ENSOcast: A Machine Learning Application for El Niño-Southern Oscillation Phase Classification

Dylan Dsouza
Halıcıoğlu Data Science Institute
University of California San Diego
San Diego, California, USA
dydsouza@ucsd.edu

***Abstract*—The El Niño–Southern Oscillation (ENSO) is a major driver of global climate variability, yet existing forecasting models are often inaccessible to non-specialists. This paper presents ENSOcast, a comprehensive machine learning application for classifying ENSO phases using aggregated climate data. The platform integrates Sea Surface Temperature (SST) anomalies, Southern Oscillation Index (SOI), and temporal features to predict monthly ENSO states (El Niño, La Niña, or Neutral). ENSOcast employs multiple algorithms on historical climate data spanning 1982-2024 including Random Forest Classifiers, eXtreme Gradient Boosting (XGBoost), One-Dimensional Convolutional Neural Networks (1-D CNNs), Long Short-Term Memory (LSTM) networks, and Ensemble Learning. The baseline Random Forest classifier achieves 82% testing accuracy, demonstrating the effectiveness of engineered climate features for ENSO phase prediction. Advanced models show comparable performance, with ensemble methods providing robust predictions through complementary algorithm strengths. The interactive web application enables climate researchers to explore ENSO patterns, customize model training parameters, and visualize global sea surface temperature dynamics.**

***Index Terms*—El Niño–Southern Oscillation (ENSO), climate prediction, machine learning, interactive visualization, exploratory analysis**

I. Introduction

The El Niño-Southern Oscillation (ENSO) represents one of the most significant climate phenomena affecting global weather patterns, agricultural productivity, and economic stability. ENSO operates through coupled ocean-atmosphere interactions in the Pacific, manifesting in three distinct phases:

- **El Niño:** Rainfall increases in the eastern Pacific region, the ocean warms up, and trade winds weaken. This often results in destructive floods in California and droughts from Australia to South Asia.
- **La Niña:** Pacific Ocean cools down and trade winds strengthen, which typically causes droughts in South America and leads to the formation of intense hurricanes in the Atlantic region.
- **Neutral:** Ocean temperatures remain near average, and weather patterns stay normal. This brings about typical seasonal weather worldwide.

Traditional ENSO monitoring relies on oceanic and atmospheric indices, primarily the Oceanic Niño Index (ONI) and Southern Oscillation Index (SOI). While these indices provide reliable phase identification, machine learning approaches offer potential advantages in pattern recognition, multivariate analysis, and automated classification systems.

Advances in climate prediction have demonstrated the effectiveness of machine learning methods, such as deep learning, for atmospheric and oceanic prediction tasks. However, existing ENSO prediction systems often focus on forecasting future states rather than robust classification of current conditions using multiple climate indicators.

This paper introduces ENSOcast, a machine learning application designed to classify ENSO phases using engineered climate features. ENSOcast addresses several key challenges, including handling irregular temporal patterns in climate data, integrating multiple climate indicators with different temporal characteristics, and providing intuitive and easily interpretable results for climate science applications.

II. Methodology

A. Data Sources and Preprocessing

ENSOcast utilizes multiple climate datasets, courteousy of the National Oceanic and Atmospheric Administration (NOAA), spanning January 1982 to December 2024:

Sea Surface Temperature (SST): Monthly SST data from the Niño 3.4 region (5°N-5°S, 170°W-120°W) obtained from NOAA's OI SST V2 High Resolution Dataset (provided by the NOAA Physical Sciences Laboratory, Boulder, Colorado, USA). SST anomalies are calculated by subtracting the monthly climatology (mean SST for each month across all years) from the observed SST, i.e.,

$$\text{SST Anomaly} = \text{SST}_{\text{observed}} - \text{SST}_{\text{climatology}}.$$

Southern Oscillation Index (SOI): Monthly SOI values representing standardized sea level pressure differences between Tahiti and Darwin, Australia, obtained from the Climatic Research Unit (CRU), University of East Anglia, UK.

Oceanic Niño Index (ONI): 3-month running mean of SST anomalies in the Niño 3.4 region, provided by the NOAA Climate Prediction Center (CPC), based on ERSST.v5 SST anomalies. These serve as the ground truth for ENSO phase classification:

- **El Niño:** ONI $\geq$ +0.5°C for 5 consecutive overlapping seasons
- **La Niña:** ONI $\leq$ -0.5°C for 5 consecutive overlapping seasons

- **Neutral:** ONI between -0.5°C and +0.5°C

Ocean Heat Content (OHC): Seasonal (3-month) subsurface temperature anomalies integrated over the upper 700m depth, provided by Ocean Heat Content, Salt Content, and Sea Level Anomalies, a joint project between the National Centers for Environmental Information (NCEI) and NOAA Global Ocean Monitoring and Observing (GOMO).

