Logistic Regression Classifier Report

Project: Multi-Class Classification of Personality Types (MBTI)

1. Executive Summary

The Logistic Regression Classifier (Multinomial/Softmax) was trained to predict 16 MBTI personality types based on 60 survey questions.

2. Model Configuration

Data Split:

- Training: 70% (42,023 samples)
- Validation: 15% (8,976 samples)
- Test: 15% (9,000 samples)

Hyperparameters:

- Algorithm: Logistic Regression (Multinomial)
- Solver: L-BFGS
- Regularization (C): 1.0
- Max Iterations: 1000 (converged in 24)
- Feature Scaling: StandardScaler

3. Performance Metrics

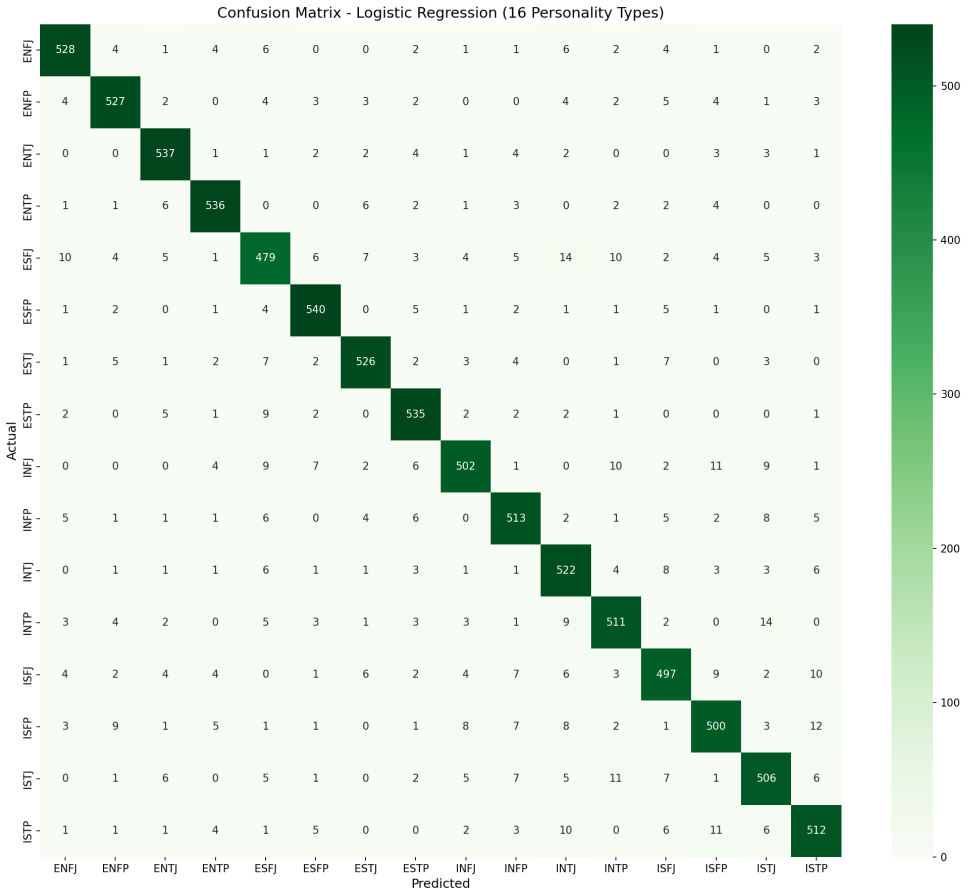
Measure	Score	Notes
Test Accuracy	91.90%	Overall classification accuracy
Validation Accuracy	92.12%	Validation set performance
Train Accuracy	92.35%	Training set performance
Top-3 Accuracy	98.62%	Correct type in top 3 predictions
Macro F1-Score	0.9189	Balanced across all 16 classes

