

Forest Cover Type Prediction - Feature Importance Report

Top 10 Feature Importance - Analysis

The bar chart below displays the top 10 features used by the Random Forest classifier to predict forest cover types. The feature importance score reflects how valuable each feature was in building the decision trees.

Key Observations:

1. Elevation is the most important feature with the highest importance score (> 0.20).
-> It strongly influences vegetation zones based on altitude.
2. Horizontal distances to roadways, fire points, and hydrology play significant roles.
-> These features indicate accessibility and human disturbance, which affect forest composition.
3. The 'Id' column appears as a top feature, but this is misleading.
-> It's just a row number and should be dropped to avoid bias.
4. Hillshade (sunlight index) and Aspect (slope direction) also impact predictions, likely due to their influence on microclimate and sunlight exposure.
5. Wilderness_Area4 is among the top 10, implying that this particular protected region has distinct vegetation patterns.

Suggested Action:

- Drop 'Id' from the feature list and retrain the model to get a more generalizable and trustworthy result.

Forest Cover Type Prediction - Feature Importance Report

