How Random is your Forest?

BenJAMIN Lin

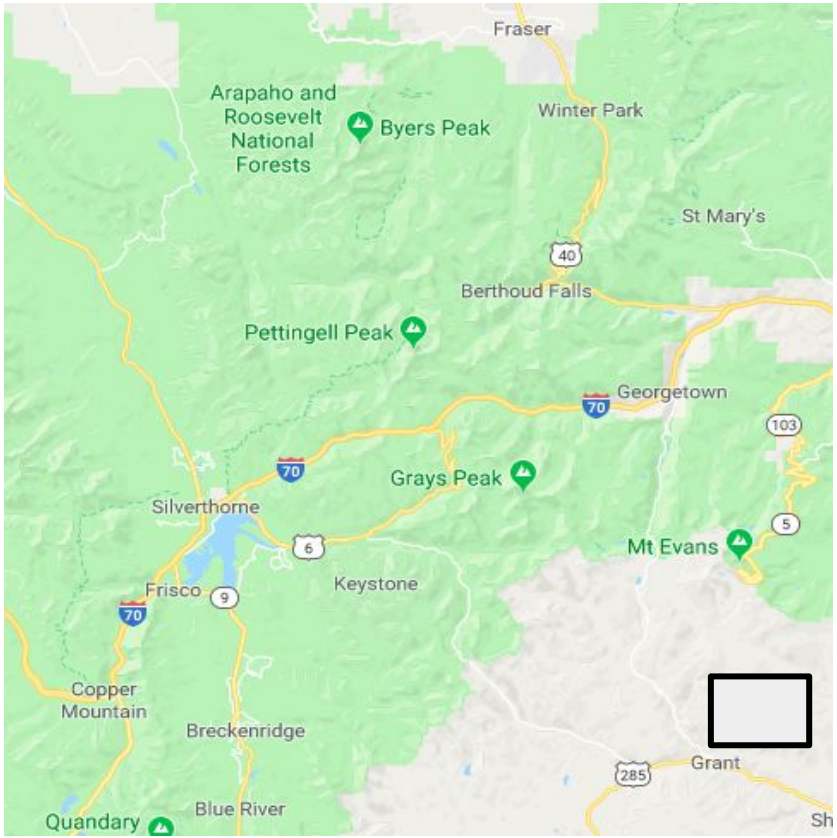
5/16/2018

Agenda

1. Background & Objectives
2. Exploratory Data Analysis
3. Models Overview
4. Conclusion

Background & Objectives

Objectives



Predict the predominant tree type

EDA

