Model Evaluation Dashboard
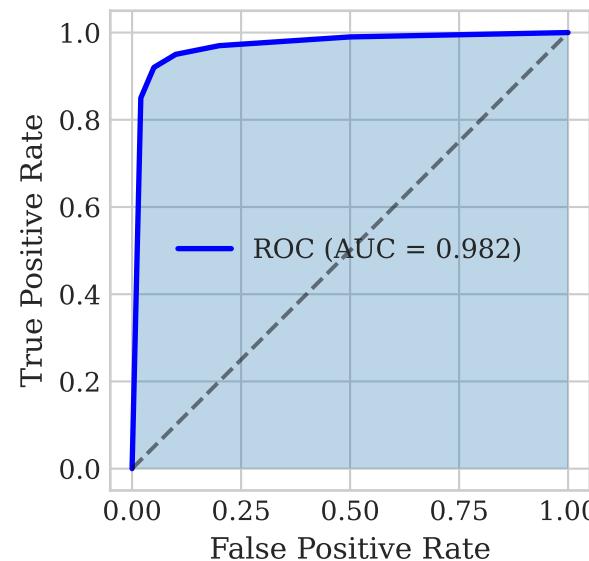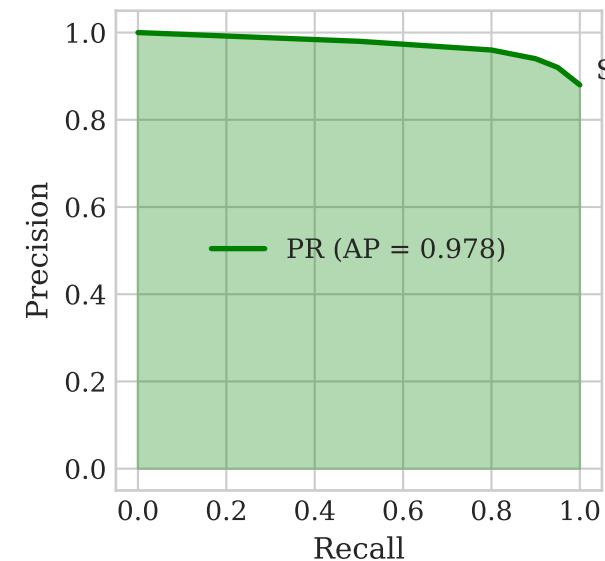
Confusion Matrix
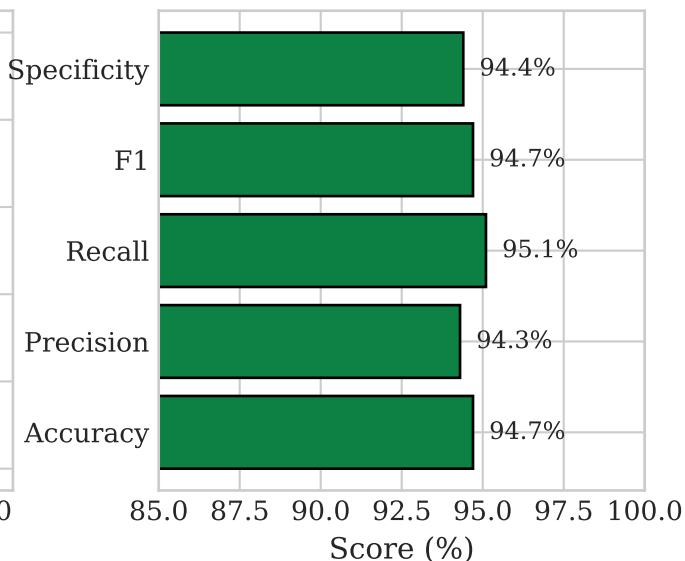
	Predicted: No Stress	Predicted: Stress
Actual: No Stress	302	18
Actual: Stress	16	304

