UNIVERSITI TEKNOLOGI MARA

DELIENA TASHA BINTI ABDUL RAHIM

Coral Reefs Bleaching Prediction

SUPERVISED BY SITI NUR KAMALIAH KAMARUDIN

Source: The Ocean Agency / Ocean Image Bank – *Coral bleaching in the Maldives, 2016*

According to NASA, 70 to 90% of coral reefs may disappear by 2050

Coral reefs are increasingly threatened by climate change. If preventative measures and conservation efforts are not prioritized, we may live in a world without coral reefs in our lifetime.

Introduction

What is coral bleaching?

Coral bleaching is a stress response to rising sea temperature (SST) where they begin to expel a symbiotic algae that gives them colors. Without this algae, corals are more susceptible to diseases and eventually die.

Source: The Ocean Agency – *Coral Bleaching, Lizard Island, Great Barrier Reef, Before (March 2016) & After (May 2016)*

- Home to 25% to marine life
- Support tourism, fisheries, medicine, and shoreline protection
- One degree increase in temperature for four weeks can initiate bleaching and in more than eight weeks, corals are unable to recover.

Problem Statement

The Loss of Coral Reefs Affect Us All

Impact on Corals

During the most intense and severe bleaching event, 84% coral reefs around the world were effected.

Economic Losses

International Coral Reef Initiative estimate a loss of $500 billion annually by 2100.

Traditional Monitoring

The methods to monitor coral reefs are short-term and reactive rather than preventative.

Scope

- **Dataset:** Global Bleaching and Environmental Data (BCO-DMO) - 41,361 instances, 63 features
- **Models:** Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT), Multilayer Perceptron (MLP), Long Short-Term Memory (LSTM), Extreme Gradient Boosting (XGBoost)
- **Evaluation Metrics:** Accuracy, MAE, RMSE, and MAPE

Significance

Supports biodiversity, climate resilience, and socio-economic stability (tourism/fisheries)

Predicting bleaching events enables early interventions

Contributes to United Nations Sustainable Development Goal 14: Life Below Water

"Conserve and sustainably use the oceans, seas and marine resources for sustainable development."

Research Questions and Objectives

Which environmental factors are most indicative of coral bleaching events?

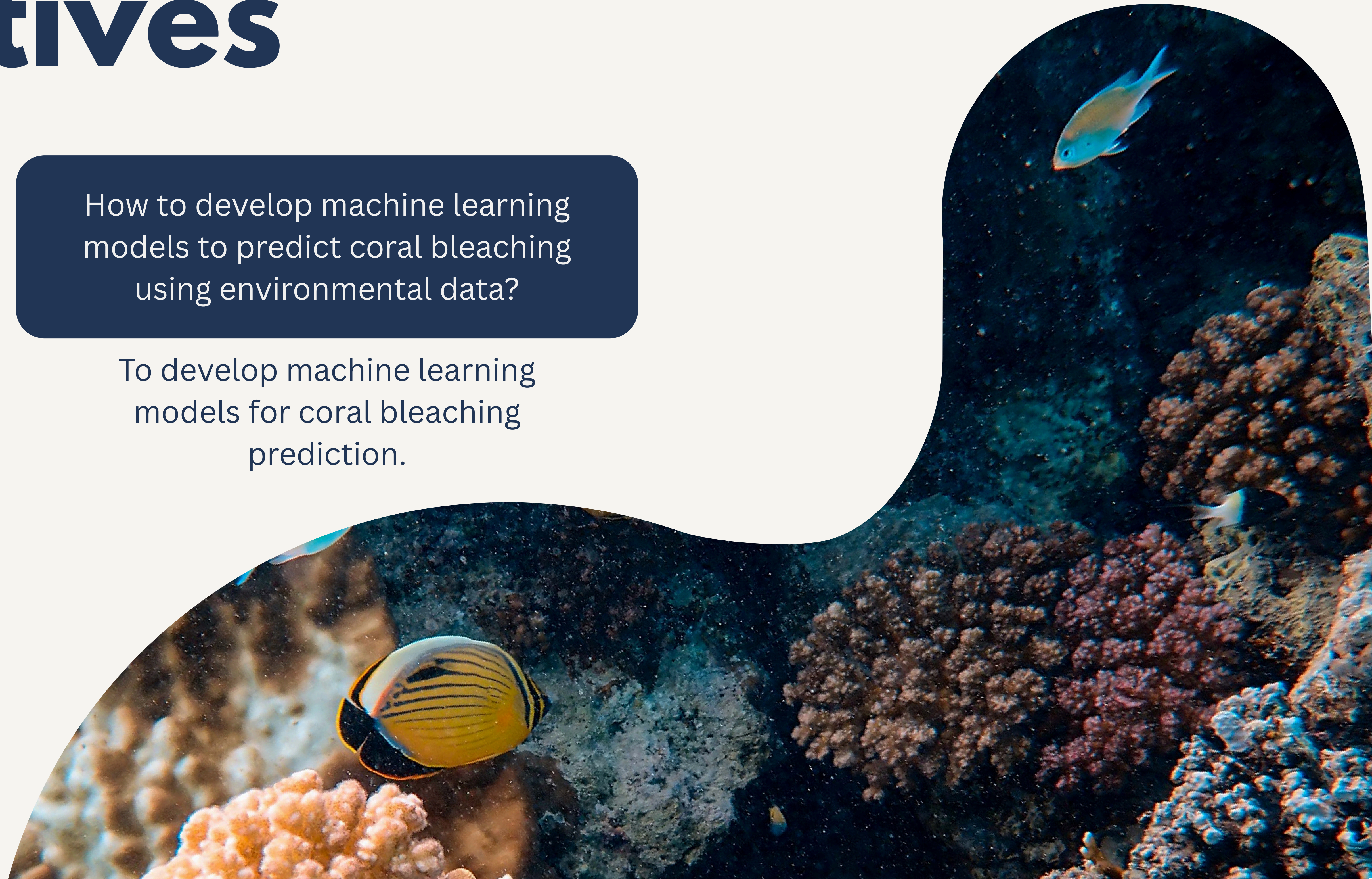
To identify key environmental factors that influence coral bleaching.

