A COMPARATIVE STUDY OF CLASSICAL AND MODERN BOOSTING ALGORITHMS

Team 13
Tony Luo, Will Kim, Sichen Li, Shang Peng, Jiaheng Guo

MOTIVATION

- We have studied AdaBoost and Gradient Boosting as ensemble methods.
- In 2025, machine learning includes highly advanced models such as deep neural networks.
- This raises a natural question:

  Have boosting algorithms themselves evolved as well?
- If so, what advancements have been made, and how do they compare to the classical methods?

RESEARCH QUESTIONS

1 What are the key modern advancements in boosting algorithms beyond AdaBoost and Gradient Boosting

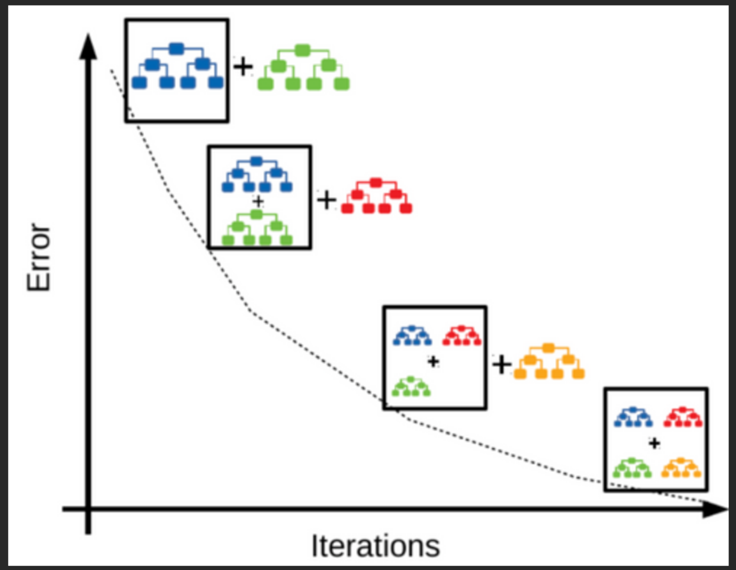
2 Under a unified preprocessing and evaluation pipeline, how do classical and modern boosting methods compare in predictive performance across diverse classification datasets?

3 Do modern boosting frameworks offer measurable improvements in computational efficiency relative to AdaBoost and Gradient Boosting?

4 How consistent are the predictive performance metrics (AUC, accuracy, F1) of these algorithms across cross-validation folds and datasets?