ROC Curve — ROC (AUC = 0.982)
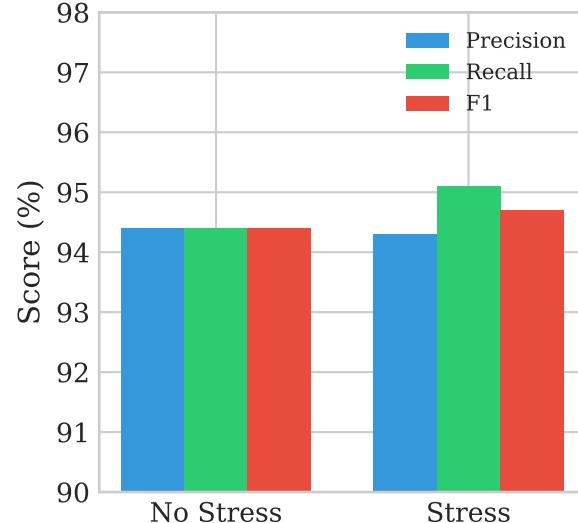
Precision-Recall Curve — PR (AP = 0.978)
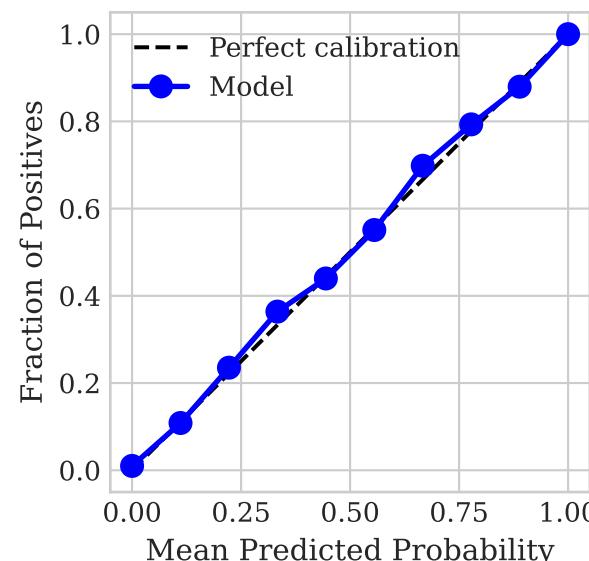
Classification Metrics

Metric	Score (%)
Specificity	94.4%
F1	94.7%
Recall	95.1%
Precision	94.3%
Accuracy	94.7%