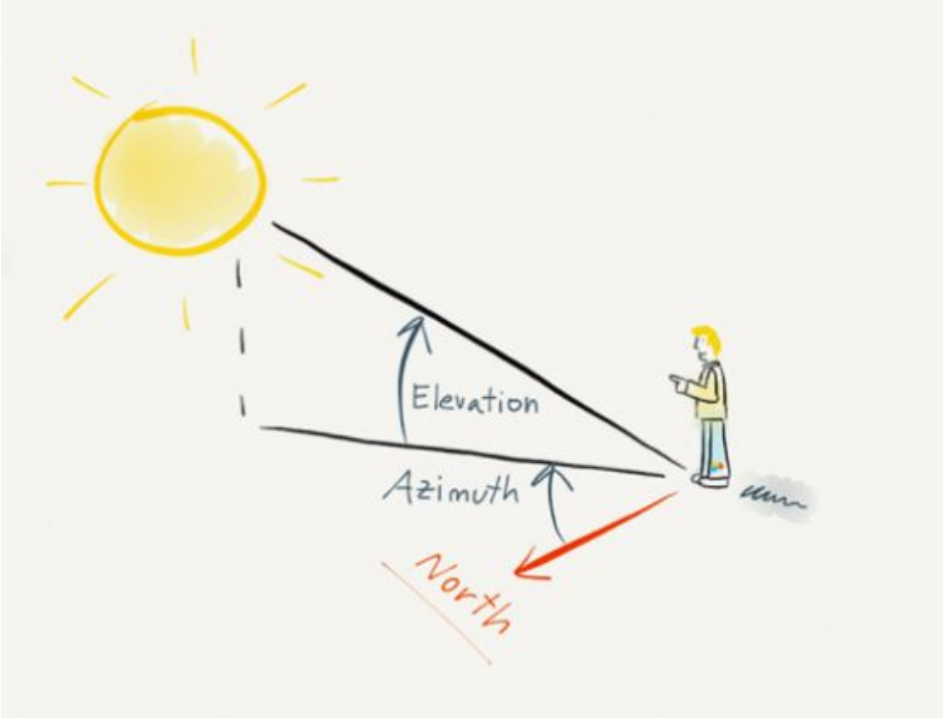
Data

Cartographic Variables:

- Elevation
- Slope
- Aspect (Azimuth)
- Distance to water, road, fire

Environmental Variables:

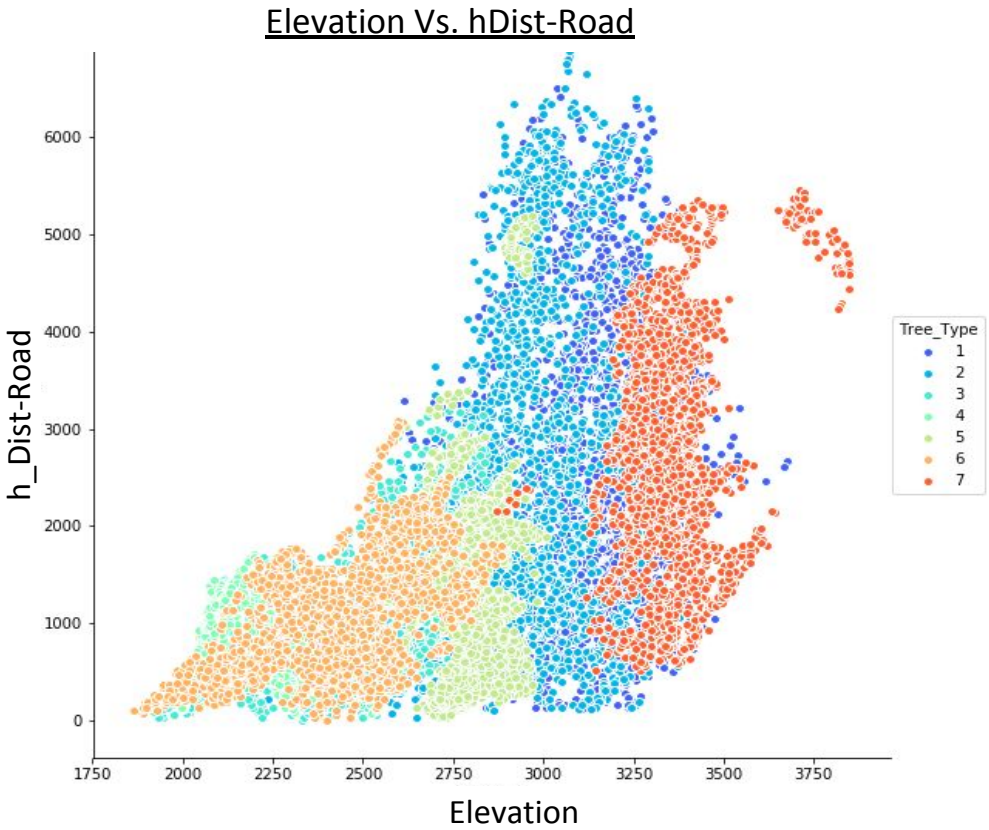
- Tree Type* (7)
- Wilderness areas* (4)
- Soil types* (40)
- Shade levels



