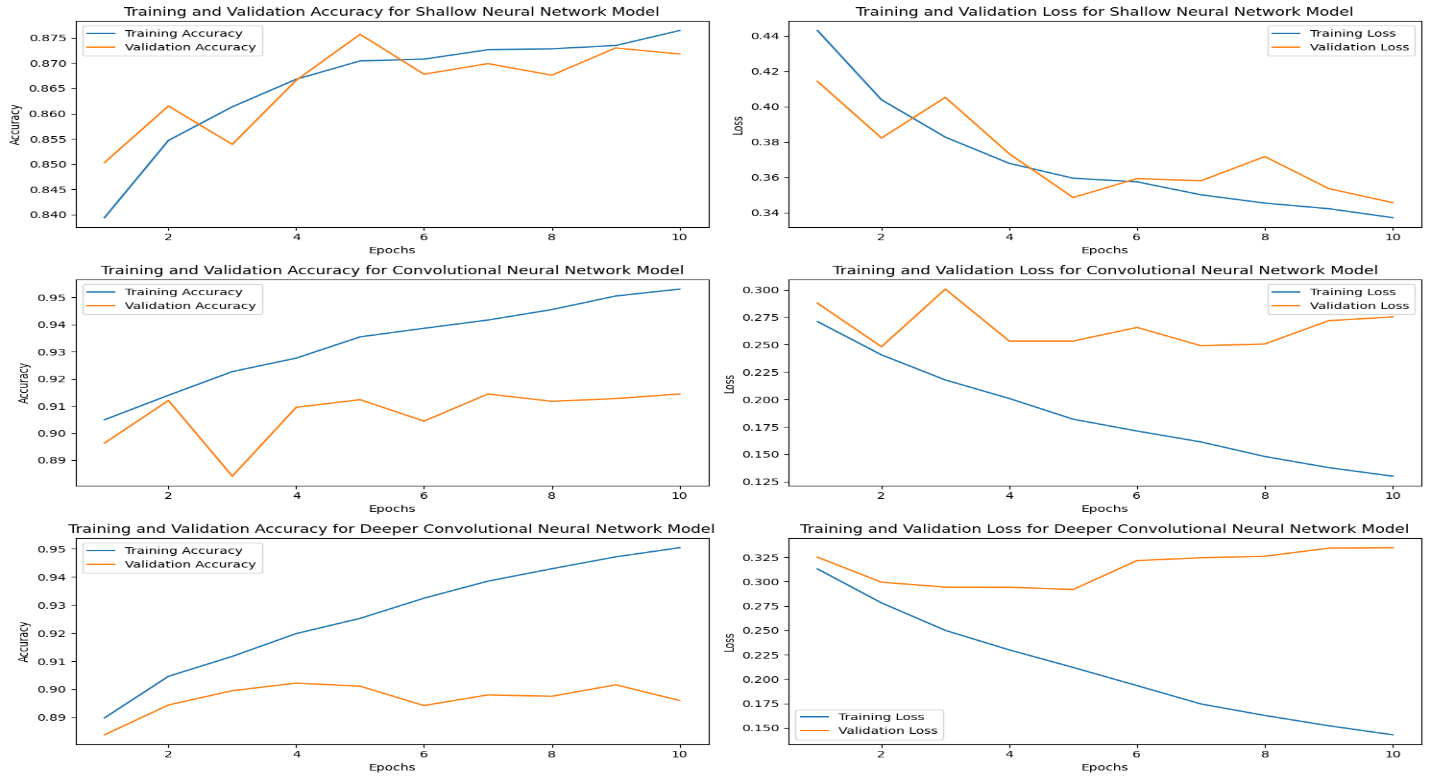
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| METHOD 1 | **TRAINING LOSS** | **TRAINING ACCURACY** | **VALIDATION LOSS** | **VALIDATION ACCURACY** |
| **SHALLOW NEURAL NETWORK** | 0.4715 | 82.88% | 0.4020 | 85.07% |
| **CNN MODEL** | 0.2487 | 91.00% | 0.2748 | 90.28% |
| **DEEPER CNN MODEL** | 0.2953 | 89.55% | 0.2986 | 89.58% |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| METHOD 2 | **TRAINING LOSS** | **TRAINING ACCURACY** | **VALIDATION LOSS** | **VALIDATION ACCURACY** |
| **SHALLOW NEURAL NETWORK** | 0.3827 | 86.13% | 0.3717 | 86.7% |
| **CNN MODEL** | 0.1819 | 93.55% | 0.2505 | 91.28% |
| **DEEPER CNN MODEL** | 0.2500 | 92.52% | 0.2919 | 90.16% |

