Statistical Comparison: Base Model vs Adapted Model
(All Metrics with 95% Confidence Intervals)

All improvements are statistically significant
($p < 0.05$, 95% CI from 5-fold CV)

- Base Model (5-Fold CV)
- Adapted Model (TTT) (5-Fold CV)

Accuracy
- 0.699±0.028
- +28.9%
- 0.904±0.039 (shown near Precision)

Precision
- 0.699±0.028
- +29.4%
- 0.904±0.039

Recall
- 0.699±0.028
- +38.4%
- 0.967±0.017

F1-Score
- 0.685±0.029
- +26.2%
- 0.864±0.064

MCC
- 0.463±0.066
- +59.6%
- 0.740±0.118

Score

Performance Metrics