Appendix: Tension PTX Early Detection

DataAnalytics Team7

A Experimental Configuration

A.1 Computing Environment

A.1.1 Hardware Specifications

- **CPU**: AMD Ryzen 9 5900X (12-core, 24-thread)
- **Memory**: 64GB RAM
- **GPU**: NVIDIA GeForce RTX 3080
- **Operating System**: Linux x86_64

A.1.2 Software Configuration

The experiments were conducted with the following computational setup:

- **XGBoost**: CUDA-accelerated training using GPU
- **LightGBM**: CPU-only training (GPU Tree Learner not enabled)
- **Cross-Validation**: 5-fold stratified cross-validation
- **Optuna Trials**: 100 trials per model
- **Feature Selection**: Permutation importance-based (3-fold CV)
- **Total Features**: 156 engineered features to 76 selected features (36.7% reduction)

A.2 Hydra Configuration

The complete Hydra configuration used for the experiments:

Table 1: Data and Pipeline Configuration

Parameter	Value
Features Path	data/features_preprocessed.csv
Output Directory	data
Meta Columns	subject_id, hadm_id, label
CV Folds	5
Resampling Method	none
Cost-Sensitive Learning	True
Use GPU	True
Target Recall	0.8
Threshold Method	f1
Random Seed	42

Table 2: Imputation and Threshold Configuration

Parameter	Value
Imputer Strategy	median
Iterative Max Iterations	10
Default Threshold	0.5

Table 3: Feature Importance Configuration

Parameter	Value
Top N Features	20
Permutation Enabled	True
Permutation Repeats	5
Importance Threshold	0.0
Minimum Features	10
RFE Enabled	False
Null Importance Enabled	False

Table 4: Optuna Hyperparameter Optimization Configuration

Parameter	Value
Number of Trials	100
Timeout	null
Number of Jobs	1
Sampler Type	TPESampler
Sampler Seed	42
Startup Trials	10
Multivariate	False
Pruner Type	MedianPruner
Pruner Startup Trials	5
Warmup Steps	2
Interval Steps	1
Optimization Direction	maximize
Optimization Metric	recall
Show Progress Bar	True

Table 5: Ensemble Configuration

Parameter	Value
Soft Voting Metric	AUROC
Stacking Solver	lbfgs
Stacking C	1.0
Stacking Random State	42

B Optimal Hyperparameters

B.1 XGBoost Hyperparameters

The optimal hyperparameters found by Optuna for XGBoost (Trial 69, Best Recall: 1.0164):

Table 6: Optimal XGBoost Hyperparameters

Hyperparameter	Value
n_estimators	158
max_depth	4
learning_rate	0.041089
subsample	0.881951
colsample_bytree	0.617206
reg_alpha	6.295268
reg_lambda	0.177507
min_child_weight	1
gamma	0.587449
pos_weight_multiplier	13.208315
tree_method	hist
device	cuda
objective	binary:logistic
eval_metric	auc
early_stopping_rounds	30

B.2 LightGBM Hyperparameters

The optimal hyperparameters found by Optuna for LightGBM (Trial 88, Best Recall: 1.0837):

Table 7: Optimal LightGBM Hyperparameters

Hyperparameter	Value
n_estimators	263
max_depth	7
learning_rate	0.012187
subsample	0.651025
colsample_bytree	0.999829
reg_alpha	9.807521
reg_lambda	5.234899
min_child_samples	27
num_leaves	48
min_split_gain	0.688953
pos_weight_multiplier	17.688157

B.3 Ensemble Weights

Table 8: Global Optimization Ensemble Weights

Model	Weight
XGBoost	0.96
LightGBM	0.04

C Feature Importance

C.1 XGBoost Feature Importance

Complete feature importance scores from XGBoost model (76 features):

Table 9: XGBoost Feature Importance (All 76 Features)

