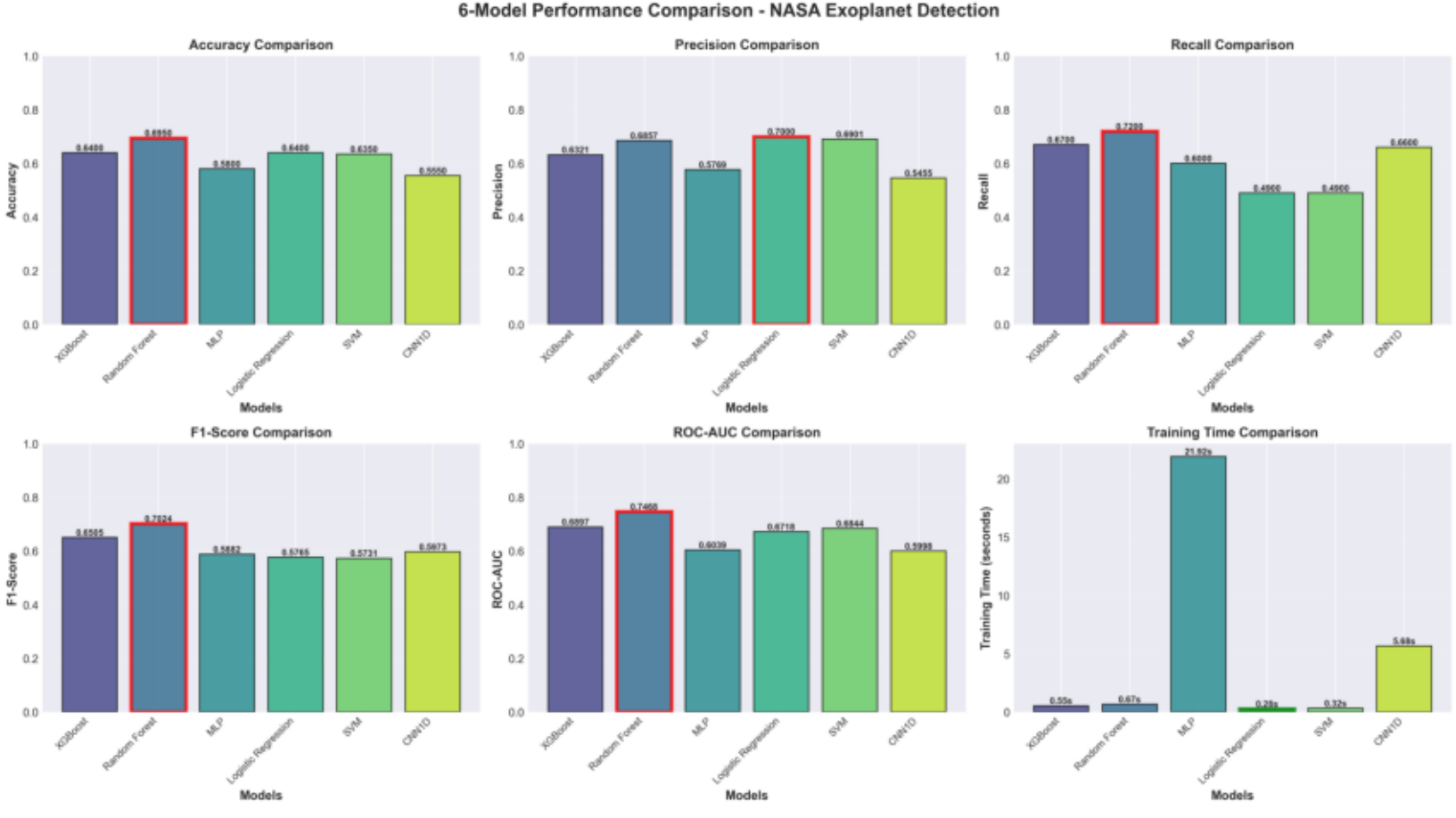
NASA Exoplanet Detection

6-Model Machine Learning Pipeline

Final Comparison Report

```
Report Generated: 2025-10-05 22:29:26

Dataset Information:
• Total Samples: 1,000
• Training Set: 600 samples
• Validation Set: 200 samples
• Test Set: 200 samples

Pipeline Summary:
• Total Models Tested: 6
• All Models Successful: [OK]
• Total Training Time: 29.42 seconds

Best Performers:
• Highest Accuracy: Random Forest (0.6950)
• Highest ROC-AUC: Random Forest (0.7468)
• Fastest Training: Logistic Regression (0.28s)

Average Performance:
• Accuracy: 0.6242
• ROC-AUC: 0.6661
```

