# Results

## Overview

This document provides an interpretation and summary of the results from various machine learning models used for the MNIST handwritten digit classification task. The models tested include Logistic Regression, RidgeClassifier, and Random Forest. The performance metrics, cross-validation results, and confusion matrices are analyzed to determine the best-performing model for deployment.

## Results Summary

The table below summarizes the performance of the models based on accuracy, F1-Score (Test), cross-validation mean F1-Score, and standard deviation. Random Forest emerged as the best-performing model with the highest accuracy and F1-Score, indicating its robustness for this task.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Experiment | Run Name | Accuracy | F1-Score (Test) | CV Mean F1-Score | CV Std F1-Score | Best Params |
| Experiment\_1\_Logistic\_Regression | incongruous-jay-494 | 0.935881 | 0.935072 | 0.933259 | 0.001509 | {'log\_reg\_\_C': 0.1} |
| Experiment\_1\_Logistic\_Regression | bittersweet-ant-96 | 0.932143 | 0.931239 | 0.929033 | 0.002629 | {'log\_reg\_\_C': 0.01} |
| Experiment\_2\_RandomForest | sincere-panda-125 | 0.998571 | 0.998547 | 0.997933 | 0.000441 | {'classifier\_\_max\_depth': None, 'classifier\_\_min\_samples\_split': 2, 'classifier\_\_n\_estimators': 200} |
| Experiment\_2\_RidgeClassifier | silent-grouse-455 | 0.852381 | 0.849716 | 0.854142 | 0.001552 | None |

## Observations

1. Logistic Regression achieved an average accuracy of ~93%, making it a reliable linear classifier for this task.  
2. Ridge Classifier showed lower accuracy (~85%) compared to Logistic Regression and Random Forest, indicating its limited performance for high-dimensional data like MNIST.  
3. Random Forest significantly outperformed other models with an accuracy of ~99.86% and an F1-Score of 0.9985. Its hyperparameter tuning helped achieve optimal results.  
4. Cross-validation results for Random Forest indicate consistent performance with minimal variance.

## Conclusion

Based on the results, Random Forest is selected as the final deployment model due to its superior performance metrics. Its ability to handle high-dimensional data and deliver near-perfect classification accuracy makes it the ideal choice for this task.

## Results on PCA-Reduced Data

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Accuracy | F1-Score | Precision | Recall |
| Logistic Regression | 0.9268 | 0.9257 | 0.9266 | 0.9268 |
| Ridge Classifier | 0.8547 | 0.8520 | 0.8558 | 0.8547 |
| Random Forest | 0.9999 | 0.9999 | 0.9999 | 0.9999 |
| XGBoost | 0.9999 | 0.9999 | 0.9999 | 0.9999 |

## Analysis of PCA-Reduced Data Results

The results from the PCA-reduced data demonstrate that dimensionality reduction did not significantly degrade the performance of the models. Random Forest and XGBoost achieved near-perfect accuracy and F1-Scores, confirming their robustness even with fewer features. Logistic Regression and Ridge Classifier showed slightly lower performance, but Logistic Regression maintained reasonable accuracy and F1-Scores, making it a viable option for faster computation.

## Performance on Noisy Data

The following table summarizes the performance of models on noisy data:

|  |  |  |
| --- | --- | --- |
| Model | Accuracy | F1-Score |
| Logistic Regression | 0.9246 | 0.9235 |
| Random Forest | 0.9840 | 0.9839 |
| XGBoost | 0.9910 | 0.9910 |
|  |  |  |

Analysis of Noisy Data Results:

1. Logistic Regression maintained reasonable accuracy and F1-Scores on noisy data, indicating its robustness to minor perturbations in the data.

2. Random Forest and XGBoost demonstrated superior performance, achieving near-perfect accuracy and F1-Scores even with noise, highlighting their resilience to noisy data.