Project Proposal : Prediction of the Location and Timing of Coral Bleaching

2491014 박예나, 2491024 이유진, 2391041 김연희, 2391042 이시현, 2391046 김유림

1. Business Value

1.1 Background

This project originated from concerns about marine ecosystem problems caused by serious climate change. Coral bleaching is one of the marine environmental problems where corals become damaged due to rising sea surface temperatures. In recent years, this problem has been increasingly observed around the world, including Thailand, Australia, and even Korea. The worsening of coral bleaching poses a serious threat not only to environments but also to food security and regional economies, thereby highlighting the need for predictive models.

1.2 Problem Definition & Value Proposition

The main goal of this project is to develop a model to predict the bleaching phenomenon based on the analysis of the association between marine environmental changes and coral bleaching. Through this, it aims to raise awareness of coral bleaching and to come up with appropriate countermeasures. The results derived from this model will contribute not only to the preservation of marine ecosystems, but also to the development of solutions to coastal area problems associated with them.

2. Available Data

2.1 Data Acquisition

The dataset used in this project is the “Changing Seas: Marine Climate and Marine Life Dataset” provided by Kaggle. It investigates factors affecting coral bleaching events in relation to marine environments over approximately eight years, from 2015 to 2023. The dataset includes information such as the latitude and longitude of observation stations, sea surface temperature and acidity, bleaching severity, the number of marine species observed during the sampling period, and occurrences of marine heatwaves.

2.2 Data Review

The data review revealed that in cases where coral bleaching did not occur, the bleaching severity values were marked as missing. Therefore, these missing values were filled with the "no bleaching" category, as originally intended in the dataset. Subsequently, Label Encoding was applied to convert

the categories low, medium, high, and no-bleaching into integer values. For the modeling process, the selected features included Date, Latitude, Longitude, Sea Surface Temperature (SST), pH Level, Bleaching Severity, and Marine Heatwave occurrence. The Species Observed feature was excluded from the modeling process.

3. Formulation

3.1 Choice of Algorithms and Rationale

This project aims to predict coral bleaching severity by classifying it into four levels using machine learning techniques. Preliminary visualizations, including scatter plots and polynomial curve fitting, revealed nonlinear trends between input variables and bleaching severity. Therefore, models like Decision Trees and Random Forests were selected to capture these patterns effectively and improve classification performance. Among these, Random Forest is particularly robust against overfitting and can handle feature interactions and variable importance well, making it a strong candidate for modeling ecological phenomena such as coral bleaching.

3.2 Input & Output Variables

In the proposed predictive modeling framework, two main stages are involved: the first is the calculation of Degree Heating Weeks (DHW), and the second is the prediction of coral bleaching severity using a machine learning model. DHW will be calculated from Sea Surface Temperature(SST), date, latitude, and longitude based on established formulas. The inputs to the coral bleaching severity prediction model include the calculated DHW, pH level, marine heatwave occurrences, latitude and longitude. The model's output is the predicted bleaching severity, categorized into four levels.

3.3 Expected Challenges and Alternatives

Coral bleaching typically occurs when corals are exposed to high temperatures over an extended period. SST data show only a fragmentary appearance of temperature variance, therefore, relying solely on instantaneous SST data is insufficient for accurately predicting bleaching severity. To overcome this limitation, we will calculate Degree Heating Weeks (DHW), which accounts for prolonged thermal stress, using Sea Surface Temperature (SST °C), date, latitude, and longitude. The resulting DHW values will then serve as a critical input variable in the subsequent model for predicting coral bleaching severity.