* = Categorical

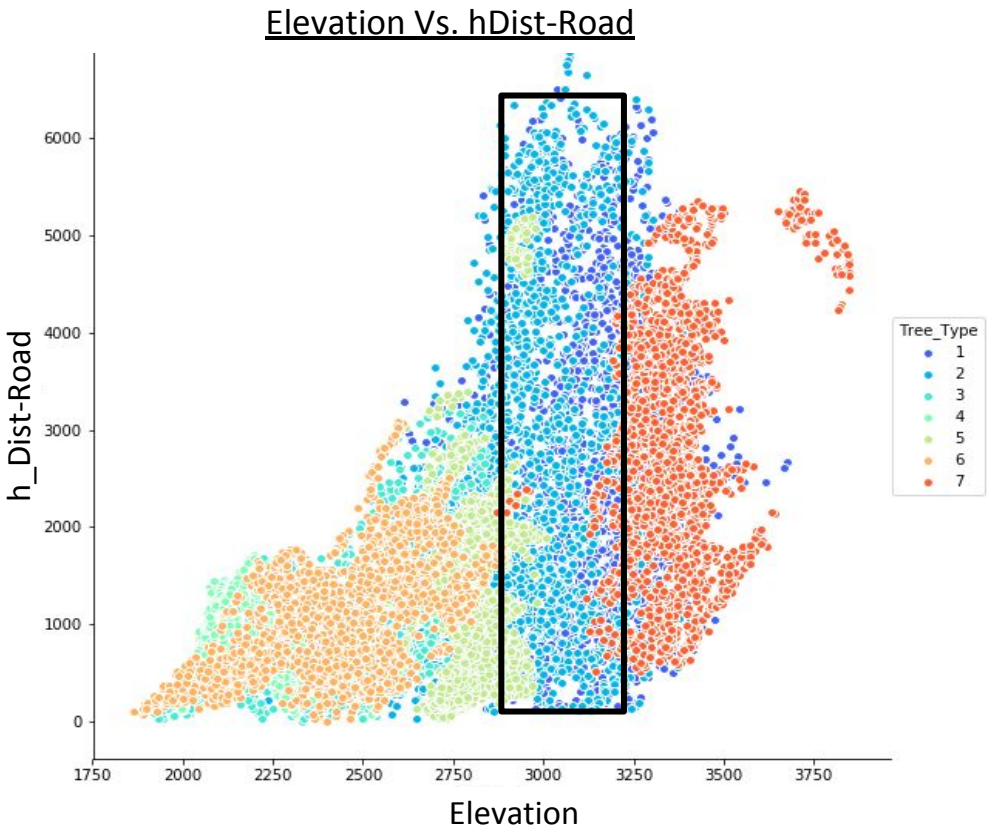
EDA

- Visualized the highest correlated features in respect to Tree Type



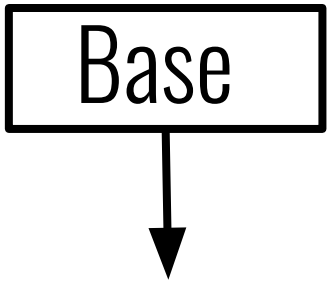
EDA

- Visualized the highest correlated features in respect to Tree Type



Models Overview

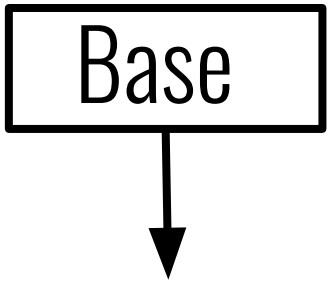
Models Overview



Gaussian N.B.

KNN

Models Overview



Gaussian N.B.
(0.54)

KNN
(0.83)

Models Overview

Gaussian N.B.
(0.54)

KNN
(0.83)

SVC (Built-in)

SVC (Manual)