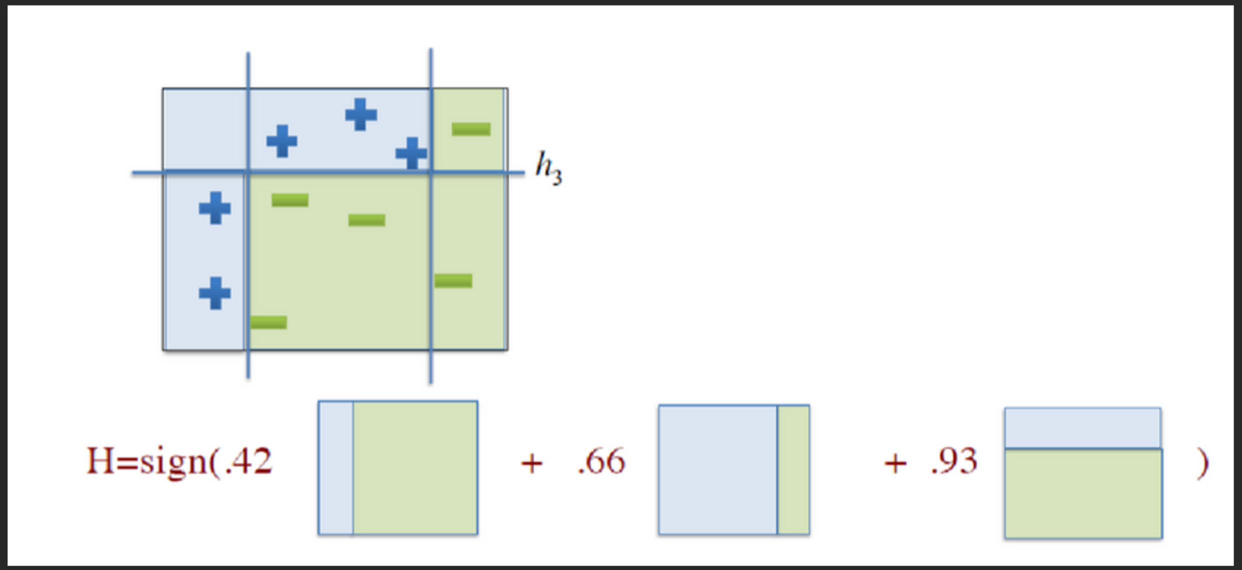
CLASSICAL BOOSTING ALGORITHM

Gradient Boosting



- Boosting as gradient descent in function space
- Fits trees to negative gradients
- Flexible but computationally heavy