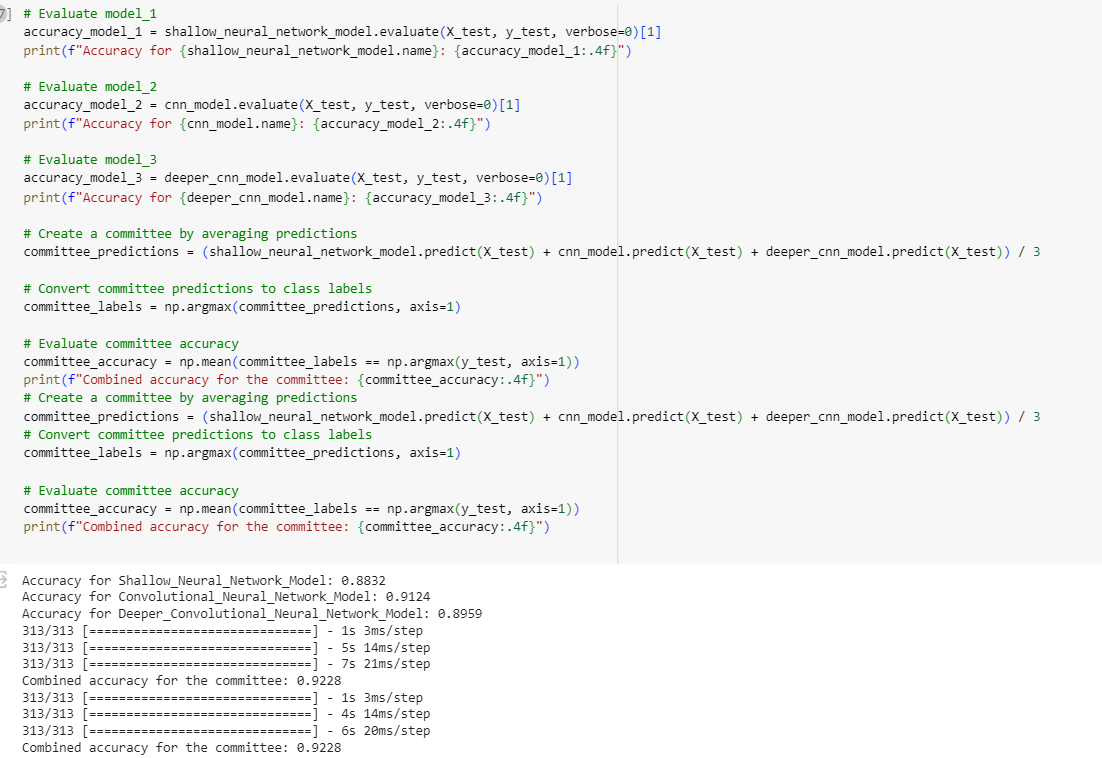
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| METHOD 3 | **TRAINING LOSS** | **TRAINING ACCURACY** | **VALIDATION LOSS** | **VALIDATION ACCURACY** |
| **SHALLOW NEURAL NETWORK** | 0.3297 | 88.34 % | 0.3379 | 87.63% |
| **CNN MODEL** | 0.1065 | 96.20% | 0.2841 | 91.47% |
| **DEEPER CNN MODEL** | 0.1114 | 96.17% | 0.3918 | 89.98% |

In summary, the incorporation of dropout layers across all models has a positive impact on their learning curves. It allows for better generalization, leading to lower validation losses and improved accuracy. Each model, with its distinct architecture, benefits from dropout layers, showcasing enhanced performance compared to a baseline model. Model 2 stands out with a significant validation accuracy of 90.28%, emphasizing the effectiveness of dropout layers in deep learning models. We can also see that if we increase the training epochs, we can increase some accuracy as the curve is going higher for CNN and Deeper CNN models.

