Report

Divy Jain(210010015), Karthik Hegde(210010022)

January 26, 2024

Abstract

This report presents a comprehensive study on a Heart Rate Prediction Model, developed as part of the AI Cure: Parsec 4.0 competition. The model aims to predict heart rates with high accuracy, leveraging machine learning techniques. Extensive data pre-processing, feature extraction, and model evaluation strategies were employed to achieve optimal performance. The effectiveness of the model is demonstrated through rigorous testing and validation procedures.

1 Approach

The development of the Heart Rate Prediction Model involved several critical steps:

Pre-Processing: The initial dataset underwent thorough cleaning, including the removal of redundant features. The data was scaled appropriately, and labels were separated to prepare for model training.

Visualization: To understand the dataset's characteristics, various visualization techniques were applied. Histogram plots of feature vectors and correlation heatmaps provided insights into the data distribution and feature relationships.

Feature Extraction: The feature set was refined by eliminating highly correlated features, reducing the total from 36 to 32. This reduction was carefully done to maintain performance, with some features being averaged and others removed entirely. Here is a summary of the features and the corresponding operations applied to them. Please note that only those modifications that yielded positive outcomes have been implemented.

SNo.	Features	Method	New Feature	Implemented
1	LF_PCT and VLF_PCT	Sum	AVG1	No
2	TP and VLF	Sum	AVG2	Yes
3	HF_LF and HF_NU	Sum	AVG3	Yes
5	MEAN_RR and MEDIAN_RR	Sum	AVG4	No
4	LF_NU	remove	-	No
6	KURT_REL_RR	remove	-	Yes
7	SKEW_REL_RR	remove	-	Yes
8	SDSD_REL_RR	remove	-	Yes
9	SDRR	remove	-	No
10	SDSD	remove	-	Yes
11	SD1	remove	-	Yes
11	HF_LU	remove	-	Yes
11	HF_NU	remove	-	Yes

Table 1: Feature Modification Summary

2 Model

Several models were explored in the study:

Linear Regression: Served as the benchmark model.

Random Forest Regressor: An ensemble model for improved predictions.

XGBRegressor: Gradient boosting framework used for its efficiency and performance.

LGBMRegressor: Emerged as the best model through evaluations.

These models were first trained on 80% of the dataset and validated on the remaining 20%. Subsequently, a K-fold cross-validation (k=5) was employed, where LGBMRegressor showed superior performance. The final model was fine-tuned using a random grid search to optimize the LGBMRegressor's parameters.

3 Results

The LGBMRegressor, after optimization, demonstrated significant efficacy in predicting heart rates. The model achieved an average Mean Squared Error (MSE) of 0.1310 in 5-fold cross-validation, indicating a high level of accuracy and reliability. These results confirm the effectiveness of the feature extraction and model optimization strategies employed in this study. We can clearly see that the LGBMRegressor model gave lowest MSE in below figures. Final accuracy[1] is **99.8815%**.

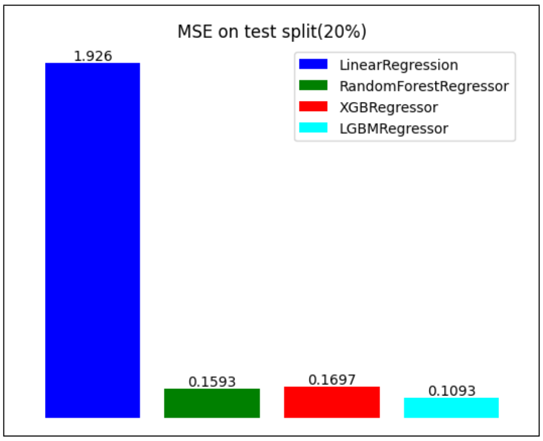

Figure 1: Result on test split (20%)

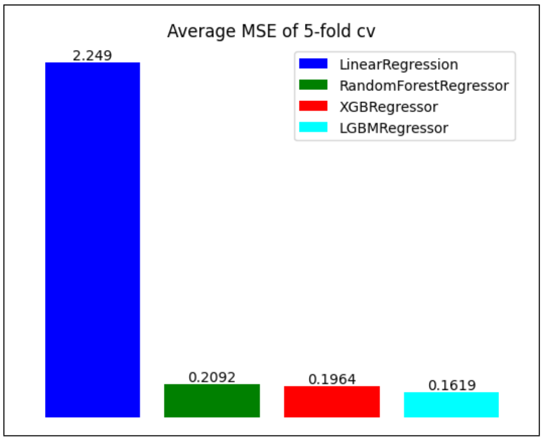

Figure 2: Results of 5-Fold Cross Validation

Metric	Value
Mean HR	74.00
Average MSE	0.1310
R^2 Score	0.9988
Explained Variance	0.9988
Model Accuracy[1]	**99.8815%**

Table 2: Performance Metrics for Optimized LGBMRegressor Model

4 Further Considerations

While the current model shows promising results, further improvements are conceivable with the availability of more data. Neural Network (NN) models, known for their capability in handling large and complex datasets, could potentially enhance the prediction accuracy. However, given the current data constraints, NN models are not immediately applicable. We could have done more focus on NNs if we had more data. Additionally, we could attempt to increase the number of iterations in the random grid search or expand the search area. However, this would likely only lead to longer training times without resulting in significant improvements in the model's performance.

[1]Model Accuracy = Explained Variance * 100