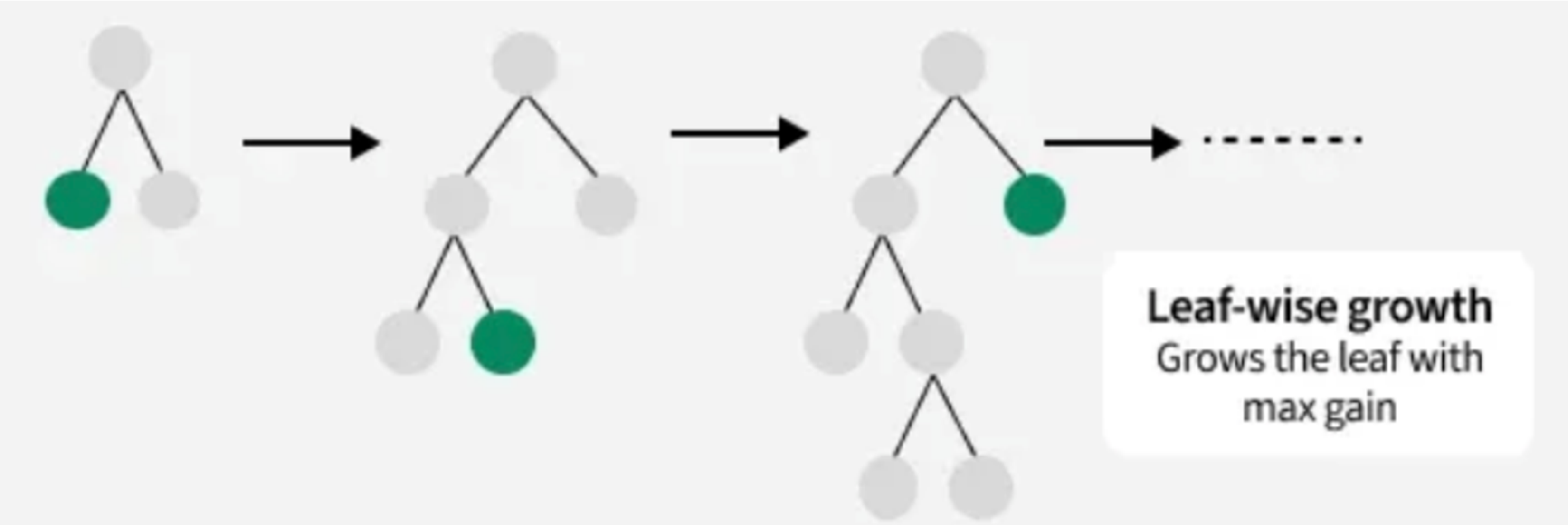
AdaBoost



- Reweights misclassified samples
- Minimizes exponential loss

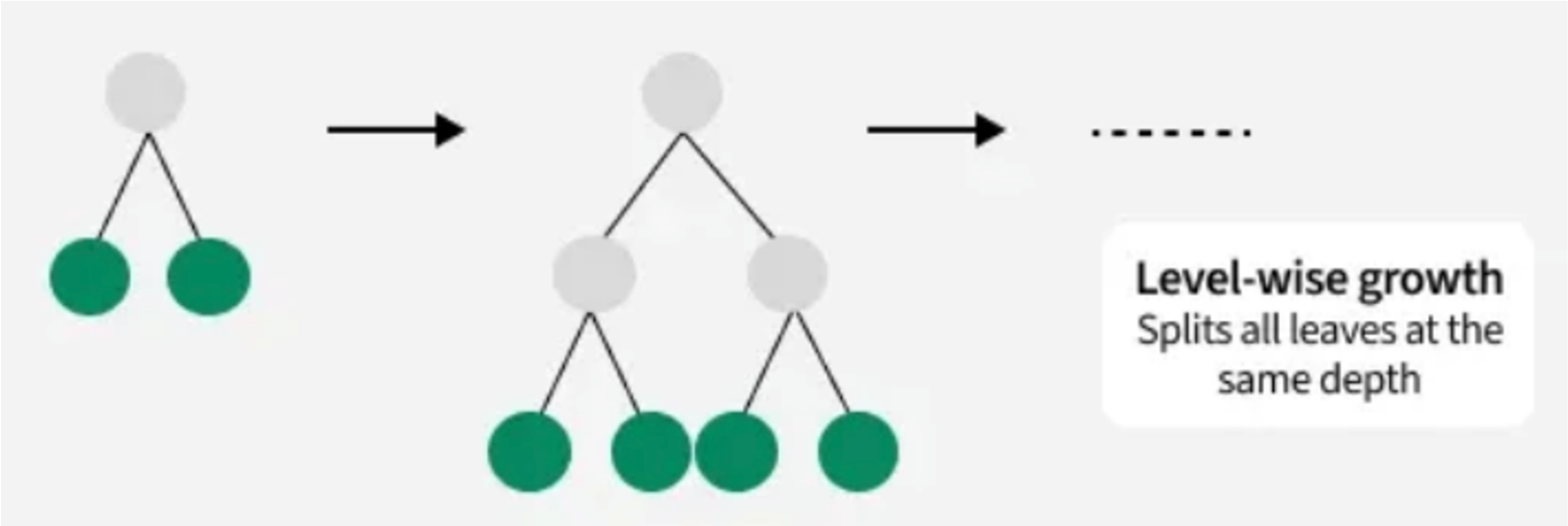
MODERN BOOSTING ALGORITHM

Light GBM



- Leaf-wise tree growth (deeper splits)
- Histogram-based splitting
- GOSS (sampling) & EFB (exclusive feature bundling)
- Extremely fast and memory efficient

XGBoost



- Regularized objective (L1/L2)
- Splitting with weighted quantile sketch
- Efficient handling of sparse features
- Parallelized tree building

DATASETS

Classification

Regression

Classification Dataset

Classficiation Datasets	#Samples	#Features	#Classes	Subject Area	Data Type
Adult Income	48,842	14	2	Social Science	Mixed
Heart Disease	303	13	2	Health	Mixed
Mushrooms	8,124	22	2	Biology	Categorical
Telco Customer Churn	7,043	33	2	Business	Mixed
Breast Cancer	569	30	2	Health	Numerical
Credit Card Fraud	284,807	30	2	Business	Numerical
IMDB Movie Review	49,582	-	2	Entertainment	Text
MNIST	70,000	784	10	Computer Science	Image
HIGGS	11,000,000	28	2	Physics	Numerical

Regression Dataset

Regression Datasets	#Samples	#Features	Subject Area
California Housing	20,640	8	Real Estate
Ames House Prices	1,460	80	Real Estate
Wine Quality	4,898	11	Business
Superconductivity	21,263	81	Physics

EXPERIMENT SETUP

Unified Preprocessing Pipeline

- Numerical features: median imputation + StandardScaler (without centering)
- Categorical features: most-frequent imputation + One-Hot Encoding
- **IMDB dataset**: TF-IDF vectorization (50k max features, unigrams + bigrams)
- **MNIST dataset**: flatten images → PCA (retain 95% variance)

Train-Test Split

- 80/20 split (stratified for classification) with fixed seed for reproducibility

EXPERIMENT SETUP

Algorithms Compared

- AdaBoost
- Gradient Boosting (GBM)
- XGBoost (histogram-based splitting, regularization)
- LightGBM (leaf-wise growth, histogram bins, GOSS/EFB)

All models trained with *Optuna* for hyperparameters tuning across datasets.