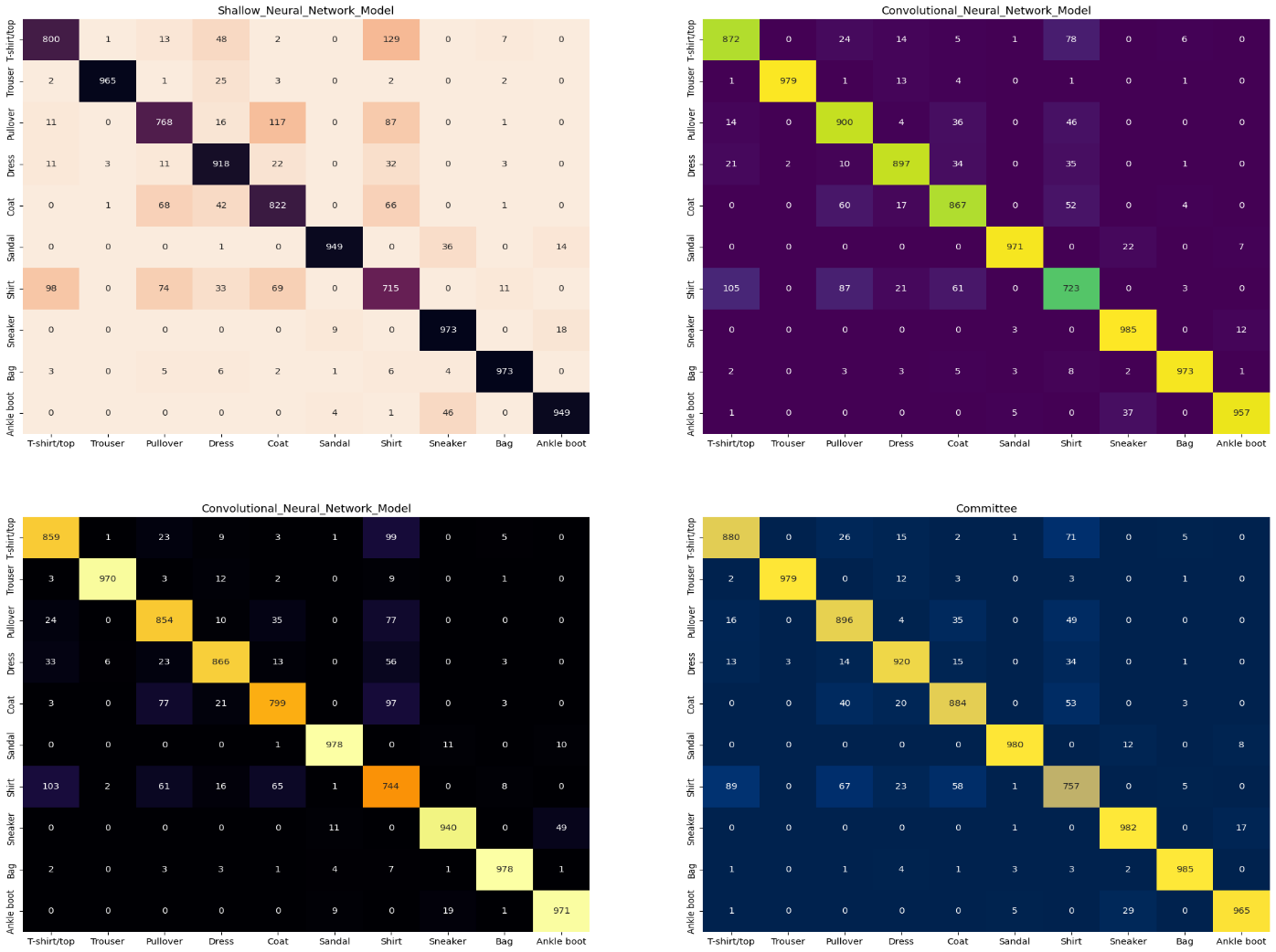
**Plotting Model Training Curves to understand how the Model Performed.**