How to develop machine learning models to predict coral bleaching using environmental data?

To develop machine learning models for coral bleaching prediction.

How to evaluate which machine learning models have the best performance for coral bleaching prediction?

To evaluate machine learning models for coral bleaching prediction using various evaluation metrics.



Authors	Title	Techniques	Findings
(Boonnam et al., 2022)	Coral Reef Bleaching under Climate Change: Prediction Modeling and Machine Learning	• Naïve Bayes, SVM, Decision Tree and K-Means Clustering • Metrics: Accuracy, Kappa Coefficient • Statistical dataset	• SVM highest accuracy (88.85%) • Sea surface temperature and pH were strongly correlated with bleaching.
(Udomchaipitak et al., 2022)	Forecast Coral Bleaching by Machine Learnings of Remotely Sensed Geospatial Data	• MLP, RF, DT, RBF-ANN • Metrics: Accuracy, 10- fold cross validation, RMSE and Kappa • Spatial dataset	• RF achieved 97.25% highest accuracy • MLP 88.07% accuracy • DT 94.49% accuracy • RBF 90.83% accuracy
(Lin et al., 2023)	Applying deep learning to predict SST variation and tropical cyclone patterns that influence coral bleaching	• CNN-LSTM, RF, ARIMA, SVM, RNN, LSTM, and GRU • Metrics: Accuracy, MAE, RMSE, MAPE	• ConvLSTM highest accuracy 99.22% • RF 92.90% accuracy • SST variations correlate with increased coral bleaching risk.

DELIENA TASHA BINTI ABDUL RAHIM

FYP | MAC-AUG 2025

Methodology

*Tools: Mendeley, Scribbr, Weka, Python, Google Colab, Google Documents

Prelimenary Study

Identify background, problem, objectives, scope and significance of the study

Knowledge Acquisition

Literature review of existing related studies.

Data Acquisition

Retrieved from Biological and Chemical Oceanography Data Management Office

Data Preprocessing

Clean dataset, feature engineering, handle missing data and, unnecessary features with Weka and Python

Model Development

Develop LSTM, SVM, RF, DT, XGboost, and MLP models using Python

Model Evaluation

Evaluate model performance using Accuracy, MAE, RMSE, and MAPE

Documentation

Prepare document containing all the processes of this project

Experiments

Data Preprocessing and Feature Selection

Experiment One
Find the top 15 features that influence the model's prediction

Top 5 features were:
1. Turbidity
2. SSTA_Standard_Deviation
3. Temperature_Mean
4. Distance_to_Shore
5. Cyclone_Frequency.

***The rest were temperature related features.**

Experiment Two
Determine the impact of data preprocessing on model's performance.

No significant difference between before and after data preprocessing

Experiment Three
The comparison of using only top 15 features verses all relevant features.