EXPERIMENT SETUP

Performance Metrics

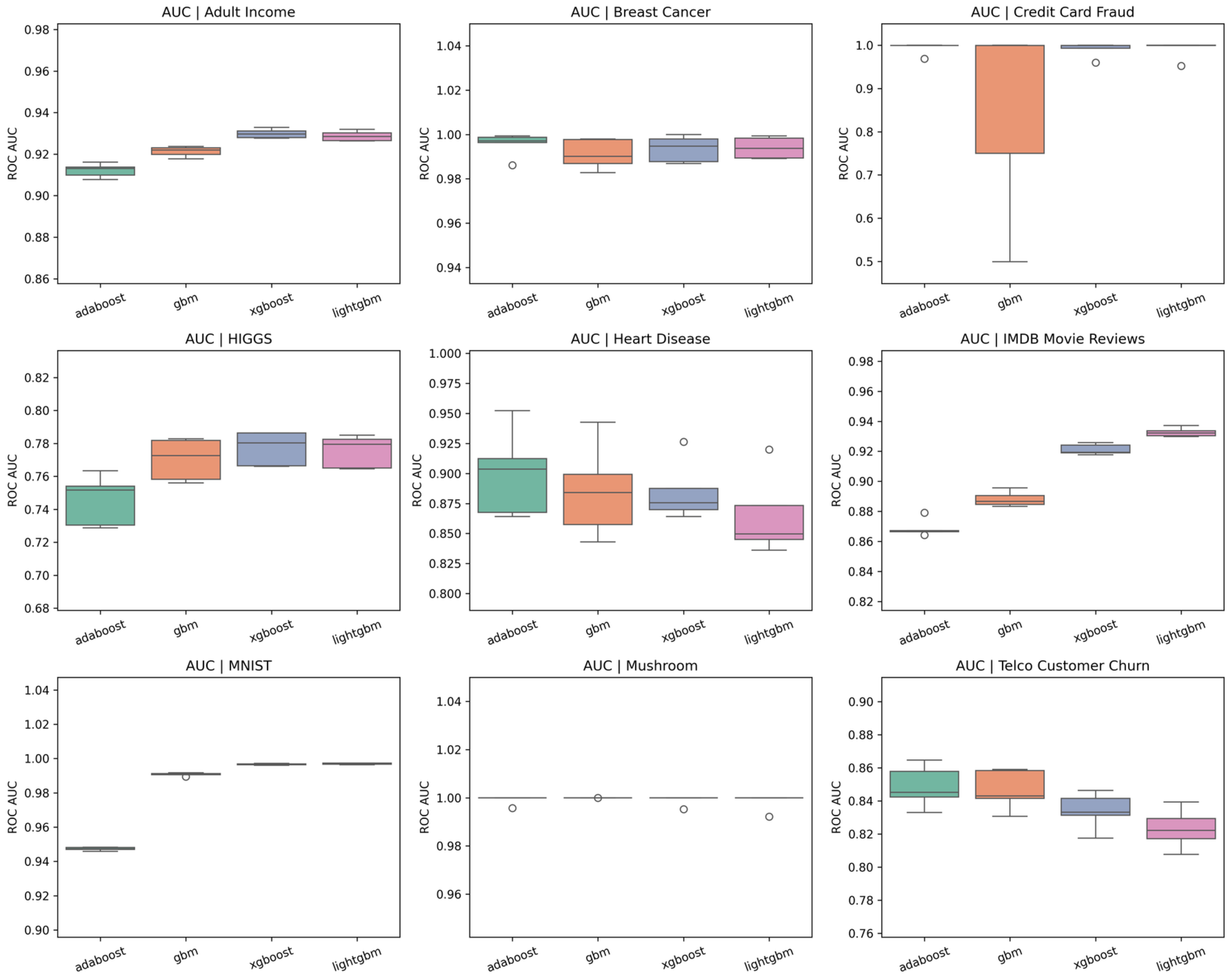
- **Classification**: ROC–AUC, Accuracy, F1-weighted
- **Regression**: RMSE, MAE, R squared
- **Training Efficiency**: total training time (seconds)