Feature	Importance
w120_RR_kurt	0.041717
all_SHOCK_INDEX_kurt	0.030034
all_RR_kurt	0.029135
w120_SpO2_mad	0.026487
w120_DBP_kurt	0.024531
w120_PP_kurt	0.024108
all_SHOCK_INDEX_mad	0.022697
all_DBP_mad	0.021445
all_SHOCK_INDEX_trend_p	0.020849
all_SpO2_kurt	0.019561
w120_SBP_kurt	0.018915
w120_HR_kurt	0.018896
all_SBP_kurt	0.017207
w120_SpO2_rmssd	0.016343
w120_PP_normality_p	0.015964
w120_SBP_trend_tau	0.015631
w120_RR_normality_p	0.015216
all_SBP_trend_tau	0.015001
w120_SpO2_trend_tau	0.014211
w120_RR_trend_p	0.014108
all_MAP_skew	0.013751
w120_PP_rmssd	0.013186
all_RR_trend_tau	0.013126
all_SHOCK_INDEX_skew	0.013052
all_HR_normality_p	0.012689
w120_SpO2_skew	0.012551
all_SpO2_skew	0.012390
all_PP_rmssd	0.012280
all_SpO2_trend_p	0.012086
w120_SpO2_normality_p	0.011837
all_MAP_kurt	0.011802
w120_HR_normality_p	0.011669
all_SBP_mad	0.011481
all_DBP_trend_tau	0.011464
all_SpO2_normality_p	0.011386
w120_PP_trend_p	0.011309
all_SHOCK_INDEX_rmssd	0.011159
w120_HR_sampen	0.011155
all_MAP_mad	0.011139
w120_SBP_skew	0.011069

Table 10: XGBoost Feature Importance (continued)

Feature	Importance
w120_RR_skew	0.011062
w120_MAP_rmssd	0.011049
all_HR_mad	0.011015
all_HR_trend_tau	0.010759
all_HR_kurt	0.010500
all_HR_rmssd	0.010475
all_PP_skew	0.010436
all_PP_mad	0.010287
all_PP_trend_tau	0.010243
all_DBP_rmssd	0.010230
w120_SpO2_trend_p	0.010227
all_SBP_skew	0.010192
all_DBP_normality_p	0.010149
all_SHOCK_INDEX_normality_p	0.010118
w120_DBP_mad	0.009885
w120_DBP_rmssd	0.009775
w120_SBP_rmssd	0.009606
w120_SBP_trend_p	0.009523
w120_HR_trend_tau	0.009342
w120_HR_skew	0.009192
w120_HR_rmssd	0.009187
all_SBP_rmssd	0.009180
all_SBP_trend_p	0.009122
w120_DBP_trend_tau	0.009087
w120_PP_skew	0.009015
w120_PP_mad	0.008514
all_SHOCK_INDEX_trend_tau	0.008490
all_MAP_rmssd	0.008429
all_DBP_kurt	0.008220
w120_RR_rmssd	0.008106
w120_MAP_mad	0.008090
all_HR_skew	0.007922
w120_MAP_trend_tau	0.007899
w120_SBP_mad	0.007748
w120_HR_trend_p	0.007700
w120_SpO2_sampen	0.006589

C.2 LightGBM Feature Importance

Complete feature importance scores from LightGBM model (76 features):

Table 11: LightGBM Feature Importance (All 76 Features)