Models Overview

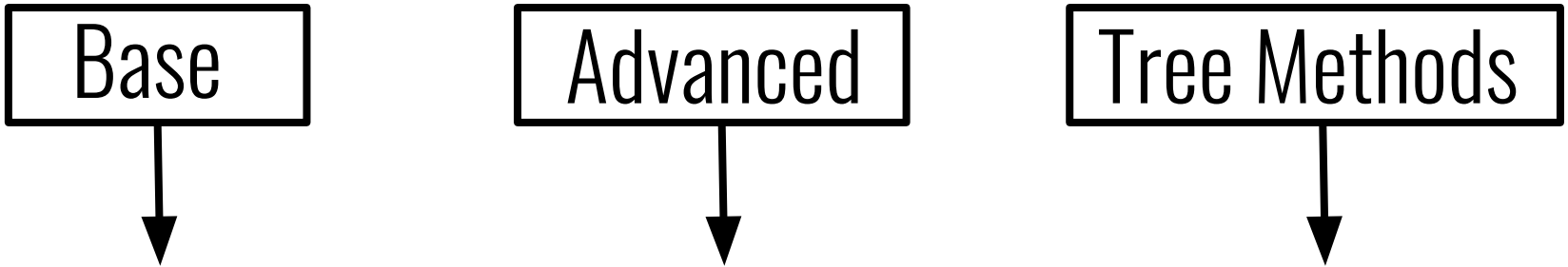
Gaussian N.B.
(0.54)

KNN
(0.83)

SVC (Built-in)
(0.84)

SVC (Manual)
(0.79)

Models Overview



Gaussian N.B.
(0.54)

KNN
(0.83)

SVC (Built-in)
(0.84)

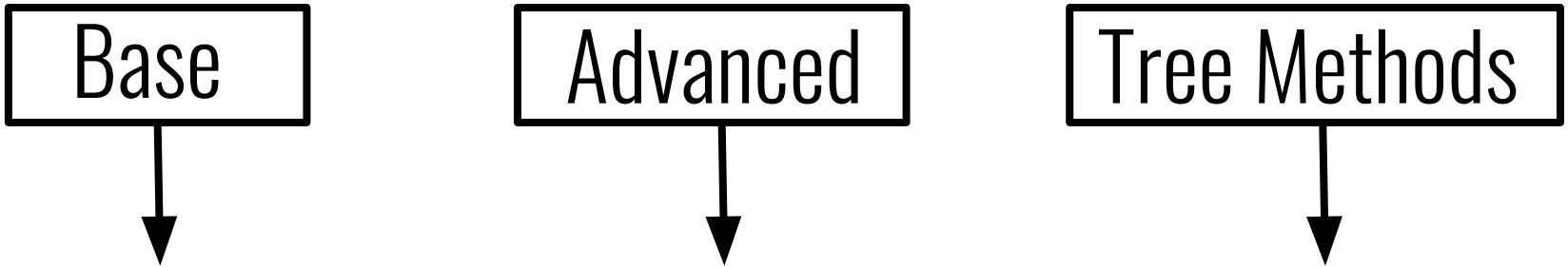
SVC (Manual)
(0.79)

Decision Tree

Extra Trees

Random Forest

Models Overview



Gaussian N.B.
(0.54)

KNN
(0.83)

SVC (Built-in)
(0.84)

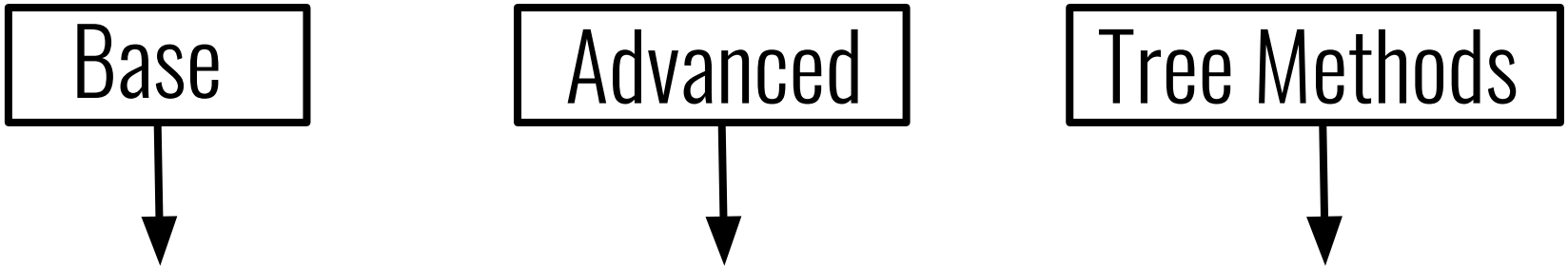
SVC (Manual)
(0.79)

