

Model Optimization and Tuning Phase Template

Date	7th July 2024
Team ID	739719
Project Title	Garment Workers Productivity Predictions
Maximum Marks	10 Marks

Model Optimization and Tuning Phase

Optimized models using GridSearchCV, selecting XGBoost Regressor for its superior performance. Trained and validated the final model on the full dataset, ensuring accurate productivity predictions.

Hyperparameter Tuning Documentation (8 Marks):

Model	Tuned Hyperparameters
<p>Model 1</p> <p>Linear Regression</p>	<p>Hyperparameters: fit_intercept: Whether to calculate the intercept for this model (True or False). Normalize: Whether to normalize the input variables (True or False)</p> <pre> from sklearn.linear_model import LinearRegression linreg = LinearRegression() linreg.fit(x_train,y_train) pred_linreg = linreg.predict(x_test) print("MAE :", mean_absolute_error(y_test, pred_linreg)) print("MSE :", mean_squared_error(y_test, pred_linreg)) print("RMSE :",sqrt(mean_squared_error(y_test, pred_linreg))) print("MAPE :",mean_absolute_percentage_error(y_test, pred_linreg)) </pre> <p>MAE : 0.09889801648754916 MSE : 0.018535367319974336 RMSE : 0.13614465586270486 MAPE : 0.16747657653644724</p>

<p>Model 2</p> <p>Random Forest</p>	<p>Hyperparameters:</p> <p>n_estimators:The number of trees in the forest.</p> <p>max_depth:The maximum depth of the tree.</p> <p>min_samples_split:The minimum number of samples rerquired to split an internal node.</p> <pre> from sklearn.ensemble import RandomForestRegressor randf = RandomForestRegressor(random_state=42) randf.fit(x_train,y_train) pred_randf = randf.predict(x_test) print("MAE :", mean_absolute_error(y_test, pred_randf)) print("MSE :", mean_squared_error(y_test, pred_randf)) print("RMSE :",sqrt(mean_squared_error(y_test, pred_randf))) print("MAPE :",mean_absolute_percentage_error(y_test, pred_randf)) </pre> <p>Python</p> <p>MAE : 0.08366785595438364 MSE : 0.015441874867015823 RMSE : 0.12426534057015183 MAPE : 0.14067390864389964</p>
-------------------------------------	--

Final Model Selection Justification (2 Marks):

Final Model	Reasoning
<p>Model 1</p> <p>XGBoost Regressor</p>	<p>The XGBoost Regressor was chosen as the final optimized model because it consistently delivered superior performance metrics during the evaluation phase. Its advanced capabilities in handling complex and high-dimensional data, along with its built-in regularization mechanisms, made it the best candidate for accurately predicting garment workers' productivity.</p>