Deep Learning Project Report

Main Objective of the Analysis

The main objective of this project is to apply deep learning techniques to a dataset of choice in order to extract meaningful insights. This includes training multiple variations of models and evaluating them to determine the best-performing model. The analysis emphasizes both accuracy and interpretability, with a focus on deriving actionable insights for stakeholders.

Dataset Description

The dataset selected for this analysis is a publicly available dataset. It contains multiple features describing input data instances and an output label that serves as the prediction target. The dataset provides sufficient variety for testing classification models. Key attributes include numerical features, categorical features, and a class label for supervised learning tasks.

Data Exploration and Cleaning

The dataset was explored to understand missing values, feature distributions, and correlations. Missing values were handled using imputation, categorical features were encoded, and continuous variables were normalized. Feature engineering steps included deriving interaction variables and removing redundant attributes.

Model Training and Variations

Three variations of deep learning models were trained: 1. A simple feed-forward neural network with two hidden layers. 2. A deeper neural network with batch normalization and dropout. 3. A convolutional neural network (CNN), adapted for tabular/image-like data. Each model was trained with different hyperparameters including learning rate, epochs, and activation functions. Performance was compared using accuracy, precision, recall, and F1 score.

Final Model Recommendation

Among the models tested, the deeper neural network with dropout and batch normalization was chosen as the final model. It provided the best balance of accuracy and generalization, while also being more robust to overfitting than the simpler models.

Key Findings and Insights

The analysis revealed that feature importance varied significantly across models, with certain features contributing disproportionately to predictions. The best-performing model achieved high accuracy and consistent validation performance, demonstrating that the selected features were meaningful predictors. This suggests that stakeholders can rely on these predictors for decision-making processes.

Next Steps

Future work should focus on expanding the dataset, incorporating additional relevant features, and experimenting with alternative architectures such as recurrent neural networks (RNNs) or transformer models. Further hyperparameter tuning and explainability methods like SHAP or LIME can help refine the model and build stakeholder trust. Additionally, real-world deployment considerations such as scalability and fairness must be addressed in subsequent iterations.