Feature	Importance
all_DBP_mad	1890.921
all_SHOCK_INDEX_mad	1755.766
all_RR_kurt	1451.096
all_SHOCK_INDEX_kurt	1299.409
w120_MAP_rmssd	1242.026
w120_HR_normality_p	1130.197
w120_SpO2_skew	1035.163
w120_SpO2_mad	1034.648
w120_SpO2_trend_tau	985.235
all_SpO2_skew	939.284
w120_SBP_skew	805.771
w120_SpO2_rmssd	799.737
w120_PP_normality_p	726.450
all_DBP_normality_p	659.169
w120_HR_trend_p	641.700
all_HR_mad	623.114
w120_SBP_trend_tau	591.215
all_RR_trend_tau	561.206
w120_DBP_trend_tau	438.620
all_HR_normality_p	389.085
w120_HR_rmssd	382.264
w120_RR_skew	348.953
all_SHOCK_INDEX_trend_p	337.694
all_SHOCK_INDEX_skew	292.501
all_SBP_kurt	273.088
all_MAP_kurt	262.351
all_MAP_skew	251.157
w120_PP_trend_p	241.899
all_PP_mad	236.646
w120_SBP_kurt	233.615
all_PP_trend_tau	225.770
w120_SBP_rmssd	214.723
w120_DBP_kurt	208.040
all_SpO2_kurt	190.059
w120_RR_kurt	183.571
w120_DBP_rmssd	176.812
all_MAP_rmssd	174.548
w120_HR_trend_tau	169.278
all_DBP_kurt	159.636
w120_DBP_mad	149.222

Table 12: LightGBM Feature Importance (continued)

Feature	Importance
all_SHOCK_INDEX_rmssd	147.889
w120_PP_skew	131.488
w120_PP_rmssd	112.416
w120_MAP_trend_tau	108.044
w120_HR_skew	104.496
all_MAP_mad	100.404
w120_RR_trend_p	89.597
w120_SpO2_normality_p	89.230
w120_HR_sampen	86.341
all_PP_rmssd	84.577
all_SBP_rmssd	82.738
all_DBP_rmssd	64.592
all_HR_skew	63.463
w120_SpO2_trend_p	51.135
all_PP_skew	49.529
w120_PP_mad	42.321
w120_RR_rmssd	39.532
w120_SBP_trend_p	34.793
all_DBP_trend_tau	34.542
all_HR_rmssd	32.260
all_HR_kurt	25.777
all_SBP_mad	18.858
w120_HR_kurt	18.629
all_HR_trend_tau	16.315
all_SHOCK_INDEX_normality_p	16.059
all_SpO2_trend_p	12.607
w120_MAP_mad	9.038
all_SBP_trend_tau	6.204
all_SpO2_normality_p	0.000
all_SHOCK_INDEX_trend_tau	0.000
all_SBP_trend_p	0.000
all_SBP_skew	0.000
w120_SpO2_sampen	0.000
w120_RR_normality_p	0.000
w120_SBP_mad	0.000
w120_PP_kurt	0.000

D Model Performance Metrics

D.1 Cross-Validation Results

Table 13: Model Performance Comparison (Mean ± SD)

Model	AUROC	AUPRC	Accuracy	Precision	Recall	F1-score
Baseline (LR)	0.4989	0.5486	0.5246	0.5417	0.4194	0.4728
XGBoost	**0.7457 (±0.0238)**	**0.6623 (±0.0563)**	**0.7678 (±0.0288)**	**0.6764 (±0.0338)**	**0.7276 (±0.1147)**	**0.7011 (±0.0516)**
LightGBM	0.7103 (±0.0310)	0.6411 (±0.0653)	0.7344 (±0.0927)	0.6512 (±0.0516)	0.6844 (±0.1144)	0.6674 (±0.0405)
Soft Voting	0.7324 (±0.0324)	0.6607 (±0.0511)	0.7167 (±0.0621)	0.6410 (±0.0625)	0.6934 (±0.0682)	0.6662 (±0.0465)
Stacking	0.7324 (±0.0324)	0.6607 (±0.0511)	0.7167 (±0.0621)	0.6410 (±0.0625)	0.6934 (±0.0682)	0.6662 (±0.0465)

D.2 Data Statistics

Table 14: Dataset Statistics

Characteristic	Value
Total Samples	900
Positive Cases (Pneumothorax)	150
Negative Cases	750
Class Ratio	1:5
Total Engineered Features	156
Selected Features	76
Feature Reduction Rate	36.7%
Missing Values	0