B. Feature Engineering

ENSOcast employs comprehensive feature engineering to capture temporal dependencies and seasonal patterns:

Temporal Lags: SOI and SST anomaly values at 1, 2, and 3-month lags to capture predictive relationships and momentum effects.

Rolling Averages: 3-month and 6-month rolling means of SST and 3-month rolling mean of SOI to smooth high-frequency variability and identify medium-term trends.

Trend Features: First-order differences of SST and SOI to capture rate-of-change information indicative of phase transitions.

Seasonal Encoding: Trigonometric transformations of calendar months using sine and cosine functions to capture annual cycles without introducing artificial ordinality:

$$month_{sin} = \sin\left(\frac{2\pi \cdot month}{12}\right) \tag{1}$$

$$month_{cos} = \cos\left(\frac{2\pi \cdot month}{12}\right) \tag{2}$$

Interaction Terms: Cross-products between SST anomalies and SOI to capture coupled ocean-atmosphere effects.

Seasonal Indicators: Binary features for December-January-February (DJF) and June-July-August (JJA) seasons when ENSO signals are typically strongest.

The final feature vector comprises 18 engineered variables designed to capture the multi-scale temporal dynamics of ENSO variability.

C. Machine Learning Models

ENSOcast leverages six machine learning approaches, including a baseline model and five advanced models utilizing the full set of engineered features:

Baseline Random Forest: A reference model using only seasonal encoding and temporal lag features. Implemented with 100 estimators and default parameters, it captures basic temporal patterns while providing interpretability through feature importance metrics.

Enhanced Random Forest: Random Forest classifier using all engineered features, configured with 300 estimators, maximum depth of 20, and bootstrap aggregation. Trained on the full feature set to model non-linear interactions and evaluated using accuracy on a held-out test set.

eXtreme Gradient Boosting (XGBoost): Gradient boosting classifier optimized for tabular data, with 300 estimators, maximum depth of 6, and learning rate of 0.1. Regularization is applied to reduce overfitting, and the model is trained on scaled features with predictions evaluated on the test set.

One-Dimensional Convolutional Neural Network (1D CNN): Deep learning architecture capturing local temporal patterns via Conv1D layers. The network consists of 64- and 32-filter Conv1D layers, batch normalization, global max pooling, dense layers with 50 units, and dropout regularization (0.3–0.5). Input sequences are dynamically generated based on available data, and the model is trained using sparse categorical cross-entropy with Adam optimizer, early stopping, and learning rate reduction callbacks.

Long Short-Term Memory Network (LSTM): Recurrent network designed for long-range temporal dependencies. Features stacked LSTM layers with 50 units each, dropout layers (0.3), and dense layers leading to a softmax output over ENSO phases. Input sequences are generated similarly to the CNN, and training uses Adam optimizer with early stopping and learning rate reduction.

Ensemble Learning: Combines predictions from Enhanced Random Forest and XGBoost models via majority voting. Each base model is trained independently on the full feature set, and the final ensemble prediction is computed by taking the statistical mode of the individual predictions, improving robustness and reducing model-specific biases.

D. Validation Strategy

Model validation employs a temporal split to respect the time-series nature of climate data. The earliest 80% of observations (1982–2015) are used as the training set, while the most recent 20% (2016–2024) form the holdout test set. This ensures models are evaluated on genuinely unseen future data, reflecting real-world deployment scenarios.

For deep learning models (1D CNN and LSTM) requiring sequential input, input sequences are dynamically constructed from the available data, with sequence lengths adapted to the dataset size. In our experiments, a typical sequence length of 6 time steps (months) was used, providing sufficient temporal context while maintaining computational efficiency. Sequences overlap to maximize training samples, and input features are scaled prior to sequence generation.

III. Implementation

A. Software Architecture

ENSOcast is implemented as a web-based application using Streamlit for the user interface, with scientific computing libraries including:

- **Data Processing:** Pandas, NumPy, xarray for climate data manipulation
- **Machine Learning:** Scikit-learn, XGBoost, TensorFlow/Keras for model implementation
- **Visualization:** Plotly, Matplotlib for interactive climate data exploration
- **Statistical Analysis:** SciPy for performance evaluation and significance testing

The application architecture supports modular model training, real-time data ingestion, and interactive parameter adjustment for research applications.

