Fashion MNIST Classification

Abhijeet Ghadiyaram, Sri Harsha Parimi Venkata
abhigs@tamu.edu, parimi@tamu.edu

November 5, 2024

Project Description

The primary goal is to leverage and enhance Convolutional Neural Networks (CNN) and Mini ResNet architectures for robust classification of the Fashion MNIST dataset while focusing on improving model interpretability through advanced techniques. The Fashion MNIST dataset, consisting of 60,000 training images and 10,000 testing images, each of 28x28 pixels and categorized into ten classes (e.g., T-shirt/top, trouser, pullover, etc.), serves as a benchmark to evaluate the effectiveness of machine learning models in fashion image classification.

Data Set

The Fashion MNIST dataset is a well-balanced, grayscale dataset structured to facilitate the testing of image classification models. Each image belongs to one of ten fashion-related categories, providing a platform for comprehensive model assessment and comparison.

Pattern Recognition Techniques

To enhance baseline performance and provide comprehensive model interpretability, the project will employ the following techniques:

- **Data Cleaning and Normalization:** Initial preprocessing ensures image integrity making data suitable for neural network training.
- **Data Visualization and Exploration:** Methods like PCA and t-SNE will be used for dimensionality reduction to visualize class separations and data distribution patterns. This helps identify any overlap among classes and serves as an indicator of potential classification challenges.
- **Feature Extraction and Model Architecture:** The CNN model will be designed with convolutional layers to capture local image features, while the Mini ResNet model will include residual blocks to improve gradient flow and support deeper networks. The feature extraction approach will

enable capturing spatial relationships essential for distinguishing between fashion categories.

- **Model Interpretability Techniques:** To provide insights into the model's decision-making, SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations) will be used. SHAP highlights the key pixel regions influencing predictions, offering global model interpretability, whereas LIME provides a localized view of how specific image regions contribute to individual predictions.

- **Classification and Validation:** The CNN and Mini ResNet architectures will be trained to classify images, using accuracy, precision, and recall as primary performance metrics. Grad-CAM (Gradient-weighted Class Activation Mapping) will be used to further visualize the areas contributing most to each prediction, adding transparency to the classification process.

Project Goals

The project's refined objectives include:

1. Achieve optimal classification performance through the tuning of CNN and Mini ResNet architectures, including hyperparameter adjustments (e.g., filter sizes, activation functions, and number of residual blocks).

2. Enhance model interpretability using SHAP and LIME to gain insights into the decision-making process of the models, ensuring that key contributing factors to predictions are identified.

3. Utilize Grad-CAM to visualize class activation maps and support LIME analysis, improving trust and transparency in practical applications.

4. Validate model robustness by applying the fine-tuned CNN to out-of-sample images distinct from the test set, confirming the model's real-world efficacy

References

1. Nocentini, O.; Kim, J.; Bashir, M.Z.; Cavallo, F. Image Classification Using Multiple Convolutional Neural Networks on the Fashion-MNIST Dataset. *Sensors* 2022, 22, 9544.

2. Xhaferra, E., Cina, E., & Toti, L. "Classification of Standard FASHION MNIST Dataset Using Deep Learning Based CNN Algorithms," 2022 International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, 2022, pp. 494-498.