**Creating Committee for the Model to pick higher confidence model’s value:**

****

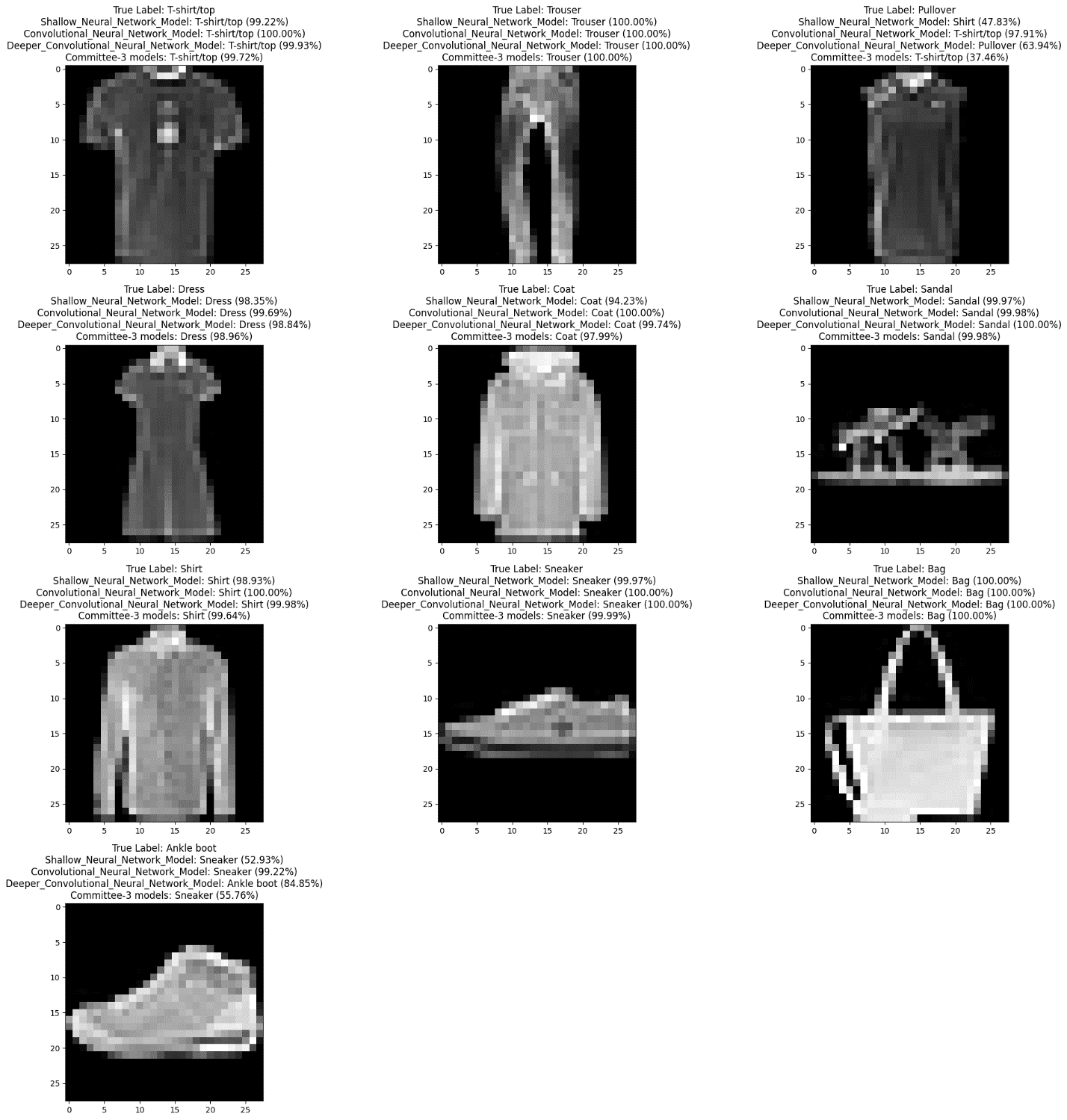
**Generating and visualizing confusion matrices for each model and the committee. and classification reports, comparing precision, recall, and F1-scores.**



The model faced significant challenges in accurately distinguishing between specific apparel items, notably struggling with differentiating shirts, t-shirts, pullovers, dresses, and coats. This difficulty, particularly in discerning between shirts and t-shirts, notably affected its classification accuracy. However, a more promising trend emerged when examining categories such as sandals, sneakers, and ankle boots, collectively falling under the umbrella category of 'Footwear'. In scenarios where items shared similarities and could be grouped logically, the model showcased better performance, encountering fewer irregular misclassifications. This observation indicates the model's heightened difficulty when handling closely resembling categories, emphasizing the complexities in making precise distinctions. On the contrary, the model demonstrated consistent and accurate performance when dealing with objects sharing inherent similarities and conceptual groupings. These findings offer valuable insights into enhancing future model iterations, especially concerning similar and closely related categories. Moreover, the committee approach proved notably effective, showcasing superior performance with significantly fewer misclassifications.