B. Interactive Features

The web interface provides several research-oriented capabilities:

Historical Analysis: Interactive exploration of ENSO patterns from 1982–2024, with customizable time ranges and the ability to filter by ENSO phase.

Model Training: Configurable training parameters, including selectable time periods, target ENSO phases, and choice of machine learning algorithm.

Performance Evaluation: Comprehensive assessment of model performance using confusion matrices, classification reports, and feature importance metrics.

Data Visualization: Global sea surface temperature (SST) maps with emphasis on the Niño 3.4 region, accompanied by temporal analyses of SST, SOI, and ONI to identify trends, anomalies, and seasonal patterns.

IV. Results and Analysis

A. Baseline Random Forest Model Performance

The Baseline Random Forest classifier achieves 82% overall accuracy on the temporal holdout test set, with class-specific performance shown in Table I.

TABLE I
BASELINE RANDOM FOREST MODEL PERFORMANCE

ENSO Phase	Precision	Recall	F1-Score	Support
El Niño	0.83	0.86	0.84	28
Neutral	0.72	0.86	0.78	36
La Niña	0.94	0.75	0.83	40
Macro Avg	**0.83**	**0.82**	**0.82**	**104**

Feature importance rankings from the Baseline Random Forest model identify the most predictive variables to be **SST_Anomaly_lag_2** (25.3%), **SST_Anomaly_lag_3** (19.6%), and **SST_Anomaly_lag_1** (19.4%).

B. Enhanced Random Forest Model Performance

The Enhanced Random Forest classifier achieves 80% overall accuracy on the temporal holdout test set, with class-specific performance shown in Table II.

TABLE II
ENHANCED RANDOM FOREST MODEL PERFORMANCE

ENSO Phase	Precision	Recall	F1-Score	Support
El Niño	0.75	0.86	0.80	28
Neutral	0.72	0.81	0.76	36
La Niña	0.94	0.75	0.83	40
Macro Avg	**0.80**	**0.81**	**0.80**	**104**

Feature importance rankings from the Enhanced Random Forest model identify the most predictive variables to be **SST_Anomaly_rolling_6m** (17.1%), **SST_Anomaly_lag_2** (15.9%), and **SST_Anomaly_lag_1** (12.4%).

C. XGBoost Model Performance

The XGBoost model achieves 83% overall accuracy on the temporal holdout test set, with class-specific performance shown in Table III.

TABLE III
XGBOOST MODEL PERFORMANCE

ENSO Phase	Precision	Recall	F1-Score	Support
El Niño	0.81	0.93	0.87	28
Neutral	0.75	0.83	0.79	36
La Niña	0.94	0.75	0.83	40
Macro Avg	**0.83**	**0.84**	**0.83**	**104**

Feature importance rankings from the XGBoost model identify the most predictive variables to be **SST_Anomaly_rolling_6m** (30.5%), **SST_Anomaly_lag_2** (19.5%), and **SST_SOI_interaction** (9.2%).

D. 1D CNN Model Performance

The 1D CNN model achieves 86% overall accuracy on the temporal holdout test set, with class-specific performance shown in Table IV.

TABLE IV
1D CNN MODEL PERFORMANCE

ENSO Phase	Precision	Recall	F1-Score	Support
El Niño	0.91	0.77	0.83	26
Neutral	0.78	0.94	0.85	33
La Niña	0.92	0.85	0.88	39
Macro Avg	**0.87**	**0.85**	**0.85**	**98**

Due to the black-box nature of deep learning models, feature importance was not explicitly calculated, as it is less straightforward compared to traditional machine learning methods. Moreover, the reported results may vary slightly across runs because of the inherent stochasticity in deep learning training processes.

E. LSTM Model Performance

The LSTM model achieves 82% overall accuracy on the temporal holdout test set, with class-specific performance shown in Table V.

TABLE V
LSTM MODEL PERFORMANCE

ENSO Phase	Precision	Recall	F1-Score	Support
El Niño	0.88	0.81	0.84	26
Neutral	0.79	0.79	0.79	33
La Niña	0.80	0.85	0.82	39
Macro Avg	**0.82**	**0.82**	**0.82**	**98**

Due to the black-box nature of deep learning models, feature importance was not explicitly calculated, as it is less straightforward compared to traditional machine learning methods. Moreover, the reported results may vary slightly across runs because of the inherent stochasticity in deep learning training processes.

F. Ensemble Learning Performance

The Ensemble Learning model achieves 81% overall accuracy on the temporal holdout test set, with class-specific performance shown in Table VI.

