ST03: Machine Learning Model Selection Report

Kateryna Makarova

1. Introduction

This report presents the evaluation of three machine learning models to predict Expected Travel Time (ETT) for vessels using AIS data. The models tested include:

- Random Forest Regressor
- Gradient Boosting Regressor
- Multi-Layer Perceptron (MLP) Regressor

Each model was trained and evaluated both before and after hyperparameter tuning. The objective is to identify the model with the lowest prediction error and best generalization performance.

2. Evaluation Metrics

The following metrics were used for model comparison:

- **Mean Absolute Error (MAE):** Average absolute difference between predicted and actual ETT.
- **Root Mean Squared Error (RMSE):** Similar to MAE, but penalizes larger errors more heavily.
- **R-squared ($\mathbf{R^2}$):** Proportion of variance in ETT explained by the model.

3. Model Performance Comparison

3.1. Before and After Tuning

Model	MAE (hr)	RMSE (hr)	R^2	Interpretation
Before Tuning				
Random Forest	0.46	0.96	0.88	Strong base model; slight overfitting possible.
Gradient Boosting	1.15	1.49	0.72	Underfits; high bias and variance.
MLP Regressor	1.21	1.56	0.69	Weak performance; capacity not well-tuned.
After Tuning				
Random Forest	0.78	1.16	0.83	Slight performance drop; possible overfitting due to tuning.
Gradient Boosting	0.22	0.31	0.90	Most improved model; excellent trade-off between bias and variance.
MLP Regressor	1.09	1.46	0.73	Slight improvement, but still outperformed by tree-based models.

3.2. Hyperparameter Tuning Details

Model	Tuned Parameters	Tuning Method	Observed Impact
Random Forest	`n_estimators`, `max_depth`, `min_samples_split`	RandomizedSearchCV + 3-fold CV (MAE)	Small improvement; already strong model, tuning led to minor robustness changes.
Gradient Boosting	`n_estimators`, `learning_rate`, `max_depth`	RandomizedSearchCV + 3-fold CV (MAE)	Significant drop in error; best trade-off post tuning.
MLP Regressor	`hidden_layer_sizes`, `activation`, `alpha`	GridSearchCV + 3-fold CV (MAE)	Small gains; performance still inferior to Gradient Boosting.

4. Conclusion

Gradient Boosting emerged as the most effective model for predicting ETT after tuning, showing significant improvements in all metrics. While Random Forest was the strongest base model, its performance decreased slightly post tuning, indicating possible overfitting. MLP Regressor showed minimal improvement and remained the least suitable choice.

For deployment and practical usage, the tuned Gradient Boosting model is recommended due to its superior balance of accuracy and stability.