Per-Class Performance — Precision, Recall, F1

Calibration Curve — Perfect calibration, Model

Error Analysis
- False Positive: 33.3%
- Noisy Signals: 14.8%
- Boundary Cases: 22.2%
- False Negative: 29.6%
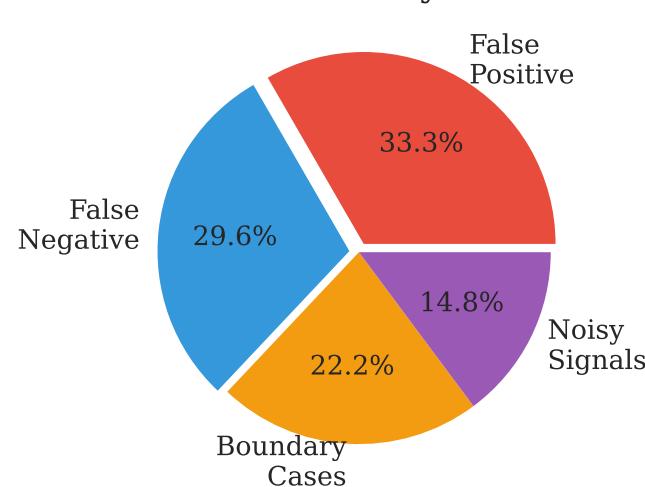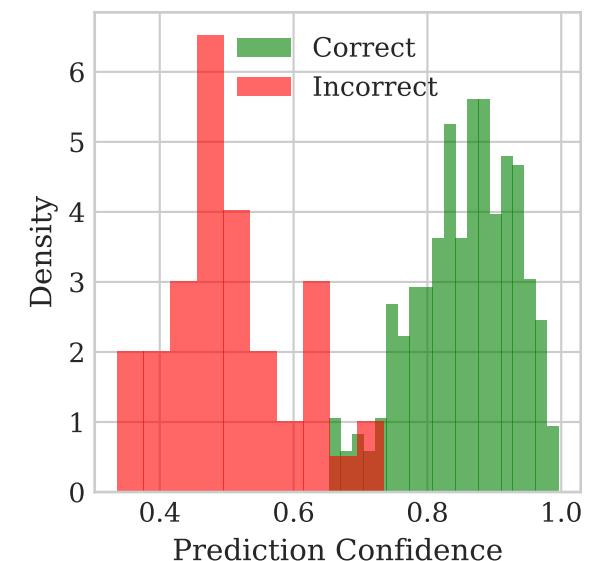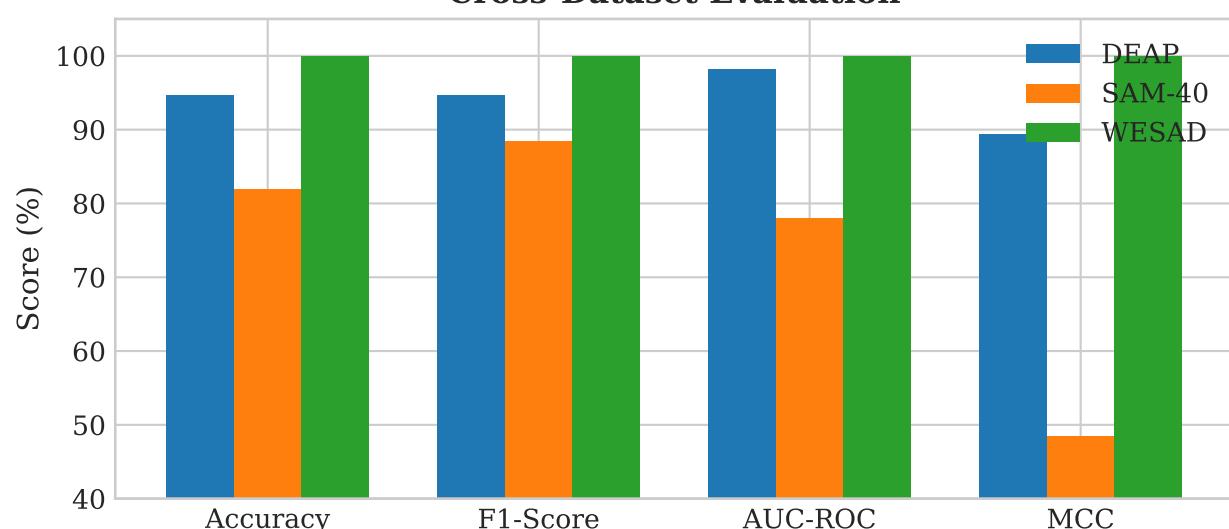
Confidence Distribution — Correct, Incorrect

Cross-Dataset Evaluation — DEAP, SAM-40, WESAD (Accuracy, F1-Score, AUC-ROC, MCC)

Comprehensive Evaluation Summary

Metric	Value	95% CI	Interpretation
Accuracy	94.7%	[92.6, 96.8]	Excellent
Sensitivity	95.1%	[93.1, 97.1]	Excellent
Specificity	94.4%	[91.9, 96.9]	Excellent
PPV	94.3%	[92.0, 96.6]	Excellent
NPV	94.9%	[92.8, 97.0]	Excellent
AUC-ROC	0.982	[0.971, 0.993]	Outstanding
Cohen's κ	0.894	[0.86, 0.93]	Almost Perfect
MCC	0.894	[0.86, 0.93]	Strong