TABLE VI
ENSEMBLE LEARNING MODEL PERFORMANCE

ENSO Phase	Precision	Recall	F1-Score	Support
El Niño	0.80	0.86	0.83	28
Neutral	0.73	0.83	0.78	36
La Niña	0.91	0.75	0.82	40
Macro Avg	**0.81**	**0.81**	**0.81**	**104**

Feature importance rankings from the Ensemble Learning model identify the most predictive variables to be **SST_Anomaly_rolling_6m** (24.1%), **SST_Anomaly_lag_2** (18.1%), and **SST_SOI_interaction** (8.6%).

G. Advanced Model Peformance Comparison

Comparative evaluation across all algorithms reveals consistent performance as shown in Table VII.

TABLE VII
PERFORMANCE OF ADVANCED MACHINE LEARNING MODELS

Model	Precision	Recall	F1-Score	Accuracy
Enhanced RF	0.80	0.81	0.80	0.80
XGBoost	0.83	0.84	0.83	0.83
1D CNN	0.87	0.85	0.85	0.86
LSTM	0.82	0.82	0.82	0.82
Ensemble	0.81	0.81	0.81	0.81

V. DISCUSSION

The comparative results demonstrate that all machine learning models are capable of effectively classifying ENSO phases from aggregated climate indicators. Among the advanced models, the 1D CNN achieved the highest overall accuracy (86%), followed closely by XGBoost (83%). The ensemble approach, while robust, did not outperform the best individual models, suggesting that model complementarities may be limited in this classification task.

The dominance of SST-related features across all tree-based models underscores the critical role of ocean surface temperature variability in ENSO dynamics. Interaction terms between SST and SOI further highlight the coupled nature of the ocean–atmosphere system, validating the inclusion of engineered features that capture cross-variable effects.

However, deep learning models introduced additional complexity without providing interpretable feature contributions. This trade-off between accuracy and interpretability is an important consideration for applications in climate science, where transparency can be as critical as predictive power. Moreover, the stochastic nature of neural networks may result in slight variability across runs, though overall performance remained consistent.

VI. CONCLUSION

This study introduced ENSOcast, a comprehensive machine learning application for ENSO phase classification. By integrating multiple climate datasets and employing feature engineering techniques, ENSOcast achieved robust classification performance across multiple model architectures. Results indicate that both traditional ensemble methods (Random Forest, XGBoost) and deep learning architectures (1D CNN, LSTM) can provide accurate classification of ENSO phases, with CNNs achieving the highest accuracy.

ENSOcast contributes to the climate science community by providing a reproducible, multi-model framework for ENSO classification, an interactive web platform for climate data visualization and model exploration, and a demonstration of how data-driven methods can support and enhance traditional ENSO monitoring approaches.

ACKNOWLEDGMENT

The author thanks the National Oceanic and Atmospheric Administration (NOAA) for providing the SST and ONI datasets, the Climatic Research Unit (CRU) at the University of East Anglia, UK for SOI data, and the National Centers for Environmental Information (NCEI) and NOAA Global Ocean Monitoring and Observing (GOMO) for making atmospheric and subsurface ocean data publicly available.

REFERENCES

[1] Y.-G. Ham, J.-H. Kim, and J.-J. Luo, "Deep learning for multi-year ENSO forecasts," *Nature*, vol. 573, no. 7775, pp. 568–572, 2019.

[2] Y.-G. Ham, J.-H. Kim, E.-S. Kim, and K.-W. On, "Unified deep learning model for El Niño/Southern Oscillation forecasts by incorporating seasonality in climate data," *Science Bulletin*, vol. 66, no. 13, pp. 1358–1366, 2021.

[3] Z. Zhang, L. Wang, and X. Liu, "Combined dynamical-deep learning ENSO forecasts," *Nature Communications*, vol. 16, art. 3456, 2025.

[4] R. W. Reynolds, T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, "Daily high-resolution-blended analyses for sea surface temperature," *Journal of Climate*, vol. 20, no. 22, pp. 5473–5496, 2007.

[5] B. Huang, P. W. Thorne, V. F. Banzon, T. Boyer, G. Chepurin, J. H. Lawrimore, M. J. Menne, T. M. Smith, R. S. Vose, and H.-M. Zhang, "Extended reconstructed sea surface temperature, version 5 (ERSSTv5): upgrades, validations, and intercomparisons," *Journal of Climate*, vol. 30, no. 20, pp. 8179–8205, 2017.

[6] T. G. Dietterich, "Ensemble methods in machine learning," in *Multiple Classifier Systems*, J. Kittler and F. Roli, Eds. Berlin, Germany: Springer, 2000, pp. 1–15.