



# **Monarch Butterfly Population Decline: analyzing threats using machine learning**

Samuel Veliveli

# Background/Overview



- **Monarch Butterfly Populations have declined by over 85% over the past two decades according to research done by center for biological diversity. Overall, the migrating populations are less than half the size they need to be to avoid extinction**
- **This project aims to analyze a dataset that lists factors affecting Monarch populations. The Project employed 3 different models: Random Forest Regressor, XGBoost Regressor, and Gradient Boost Regressor**
- **The Primary Goal of the project was to see if the models showed similarity with regards to feature importance and see which model had the most accurate prediction with regards to population growth. Feature importance would then be compared to known biological information to see the relevance/accuracy in real life applications**

# Dataset Background



- Dataset contained information regarding threats to the Monarch Population: factors included information regarding climatic conditions, habitat loss, and insecticide use
- Researchers have done similar projects using partial least-squares regressions and time-series analysis
- For our target variable that the models aimed to predict: we chose population growth
- Link to dataset: <https://royalsocietypublishing.org/doi/10.1098/rsos.170760>

# Random Forest Regressor



Accuracy Metrics for Random Forest Regressor:

Mean Squared Error: 0.22983168764000025

R-squared Score: -2.184192285951478

Feature Importance:

OWmean_max_wind	0.470579
-----------------	----------

MeanJ_NE	0.206726
----------	----------

Tempp4avg_NC	0.150339
--------------	----------

OWtotal_precip	0.101337
----------------	----------

T70p3sum_NE	0.071018
-------------	----------

dtype: float64

# XGBoost Regressor



```
Accuracy Metrics for XGBoost Regressor:  
Mean Squared Error: 0.2788547095766095  
R-squared Score: -2.8633794332394027
```

```
Feature Importance:
```

OWtotal_precip	0.467578
OWmean_max_wind	0.327523
T70p3sum_NE	0.142598
MeanJ_NE	0.041419
Tempp4avg_NC	0.020883

```
dtype: float32
```

# Gradient Boost Regressor



Accuracy Metrics for Gradient Boosting Regressor:

Mean Squared Error: 0.3312229997025212

R-squared Score: -3.5889134410155155

Feature Importance:

OWmean_max_wind	0.593446
-----------------	----------

OWtotal_precip	0.194864
----------------	----------

T70p3sum_NE	0.093332
-------------	----------

MeanJ_NE	0.091195
----------	----------

Tempp4avg_NC	0.027163
--------------	----------

dtype: float64

# Overall Conclusion



- **Random Forest Regressor performed the best out of all 3 models followed by XGBoost Regressor**
- **Max Wind Gust and total precipitation were correlated heavily with population growth. This is logical due to the Weather influence on**
  - **Migration Patterns**
  - **Resource availability**
  - **Breeding and Reproduction**
- **This project proposes that more research should be done on climatic factors and their effects on the butterfly population**

# Application



- **By utilizing 3 different models, this project was able to find a correlation between climatic factors and population growth for the monarch butterfly**
- **In addition to other factors, this project proposes that more research should be done to how climatic factors affect the population's migration route, lifestyle, and other factors that could influence population growth**
- **The author of this project hopes that similar methods will be applied to other species facing population decline**