EXPERIMENT SETUP

Evaluation Procedure

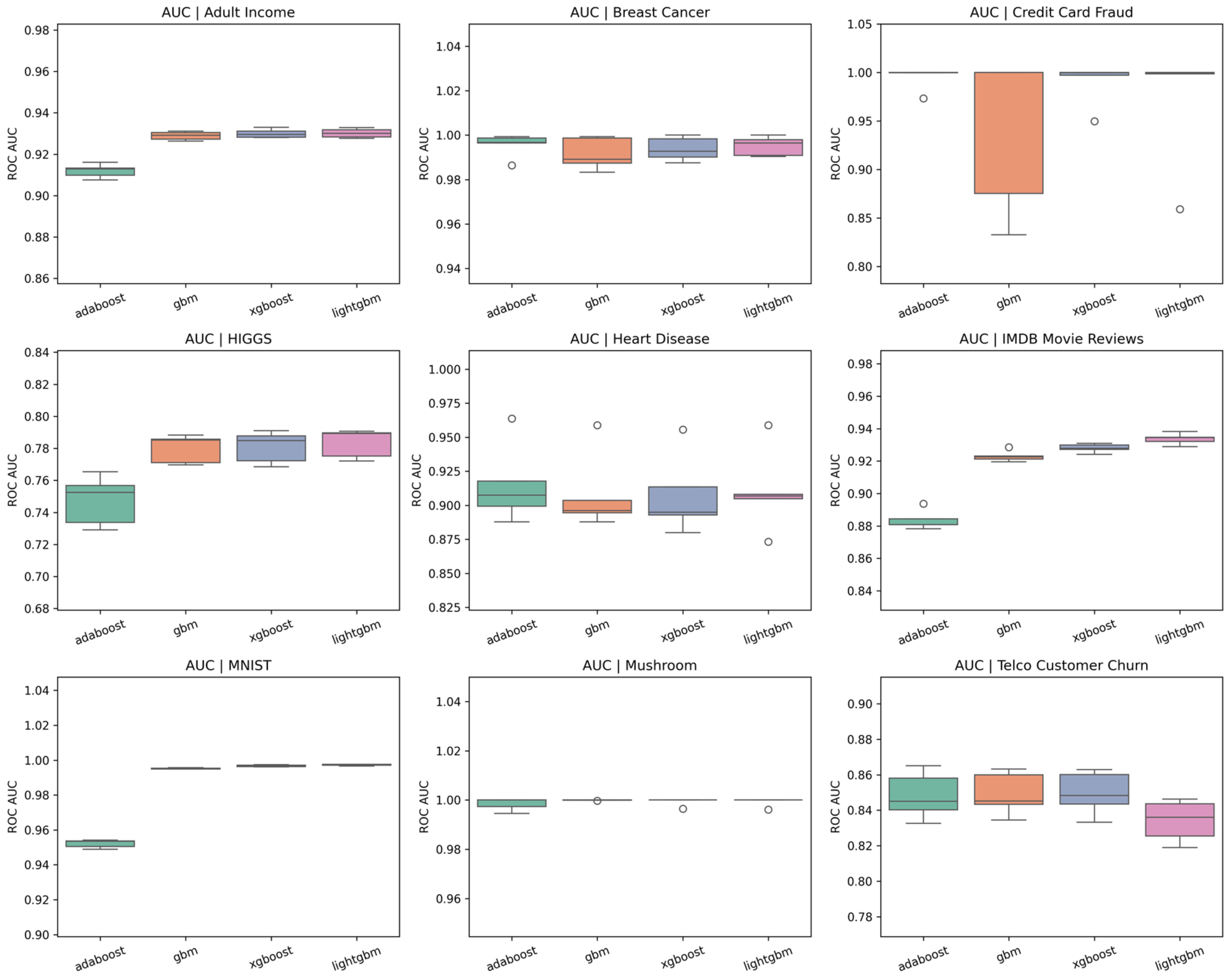
- Baseline model (no-hyperparameter tuning)
  - 5-fold Cross-Validation for each algorithm–dataset pair
  - For each fold, record the ROC, training time, etc.
- Optimized model (tuned with **Optuna** for ~20 trails)
  - 5-fold Cross-Validation for each algorithm–dataset pair
  - For each fold, record the ROC, training time, etc.