Decision Tree
(0.8)

Extra Trees
(0.84)

Random Forest
(0.87)

Models Overview



Gaussian N.B.
(0.54)

KNN
(0.83)

SVC (Built-in)
(0.84)

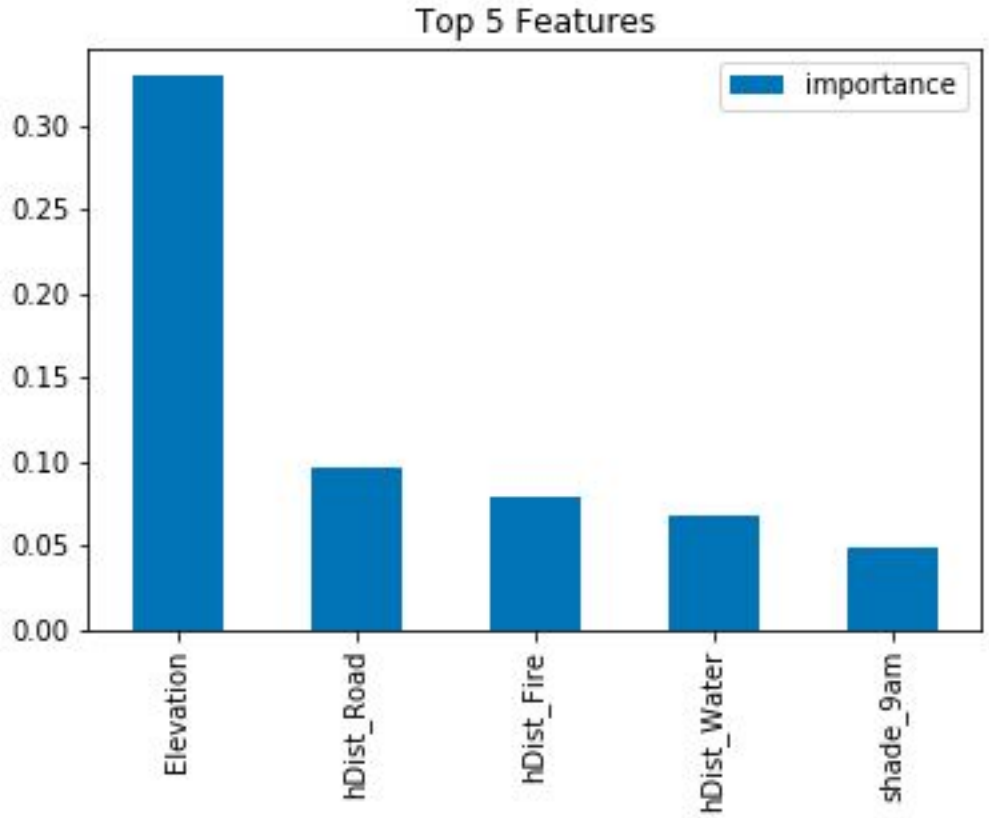
SVC (Manual)
(0.79)

Decision Tree
(0.8)