4. Visualizations

4.1 Confusion Matrix

Logistic Regression Classifier Report

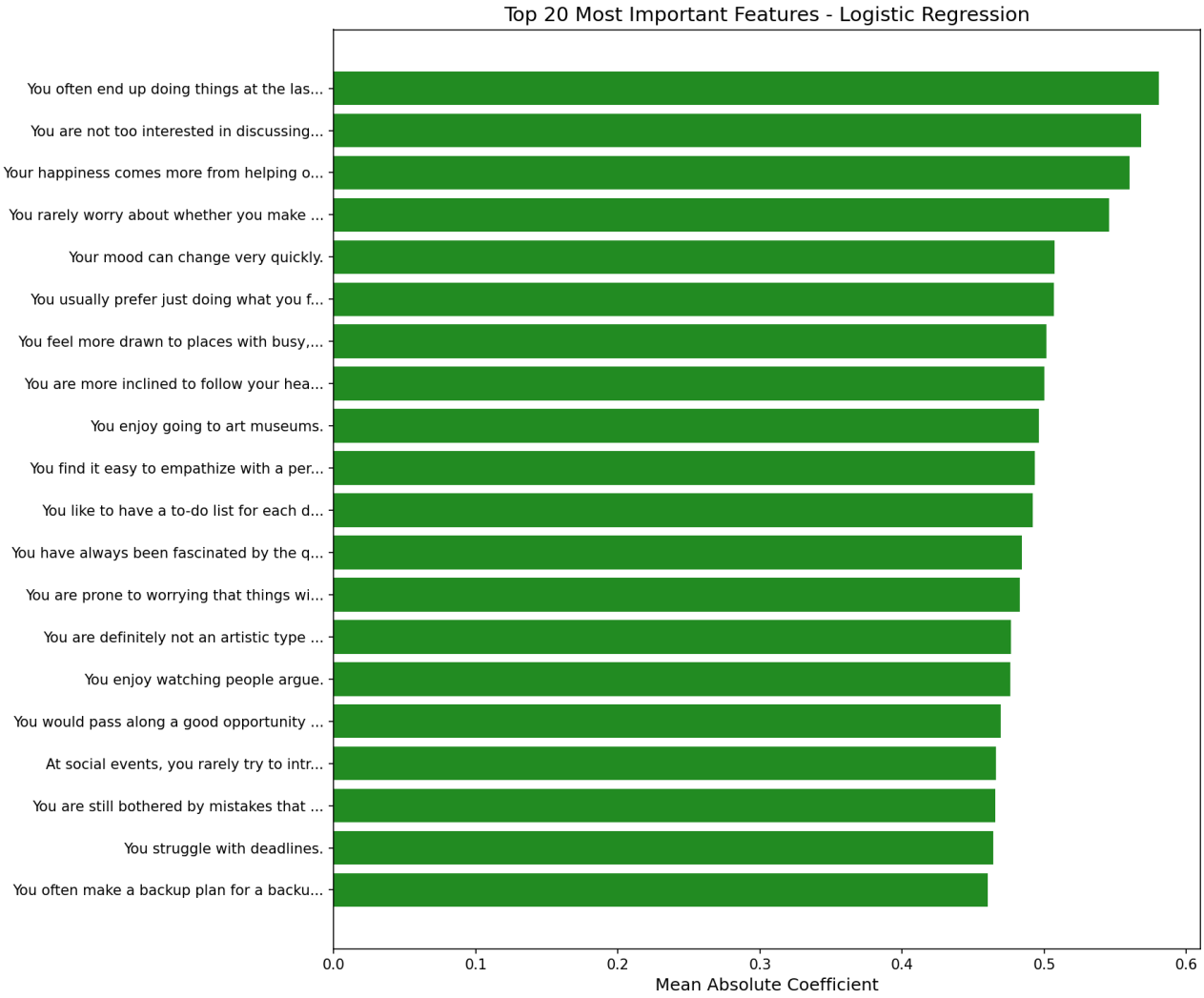
Project: Multi-Class Classification of Personality Types (MBTI)



Logistic Regression Classifier Report

Project: Multi-Class Classification of Personality Types (MBTI)

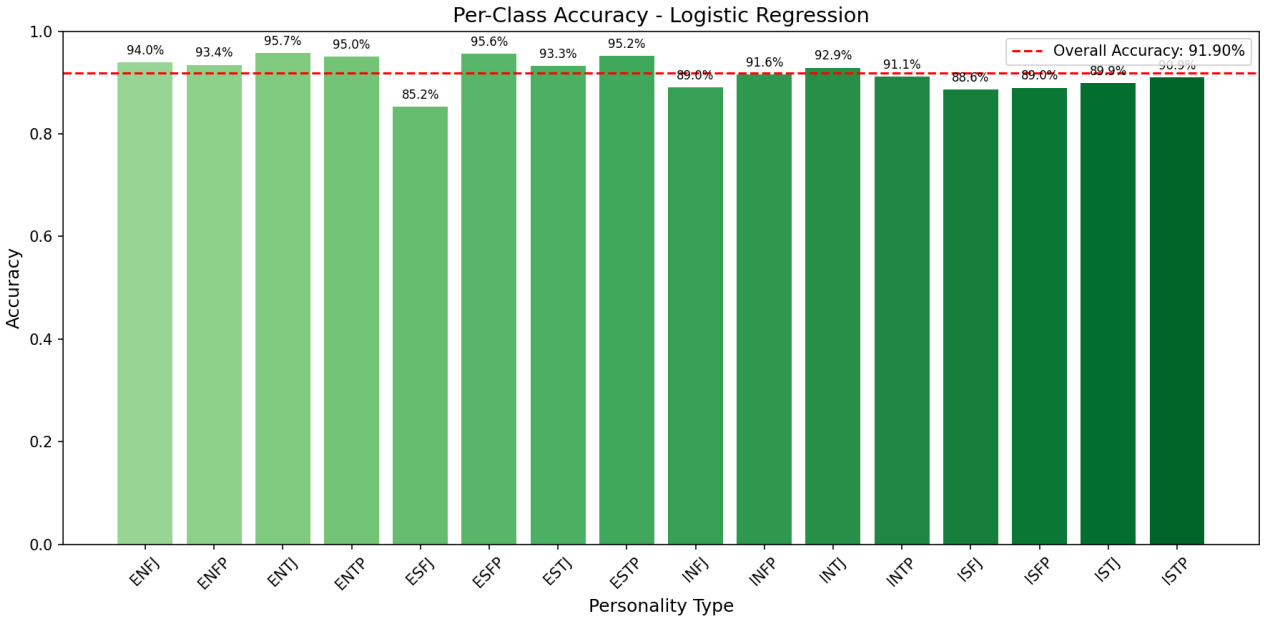
4.2 Feature Importance



4.3 Per-Class Accuracy

Logistic Regression Classifier Report

Project: Multi-Class Classification of Personality Types (MBTI)



Logistic Regression Classifier Report

Project: Multi-Class Classification of Personality Types (MBTI)

5. Conclusion

Logistic Regression achieves 91.90% test accuracy on the 16-class MBTI personality prediction task. Its coefficient-based feature importance provides clear, interpretable insights into which survey questions are most predictive of personality types.

Key Takeaways:
- Strong generalization (train-test gap only 0.45%)
- High top-k accuracy (98.62% for top-3)
- Interpretable coefficients for feature analysis
- Model simplicity and faster training time