Performance Metrics Comparison

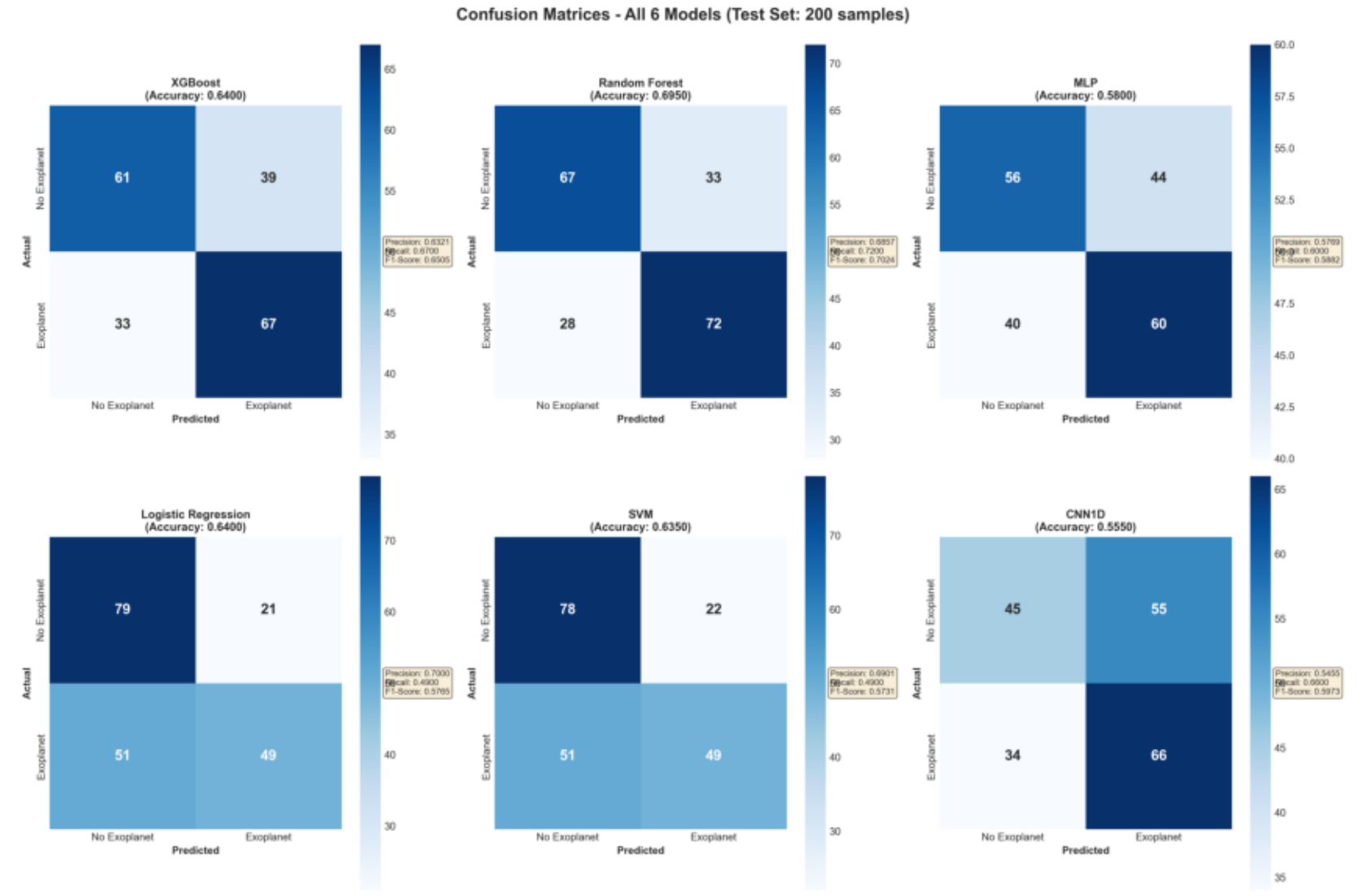


Model Performance Ranking Table
NASA Exoplanet Detection - Test Set Results

Rank	Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC	Time (s)
1	Random Forest	0.6950	0.6857	0.7200	0.7024	0.7468	0.67
2	XGBoost	0.6400	0.6321	0.6700	0.6505	0.6897	0.55
3	SVM	0.6350	0.6901	0.4900	0.5731	0.6844	0.32
4	Logistic Regression	0.6400	0.7000	0.4900	0.5765	0.6718	0.28
5	MLP	0.5800	0.5769	0.6000	0.5882	0.6039	21.92
6	CNN1D	0.5550	0.5455	0.6600	0.5973	0.5998	5.68