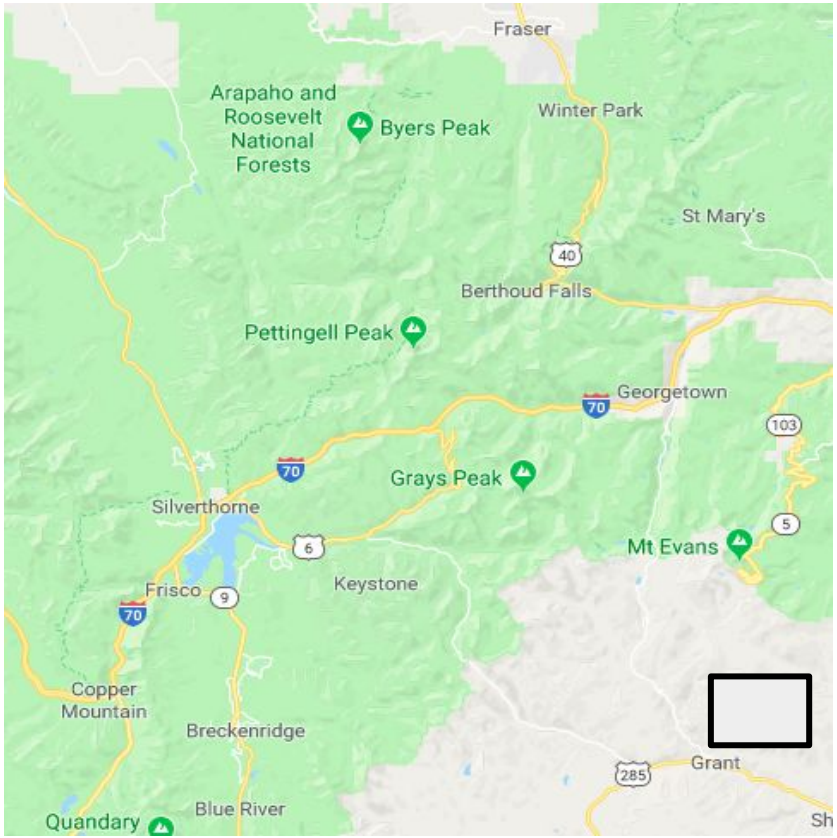
Extra Trees
(0.84)

Random Forest
(0.87)

