Edge of Tomorrow : Enhancing Efficiency through Model Integration for ENSO Index Prediction

Sonal Singh [1] Subramanyam Sahoo [2] Joel Jojo [1]

[1]Liverpool John Moores University [2]NIT Hamirpur

Introduction

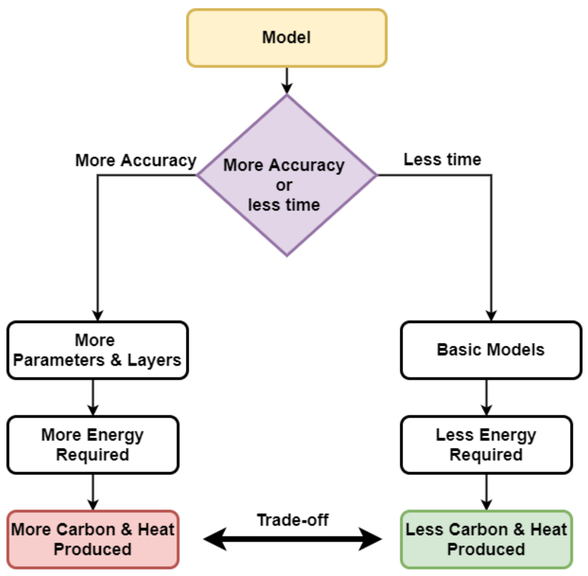

Figure 1. Accuracy-Time Trade off

Ensemble Architecture

- **Efficiency Testing** for the hypothesis was carried out using the following architecture

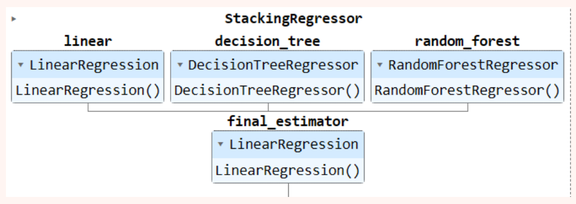

Figure 2. Ensemble Architecture

The experiment aims to accurately forecast the Nino 3.4 Index, which represents the condition of the El Niño Southern Oscillation (ENSO), using **historical Sea Surface Temperatures**.

Results

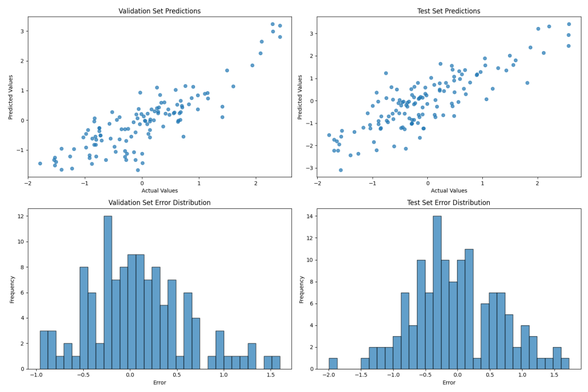

Figure 3. Linear regressor Data distribution

Metric	Validation Set	Test Set
Mean Absolute Error	0.406	0.544
Mean Squared Error	0.276	0.465
Root Mean Squared Error	0.525	0.682
R^2 Score	0.663	0.468
Explained Variance Score		0.469

Table 1. Linear Regressor Evaluation Metrics

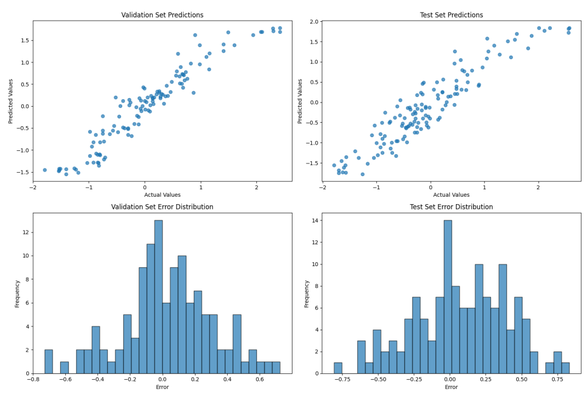

Figure 4. Stacked model data distribution

Metric	Validation Set	Test Set
Mean Absolute Error	0.214	0.274
Mean Squared Error	0.078	0.109
Root Mean Squared Error	0.269	0.329
R^2 Score	0.905	0.876
Explained Variance Score		0.876

Table 2. stacked Evaluation Metrics

Metric	Test Set
Mean Squared Error	0.119
Root Mean Squared Error	0.345
R^2 Score	0.864
Explained Variance Score	0.864

Table 3. PSO-based evaluation metrics

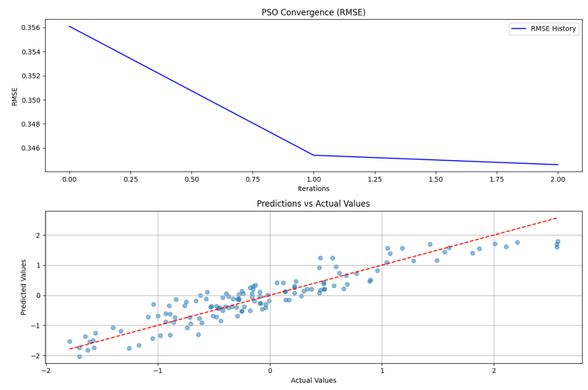

Figure 5. PSO performance

Comparison

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|\theta_i - \hat{\theta}_i\right| \tag{1}$$

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(\theta_i - \hat{\theta}_i\right)^2 \tag{2}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\theta_i - \hat{\theta}_i\right)^2} \tag{3}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(\theta_i - \hat{\theta}_i\right)^2}{\sum_{i=1}^{n}\left(\theta_i - \bar{\theta}\right)^2} \tag{4}$$

Where:

- θ_i represents the actual values ,
- $\hat{\theta}_i$ represents the predicted values,
- $\bar{\theta}$ represents the mean of the actual values,
- n is the number of observations.

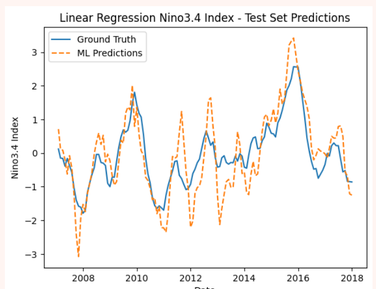

Figure 6. Linear Regressor Output

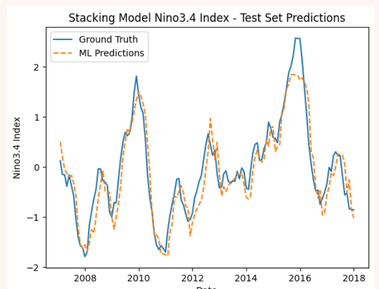

Figure 7. Stacked Regressor Output

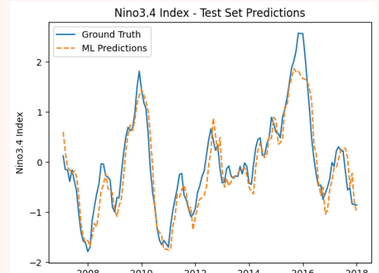

Figure 8. PSO based Output

Metric	Linear Regression	Stacked Regression	CuML Stacked Regression with PSO
Root Mean Squared Error	0.6818	0.3296	0.3446
R-squared Score	0.4682	0.8757	0.8641
Fit Time in Seconds	4.6889	1827.1074	6956.9652

Table 4. Comparison of Model Performance on Test Set

Conclusion

The stacked models deliver strong accuracy, while the simple stacking and linear regression models are faster.