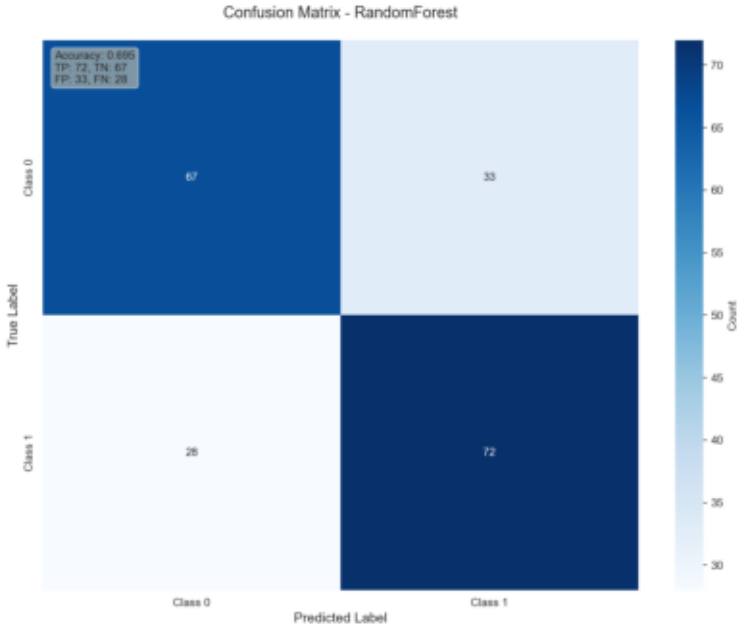
Dataset: 1000 samples | Train: 600 | Validation: 200 | Test: 200 | Total Time: 29.42s

Confusion Matrices - All Models



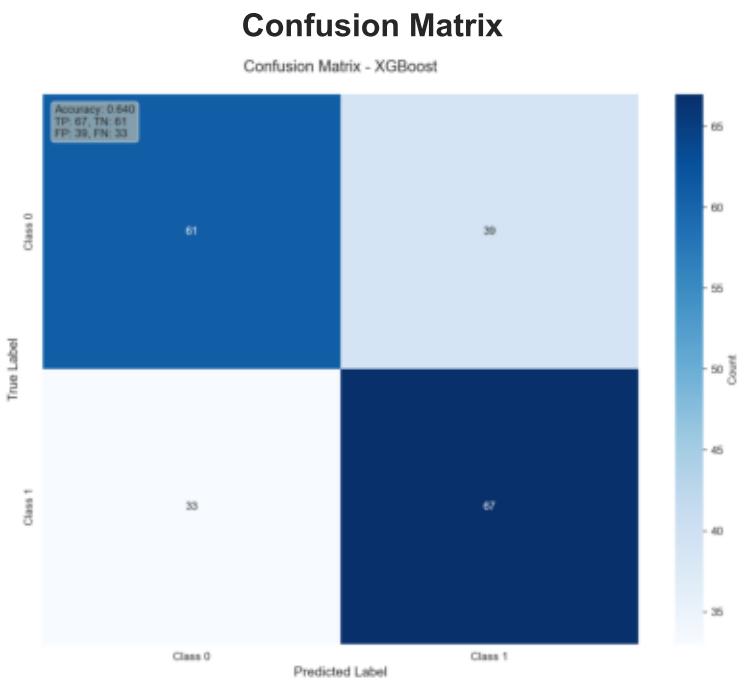
Detailed Analysis: Random Forest

Confusion Matrix




```
Performance Metrics:

Accuracy:     0.6950
Precision:    0.6857
Recall:       0.7200
F1-Score:     0.7024
ROC-AUC:      0.7468

Training Time: 0.67 seconds

Data Split:
• Train: 600 samples
• Validation: 200 samples
• Test: 200 samples
```