RESULTS ANALYSIS



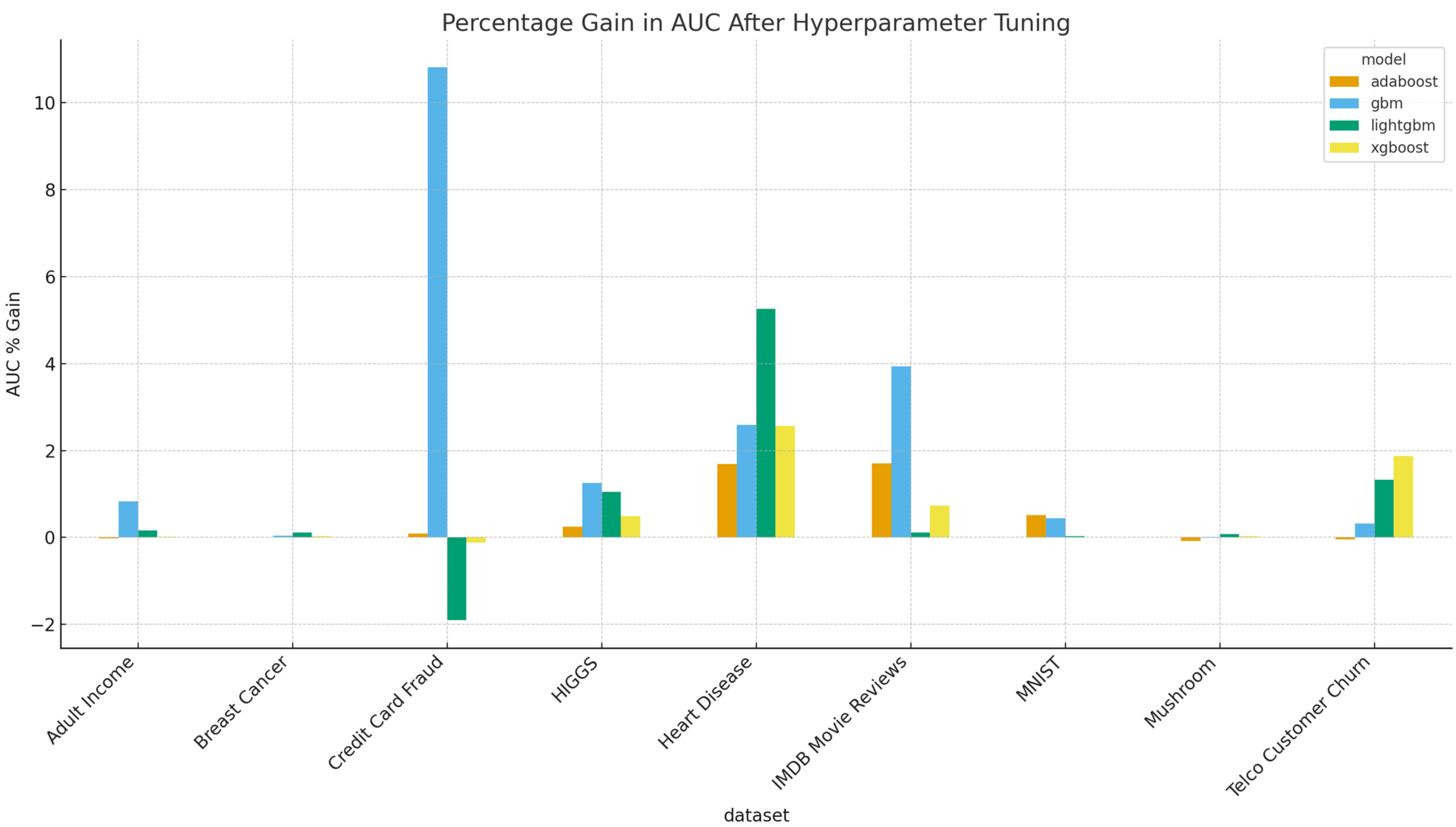
[Classification] Test AUC Scores of Baseline Models

