**Model Optimization and Tuning Phase Template**

|  |  |
| --- | --- |
| Date | 15 March 2024 |
| Team ID | xxxxxx |
| Project Title | Forecasting Economic Prosperity: Leveraging Machine Learning For GDP Per Capita Prediction |
| Maximum Marks | 10 Marks |

**Model Optimization and Tuning Phase**

The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

### Hyperparameter Tuning Documentation (6 Marks):

|  |  |  |
| --- | --- | --- |
| **Model** | **Tuned Hyperparameters** | **Optimal Values** |
| Linear Regression | N/A | N/A |
| Random Forest | n\_estimators | n\_estimators=100 |
| Support Vector Regression | kernel | kernel='rbf' |

### Performance Metrics Comparison Report (2 Marks):

|  |  |  |
| --- | --- | --- |
| **Model** | **Baseline Metric** | **Optimized Metric** |
| Linear Regression | MSE: 21883733.766837504,  R^2: -0.9441781713294894 | MSE: 21883733.766837504,  R^2: -0.9441781713294894 |
| Random Forest | MSE: 7133952.090909091,  R^2: 0.36621080852875953 | MSE: 7133952.090909091,  R^2: 0.36621080852875953 |
| Support Vector Regression | MSE: 15330417.613966553,  R^2: -0.3619733999691279 | MSE: 15330417.613966553,  R^2: -0.3619733999691279 |

### Final Model Selection Justification (2 Marks):

|  |  |
| --- | --- |
| **Final Model** | **Reasoning** |
| Random Forest | The Random Forest model was chosen as the final optimized model because it provided the lowest Mean Squared Error (MSE) and highest R^2 score. This indicates that the model performs better at predicting the target variable compared to Linear Regression and Support Vector Regression. Additionally, Random Forests are robust to overfitting and handle both numerical and categorical data well, making them suitable for the "Countries of the World" dataset. |