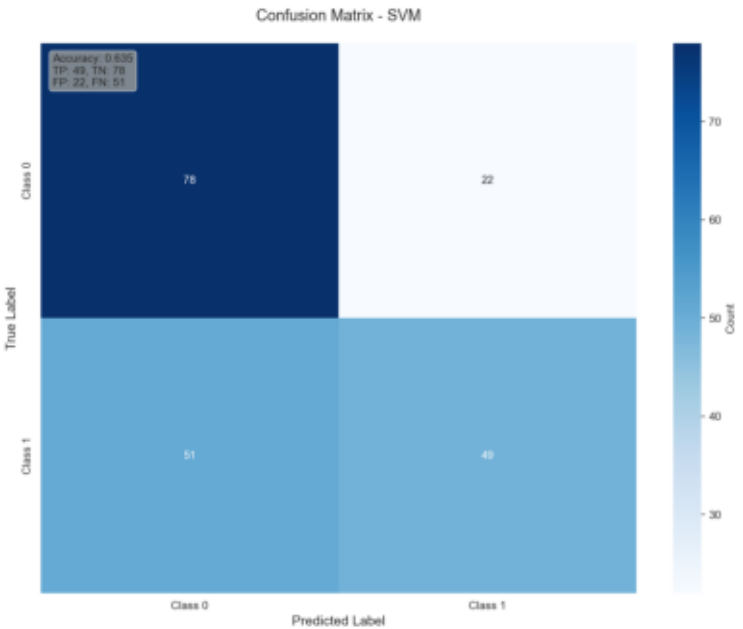
Detailed Analysis: XGBoost




```
Performance Metrics:

Accuracy:     0.6400
Precision:    0.6321
Recall:       0.6700
F1-Score:     0.6505
ROC-AUC:      0.6897

Training Time: 0.55 seconds

Data Split:
• Train: 600 samples
• Validation: 200 samples
• Test: 200 samples
```

Detailed Analysis: SVM

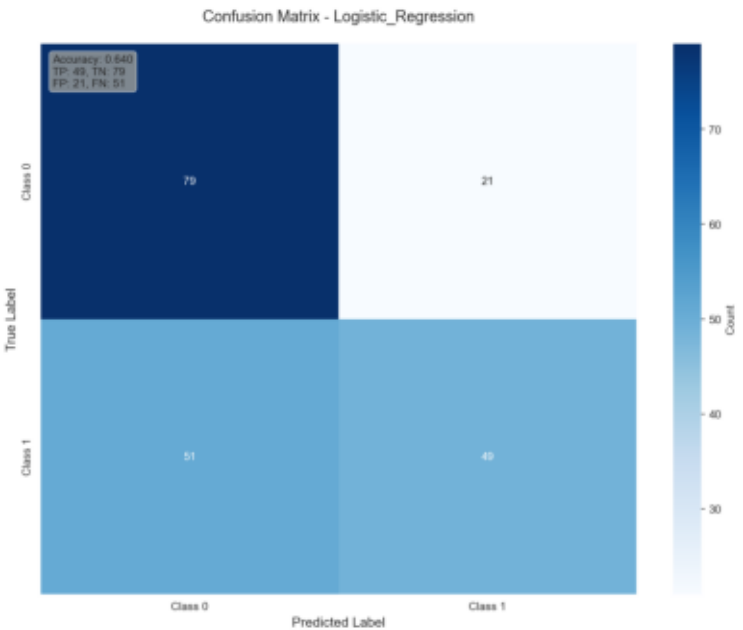
Confusion Matrix




```
Performance Metrics:

Accuracy:     0.6350
Precision:    0.6901
Recall:       0.4900
F1-Score:     0.5731
ROC-AUC:      0.6844

Training Time: 0.32 seconds

Data Split:
• Train: 600 samples
• Validation: 200 samples
• Test: 200 samples
```