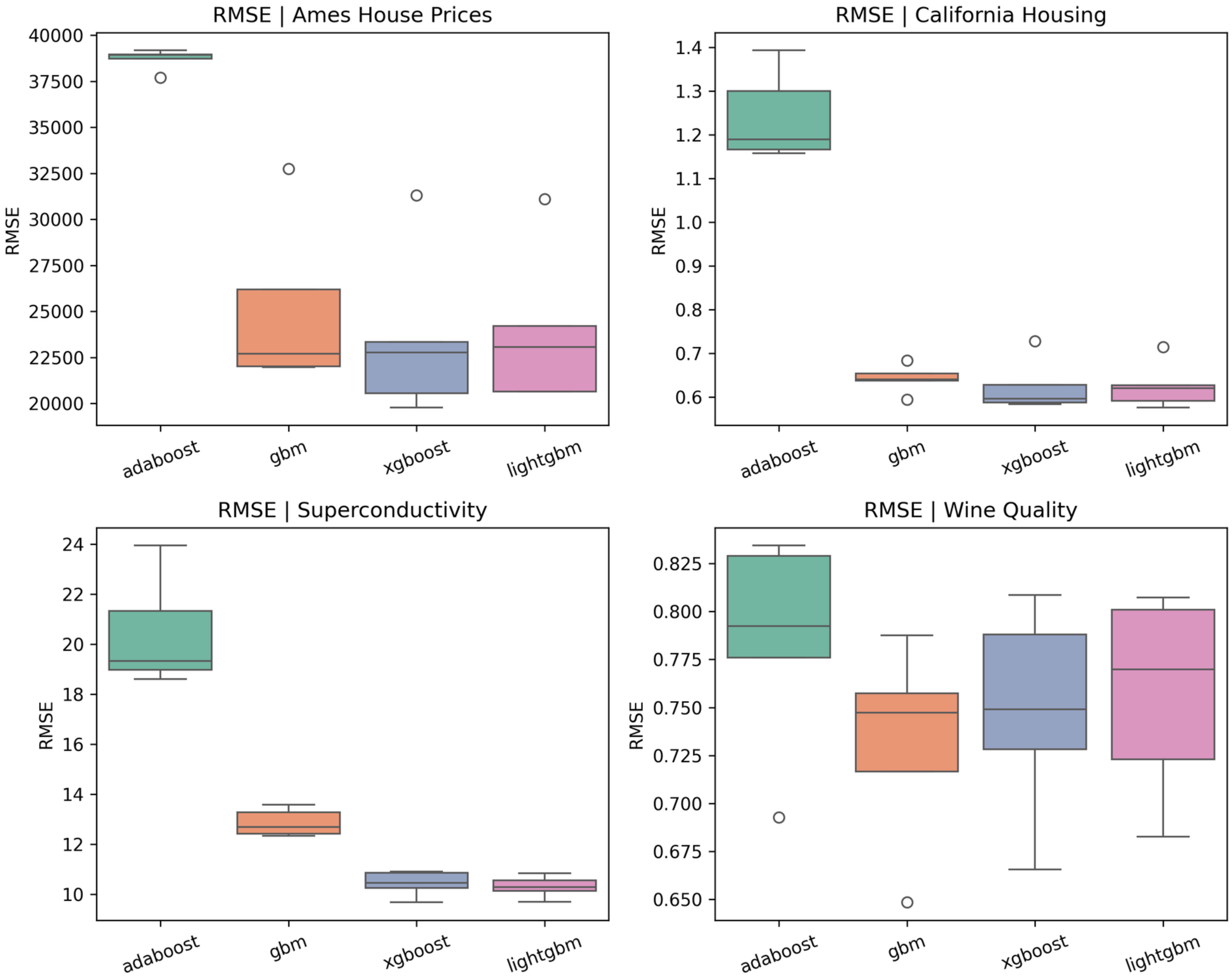
no tuning; 5-fold CV



[Classification] Test AUC Scores of Tuned Models

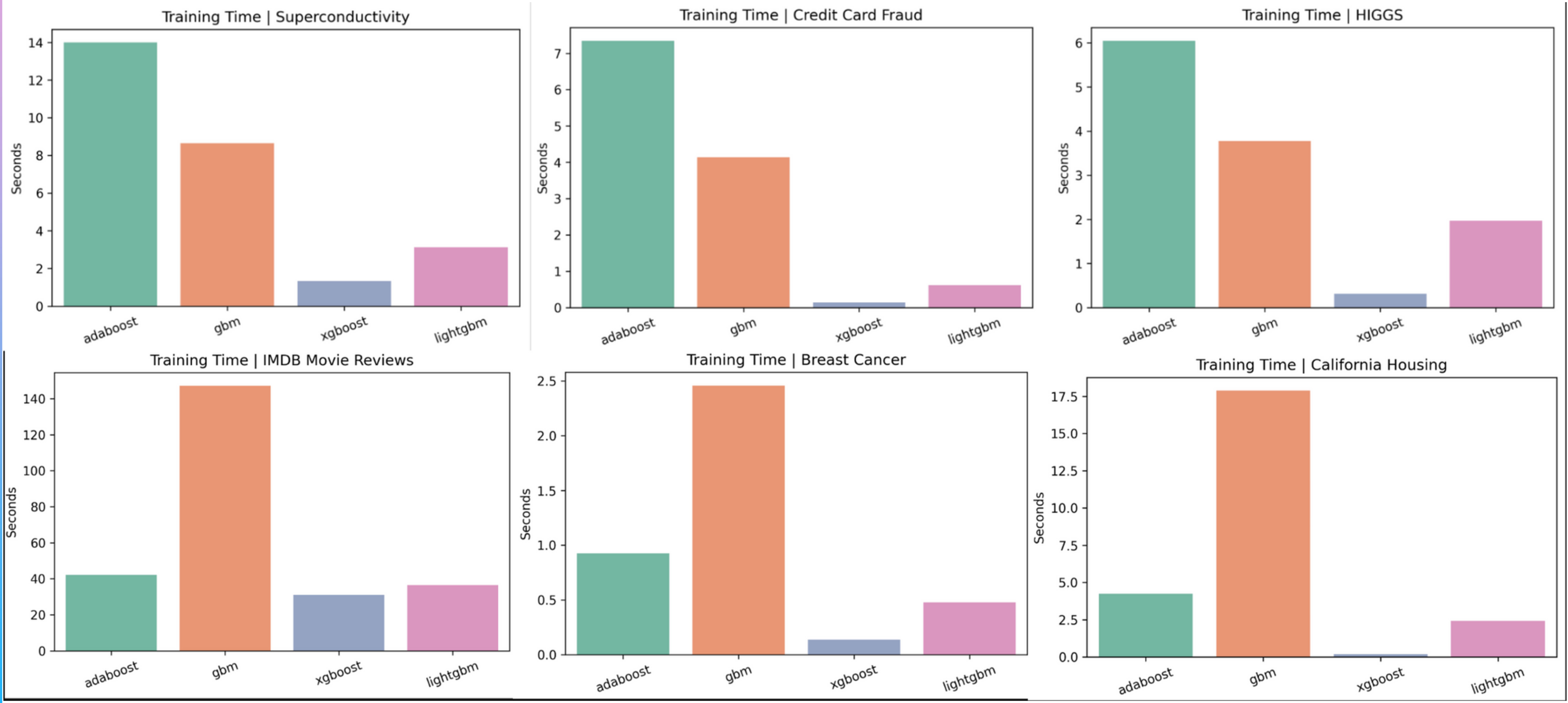
Optuna; 5-fold CV


Percentage Gain in AUC After Hyperparameter Tuning
AUC % Gain
dataset
model
adaboost
gbm
lightgbm
xgboost
10
8
6
4
2
0
−2
Adult Income
Breast Cancer
Credit Card Fraud
HIGGS
Heart Disease
IMDB Movie Reviews
MNIST
Mushroom
Telco Customer Churn



[Regression]
Test RMSE of
Tuned Models
Optuna; 5-fold CV

Selected Plots of Training Time



FINDINGS

Higher predictive performance

LightGBM and XGBoost generally achieve stronger AUC, accuracy, and F1 across datasets

Greater efficiency

Modern boosting algorithms train significantly faster due to innovative design

Scalability

Modern methods performs better in accuracy and efficiency in large-scale datasets

No dominating method!

In terms of performance, there is no single algorithm performing best across all datasets

CONCLUSION

KEY FINDINGS

- Modern boosting models (LightGBM, XGBoost) provide **clear advances** in performance, efficiency, and robustness over classical methods
- Design innovations such as leaf-wise growth, regularization, and histogram-based splitting drive these improvements

LIMITATIONS & FUTURE DIRECTIONS

Multiclass Classification

Have only considered 1 multiclass classification dataset (MNIST, and it’s a CV dataset)

Tuning Space

Modest hyperparameter tuning badget (20 Optuna trails for each algo-dataset pair)

not guarantee convergence to global optima

LIMITATIONS & FUTURE DIRECTIONS

Feature Importance

Evaluate feature-importance stability and interpretability in greater depth

Comparison with Deep Learning models

Compare boosting to deep-learning models for tabular data

(e.g., TabNet, FT-Transformer)

THANK YOU

FOR THE ATTENTION