Using all relevant features improved the results of all model in comparison to the results in experiment One and Two.

The best results of Experiment One, Two and Three

Model	Experiment	Accuracy	MAE	RMSE	MAPE
RF	Results with All Relevant Features	99.21%	0.008	0.089	0.79%
SVM		82.61%	0.174	0.417	17.39%
LSTM		86.98%	0.130	0.360	13.04%
XGboost		97.96%	0.020	0.143	2.04%
DT		98.67%	0.013	0.115	1.33%
MLP		92.57%	0.074	0.272	7.43%

Experiments

Hyperparameter Tuning

The best parameters for each models

Model	Total Experiments Conducted	Best Hyperparameters	Worst Hyperparameters
RF	24	Max Depth: 30 Max Features: log2 N Estimators: 150	Max Depth: 10 Max Features: sqrt N Estimators: 150
SVM	24	C: 10 Kernel: Poly Gamma: Scale Degree: 3	C: 10 Kernel: Sigmoid Gamma: Scale Degree: 3
LSTM	24	Units: 128 Dropout: 0.2 Batch Size: 32 Optimizer: Adam	Units: 64 Dropout: 0.5 Batch Size: 64 Optimizer: RMSprop
XGboost	24	N Estimators: 400 Max Depth: 10 Learning Rate: 0.3	N Estimators: 100 Max Depth: 3 Learning Rate: 0.1
DT	24	Criterion: Gini Max Depth: No limit Min Samples Split: 2	Criterion: Entropy Max Depth: 10 Min Samples Split: 10
MLP	24	Activation: relu Hidden Layers: (128, 64, 32) Learning Rate: 0.001	Activation: tanh Hidden Layers: 64 Learning Rate: 0.001

= 144 Experiments

*30 Epochs

Purpose

To improve the performance of each models

Results

Tuning improved model accuracy and reduced error metrics.

Results

The best results of from all experiments for each models

Model	Accuracy	MAE	RMSE	MAPE
Random Forest	99.22%	0.008	0.088	0.78%
XGBoost	99.28%	0.007	0.085	0.72%
Decision Tree	98.67%	0.013	0.115	1.33%
MLP	96.45%	0.036	0.189	3.55%
LSTM	86.30%	0.137	0.370	49.35%
SVM	87.69%	0.123	0.351	12.31%

Discussion

Six machine learning models were evaluated and compared after conducting experiments. The best results were organized into this final results table. Overall, all models showed > 80% accuracy.

Top Peformers

- XGBoost: 99.25% Accuracy and MAPE < 1%
- Random Forest: 99.22% Accuracy and MAPE < 1%
- Ensemble models (RF and XGBoost) outperform individual learners
- Decision Tree also showed high accuracy at 98.67%

Worst Performer

- LSTM had the lowest accuracy and highest error rates
- Possibly due to lack of time series and sequential data.

Data Preprocessing Experiment

- Before and After data preprocessing showed no conclusive results.
- Possibly due to the dataset being well structured and high quality.

Feature Selection Experiment

- Using all relevant features improved the performance of all models.
- Shows that the features interaction with each other effected the performance. When features were reduced some of those interactions were lost.
- This confirms that even moderately important features can contribute to prediction when combined with others.
- Top 15 ranks suggest that more than just temperature influence coral bleaching

Hyperparameter Tuning Experiments

- Improved the results of all models.
- Exception: LSTM had better accuracy but higher error rates than before parameter tuning. Possibly due to using less epoch for these experiments.

The Potential of Machine Learning in Coral Bleaching

This project demonstrates the potential of machine learning as a tool to aid the conservation of coral reefs.

Among the six models tested, ensemble based models outperform others with an almost 100% accuracy and < 1% MAPE errors. The other models still showed promise with all accuracy > 80%. These models successfully predicted coral bleaching using BCO-DMO dataset.

Future works

- Integration with satellite imagery and real-time data
- Explore hybrid deep learning architectures (e.g., CNN-LSTM)
- interactive prototype or dashboard

Source: The Ocean Agency / Ocean Image Bank - *Great Barrier Reef, Australia, 2017 Bleaching*

Conclusion

UNIVERSITI TEKNOLOGI MARA

DELIENA TASHA BINTI ABDUL RAHIM

FYP | MAC-AUG 2025

Thank You

Special thanks to family, lecturers, examiners, friends, BCO-DMO for data and The Ocean Agency for images.

xdeliena

deliena.tasha1@gmail.com

Selangor, Malaysia

Project Supervisor: Siti Nur Kamaliah Kamarudin