Detailed Analysis: Logistic Regression

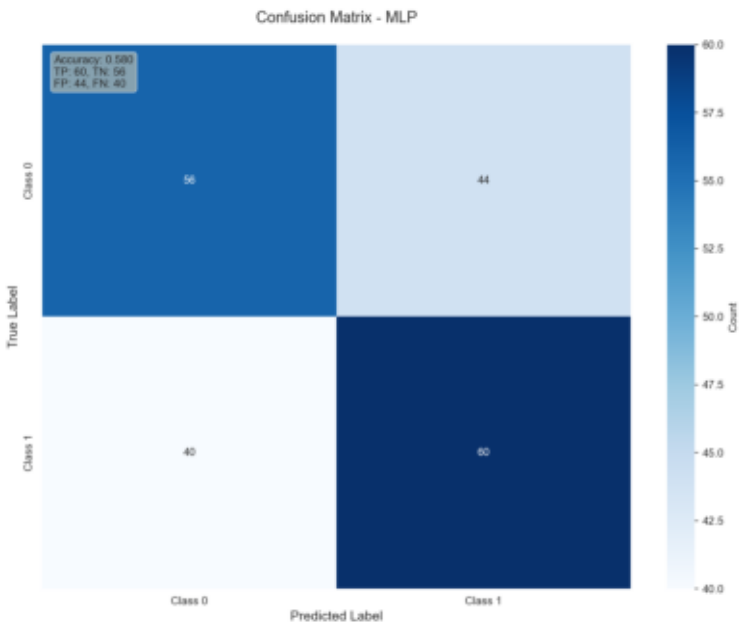
Confusion Matrix




```
Performance Metrics:

Accuracy:     0.6400
Precision:    0.7000
Recall:       0.4900
F1-Score:     0.5765
ROC-AUC:      0.6718

Training Time: 0.28 seconds

Data Split:
• Train: 600 samples
• Validation: 200 samples
• Test: 200 samples
```

Detailed Analysis: MLP

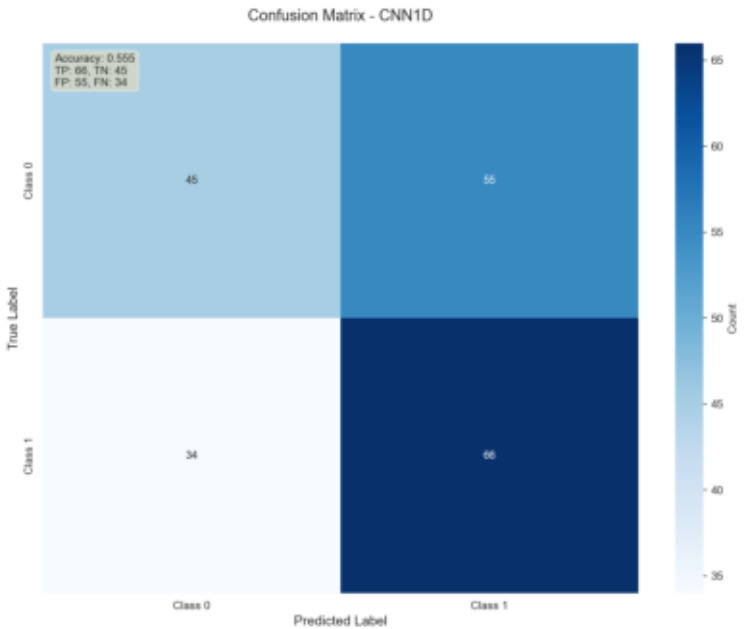
Confusion Matrix




```
Performance Metrics:

Accuracy:     0.5800
Precision:    0.5769
Recall:       0.6000
F1-Score:     0.5882
ROC-AUC:      0.6039

Training Time: 21.92 seconds

Data Split:
• Train: 600 samples
• Validation: 200 samples
• Test: 200 samples
```

Detailed Analysis: CNN1D

Confusion Matrix




```
Performance Metrics:

Accuracy:     0.5550
Precision:    0.5455
Recall:       0.6600
F1-Score:     0.5973
ROC-AUC:      0.5998

Training Time: 5.68 seconds

Data Split:
• Train: 600 samples
• Validation: 200 samples
• Test: 200 samples
```