

PROJECT DELIVERABLES



Annual Fall Foliage Prediction Machine Learning Model

By:

Khanh Van - kxv230013

Spring 2025

University of Texas at Dallas

Supervised By:

Professor Brian Wescott Ricks

Professor of Instruction - Computer Science

University of Texas at Dallas

TABLE OF CONTENTS

Executive Summary.....	2
Project Overview.....	2
Objectives.....	2
Introduction.....	3
Background & Motivation.....	3
Problem Statement.....	3
Data Collection Process.....	4
Source of Data.....	4
Archived Fall Foliage Data.....	4
Data Selection.....	4
Data Pre-Processing.....	5
Archived Fall Foliage Data.....	6
Data Selection.....	6

Executive Summary

Project Overview

The Annual Fall Foliage Prediction Machine Learning Model aims to provide accurate and timely predictions of fall foliage patterns across the United States. By utilizing historical foliage data, year-to-date weather records, and advanced machine learning techniques, the project seeks to address the growing demand for precise and user-friendly tools to plan autumn-related activities, tourism, and environmental analysis.

Objectives

1. Develop a machine learning model that predicts the timing and intensity of fall foliage at a granular geographical level.
2. Create a reliable pipeline for collecting, preprocessing, and integrating data from multiple sources, including NOAA weather data and Explore Fall fall foliage archived data.
3. Evaluate the model's performance by using different metrics including accuracy, mean squared error (MSE), or F1 scores to ensure reliability and precision.
4. Integrate model's predictions and results into a database using PostgreSQL for visualization and model-training purposes.
5. Visualize the prediction results in an intuitive and interactive format such as heatmaps or intensity map with time included.

This project not only highlights the intersection of data science, machine learning, and environmental science. But also provides practical value for stakeholders, including tourism agencies, environmental researchers, and individual travelers.

Introduction

Background & Motivation

As a computer scientist, countless hours are spent in front of monitors, navigating complex algorithms, analyzing numbers, and solving challenging problems. While this work is intellectually stimulating, it can often lead to mental fatigue and a diminished capacity for creativity and productivity. To maintain peak performance and balance, it's essential to engage in hobbies and activities that allow the mind to reset and recharge.

One such activity is exploring nature, which has been shown to reduce stress, improve focus, and spark creativity. Fall, in particular, offers a unique opportunity to connect with nature through its vibrant foliage and breathtaking scenery. The annual transformation of leaves into shades of red, orange, and yellow draws millions of people outdoors, whether for hiking, photography, or simply appreciating the beauty of the season.

Problem Statement

The timing and intensity of fall foliage vary annually due to environmental factors including temperature, precipitation, and daylight, making it difficult for travelers, photographers, and tourism agencies to plan effectively. Existing prediction methods are often too generalized, and climate change adds further uncertainty.

To address these challenges, this project leverages machine learning to create a precise and adaptable model for predicting fall foliage patterns. By integrating historical data and environmental variables, it aims to provide a reliable, user-friendly solution for planning and optimizing fall-related activities.

This project aims to develop a machine learning model to predict the timing and intensity of fall foliage across the United States. It involves collecting and integrating historical foliage and weather data, training and optimizing a predictive model, and presenting results through interactive visualizations like maps and charts.

Data Collection Process

Source of Data

This Machine Learning Model will rely solely on two different datasets:

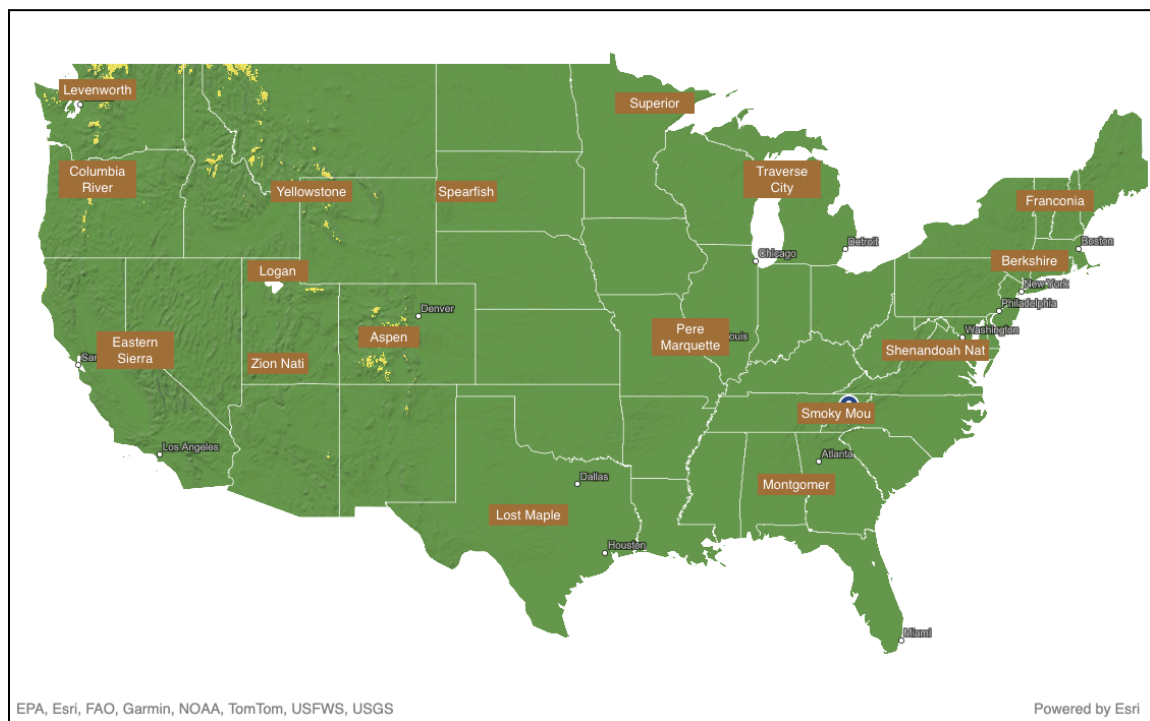
- Archived Fall Foliage Data - [Explore Fall](#)
- Year-to-Date Weather Data - [NOAA](#)

