Forest Cover Type Prediction

Overview & Dataset:

Aimed at environmental conservation, this project predicts the tree species of 30m x 30m forest patches in the **Roosevelt National Forest**. The dataset includes 54 features covering elevation, slope, aspect, and distance to hydrology, along with soil type and wilderness area designations.

Technologies & Model:

A **Random Forest Classifier** was selected for its ability to handle large feature sets and capture non-linear relationships between topographic data and vegetation. The model processes 56 variables across 7 different cover types (e.g., Spruce/Fir, Aspen, Ponderosa Pine).

Performance & Real-Time Inference:

The model achieved a robust **accuracy of 0.8724**. Feature importance analysis revealed that **Elevation** was the most significant predictor of species distribution, followed by Horizontal Distance to Roadways and Fire Points.

Conclusion:

This project demonstrates how topographic and environmental data can be used for automated ecological mapping and forest management at scale.