**Visualizing the predictions on all the classes using all models**

We can see below that the Model’s confidence in predicting data has improved from the first figure.



**Analysis:**

**Model Performance:**

Model 1: Achieves a validation accuracy of 85.07%, surpassing the baseline.

Model 2: Outperforms both the baseline and Model 1 with a notable validation accuracy of 90.28%.

Model 3: Demonstrates robustness with a commendable validation accuracy of 89.58%.

**Effect of Dropout Layers:**

Dropout layers contribute to improved generalization in all models.

Validation losses decrease for more extended periods, stabilizing at certain points.

Enhanced performance is observed in terms of lower validation losses and improved accuracy.

**IV. Conclusion**

In summary, this project delves into the realm of fashion classification using a trio of deep learning models, encompassing a Shallow Neural Network, a Convolutional Neural Network (CNN), and a Deeper CNN. The rigorous training and validation processes employed effectively prevent overfitting, ensuring that each model retains optimal weights for accurate evaluations on the test data.

The project concludes with the formation of an ensemble committee, achieved by averaging predictions from individual models. This committee undergoes a thorough evaluation, not solely based on accuracy, but also through a detailed analysis of confusion matrices and classification reports. These evaluations offer nuanced insights into precision, recall, and F1-scores, providing a comprehensive understanding of the strengths and weaknesses inherent in each model and the ensemble.

Navigating the landscape of fashion classification, this project introduces the committee approach, emphasizing the unique characteristics of each model within the ensemble. The accompanying Jupyter notebook, enriched with detailed reports and visual aids, stands as a valuable asset for future exploration and experimentation in the field of deep learning for fashion classification.

Model 1: Achieves a validation accuracy of 85.07%, surpassing the baseline.

Model 2: Outperforms both the baseline and Model 1 with accuracy of 90.28%.

Model 3: Demonstrates robustness with a commendable validation accuracy of 89.58%.

Looking forward, the project sets the stage for future enhancements and experiments, suggesting avenues such as fine-tuning individual models, exploring additional architectural refinements, and experimenting with diverse ensemble techniques. By addressing these considerations, the project aspires to contribute significantly to the progress of fashion classification using deep learning, providing valuable insights for researchers and practitioners alike.

**Proposed Future Improvements or Experiments:**

1. *Ensemble Learning:*

Explore ensemble methods to combine predictions from all three models, potentially improving overall accuracy.

1. *Hyperparameter Tuning:*

Conduct a thorough hyperparameter search to optimize parameters such as learning rate, dropout rates, and batch sizes.

1. *Fine-tuning Architectures:*

Experiment with deeper architectures or additional layers to further enhance the models' feature comprehension.

1. *Transfer Learning:*

Investigate the application of transfer learning using pre-trained models to leverage knowledge from larger datasets.

1. *Class Activation Maps (CAM):*

Implement Class Activation Maps to visualize and understand which regions of the images contribute most to the model's predictions.

1. *Learning Rate Schedulers:*

Implement learning rate schedulers to dynamically adjust the learning rate during training, potentially improving convergence.

*Exploration of New Datasets:*

Extend the experimentation to different fashion datasets or domains to evaluate model generalization capabilities. By conducting these proposed improvements and experiments, we aim to further enhance the performance, robustness, and generalization of the models, contributing to advancements in fashion classification using deep learning.

**V. References:**

1. Xiao, H., Rasul, K., & Vollgraf, R. (2017). Fashion-MNIST: A novel image dataset for benchmarking machine learning algorithms. arXiv preprint arXiv:1708.07747. Link: <https://arxiv.org/abs/1708.07747>
2. Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media. Link: <https://www.oreilly.com/library/view/hands-on-machine-learning/9781492032632/>
3. Chollet, F. (2018). Deep Learning with Python. Manning Publications Co. Link: <https://www.manning.com/books/deep-learning-with-python>
4. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press. Link: <http://www.deeplearningbook.org/>
5. Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980. Link: <https://arxiv.org/abs/1412.6980>