Archived Fall Foliage Data

Data Selection

With a time constraint of roughly 3 months, this project will be focusing on only a few key and major destinations rather than attempting to cover the entire U.S. This approach allows for more efficient travel planning, reduces the stress of overextending, and ensures a richer tourism experience at iconic spots known for their stunning autumn colors.

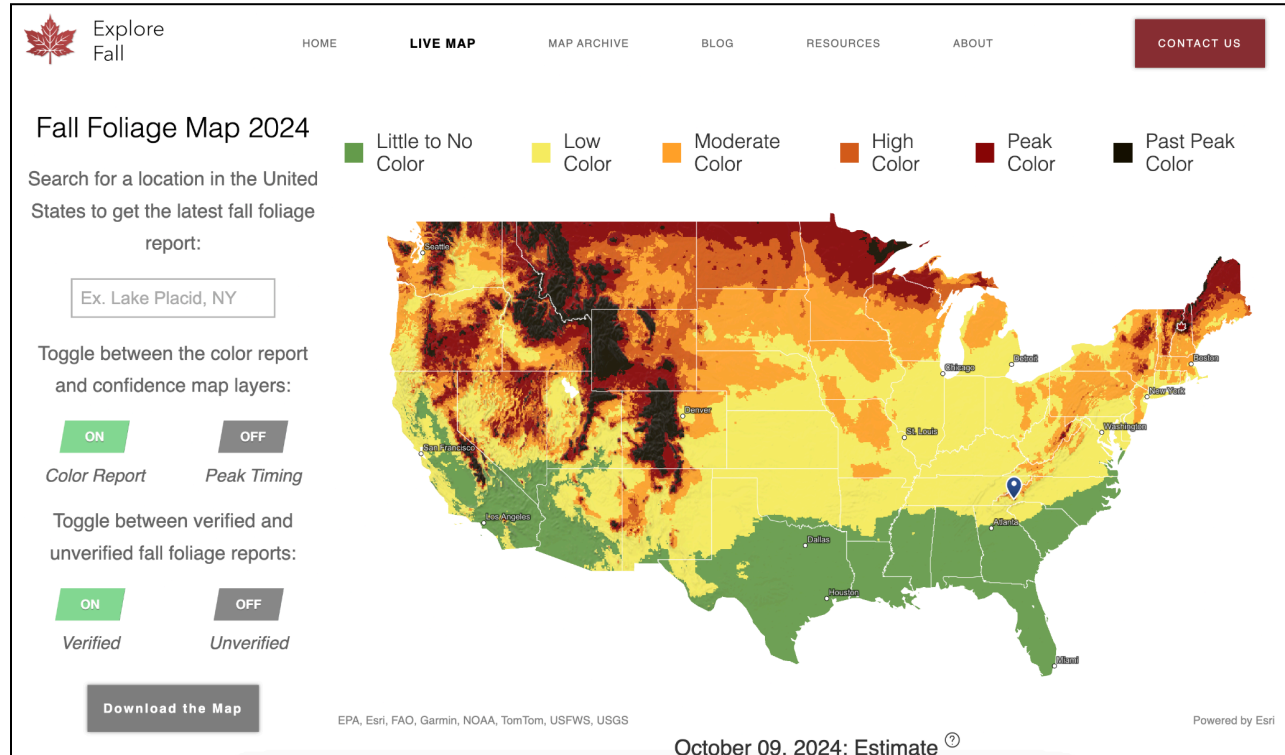
The following destinations are the most scenic and popular choices for fall foliage sightseeing according to *Thrillist*, *US News Travel*, *Culture Trekking*, and *Lonely Planet*:



(Selected locations for the project)

Data Pre-Processing

The process of data preprocessing for Archived Fall Foliage Data will be done manually since there is no existing formatted dataset for Annual Fall Foliage. Instead, there will be an interactive map from Explore Fall that shows the intensity of Fall colors with dates indicator.



[Explore Fall Interactive Fall Intensity Map](#)

This project will be utilizing Fall Intensity Data from the past 10 years from Explore Fall. The Processed Data will includes:

- Locations - Aspen, Traverse City, Stowe, etc.
- The Time - Early (1st - 9th), Mid (10th - 19th), Late (20th - 31st).
- Months - September, October, November
- Intensity - Little to No Color, Low - Moderate, High - Peak, Past-Peak

**Data will be recorded into Structured Query Language (SQL) file format.
Updating and modifying using PostgreSQL for better efficient use of the
Machine Learning process.**

Year-to-Date Weather Data

Data Selection