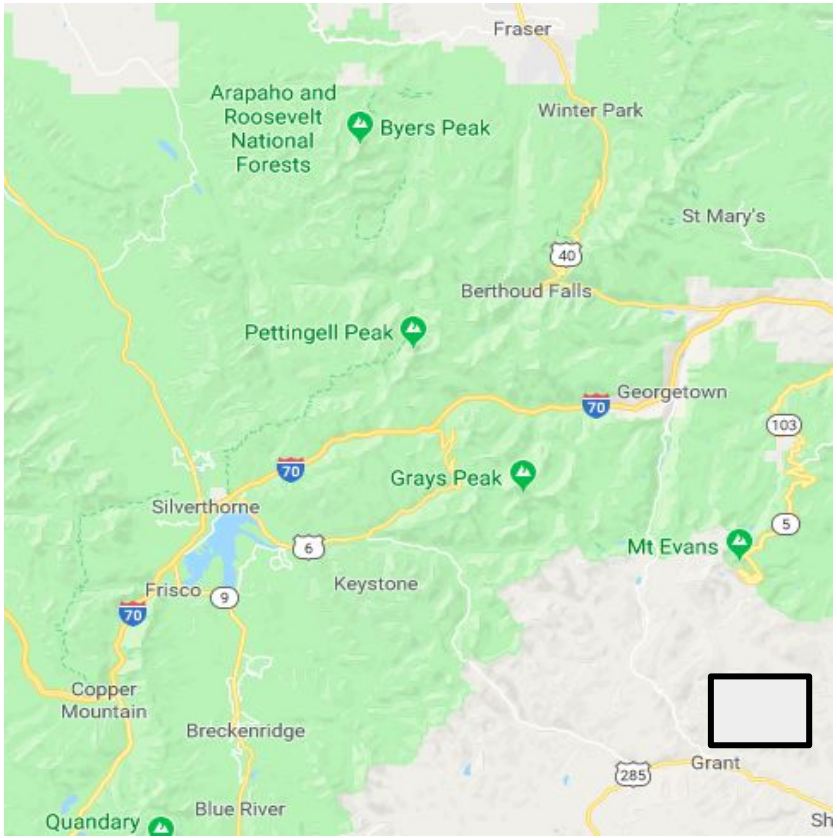
Feature Importance



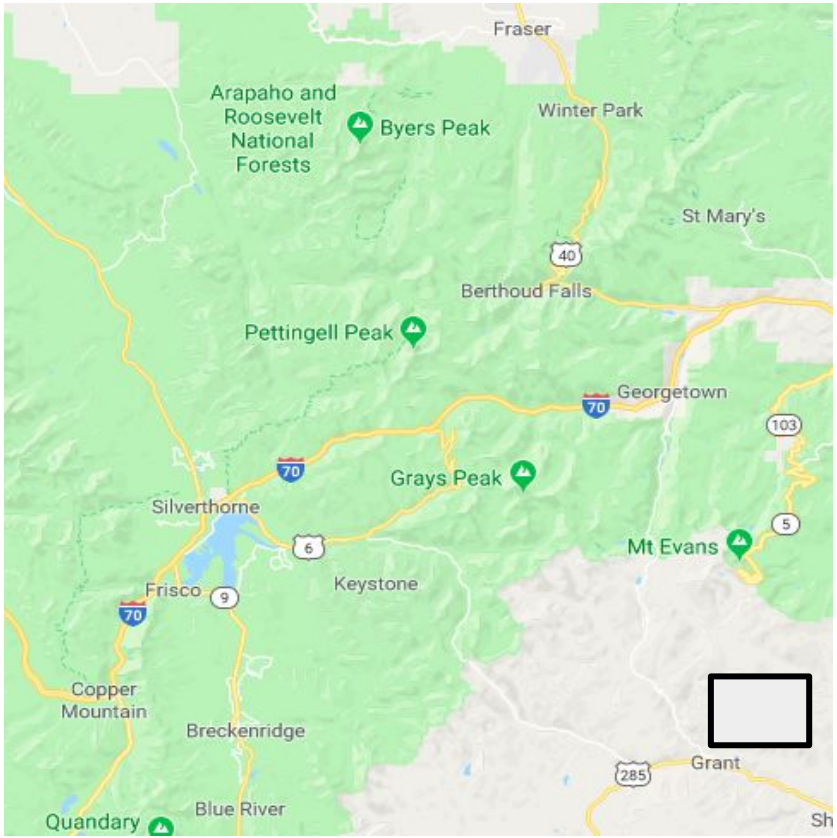
Conclusion

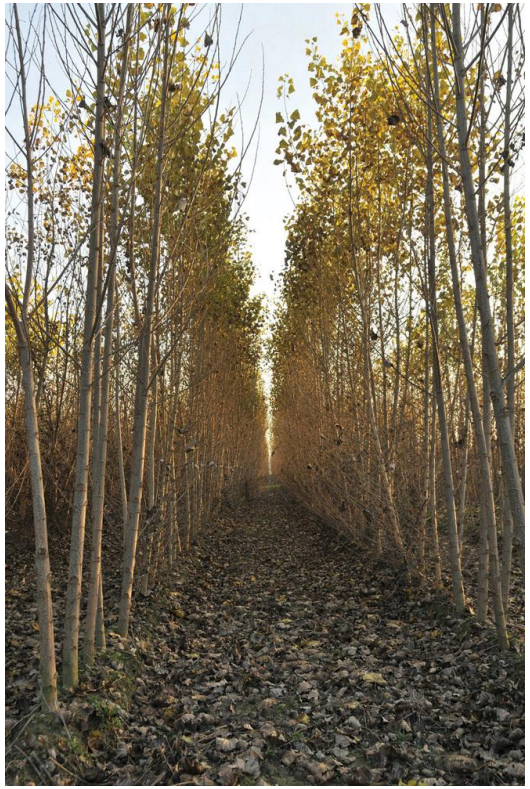


Predict the predominant tree type



Tree Type 5 - Cottonwood!


Fraser
Arapaho and Roosevelt National Forests
Byers Peak
Winter Park
St Mary's
40
Berthoud Falls
Pettingell Peak
Georgetown
70
103
70
Grays Peak
Silverthorne
6
5
Mt Evans
Keystone
Frisco
9
70
Copper Mountain
Breckenridge
Grant
285
Blue River
Quandary
Sh



Thank you!

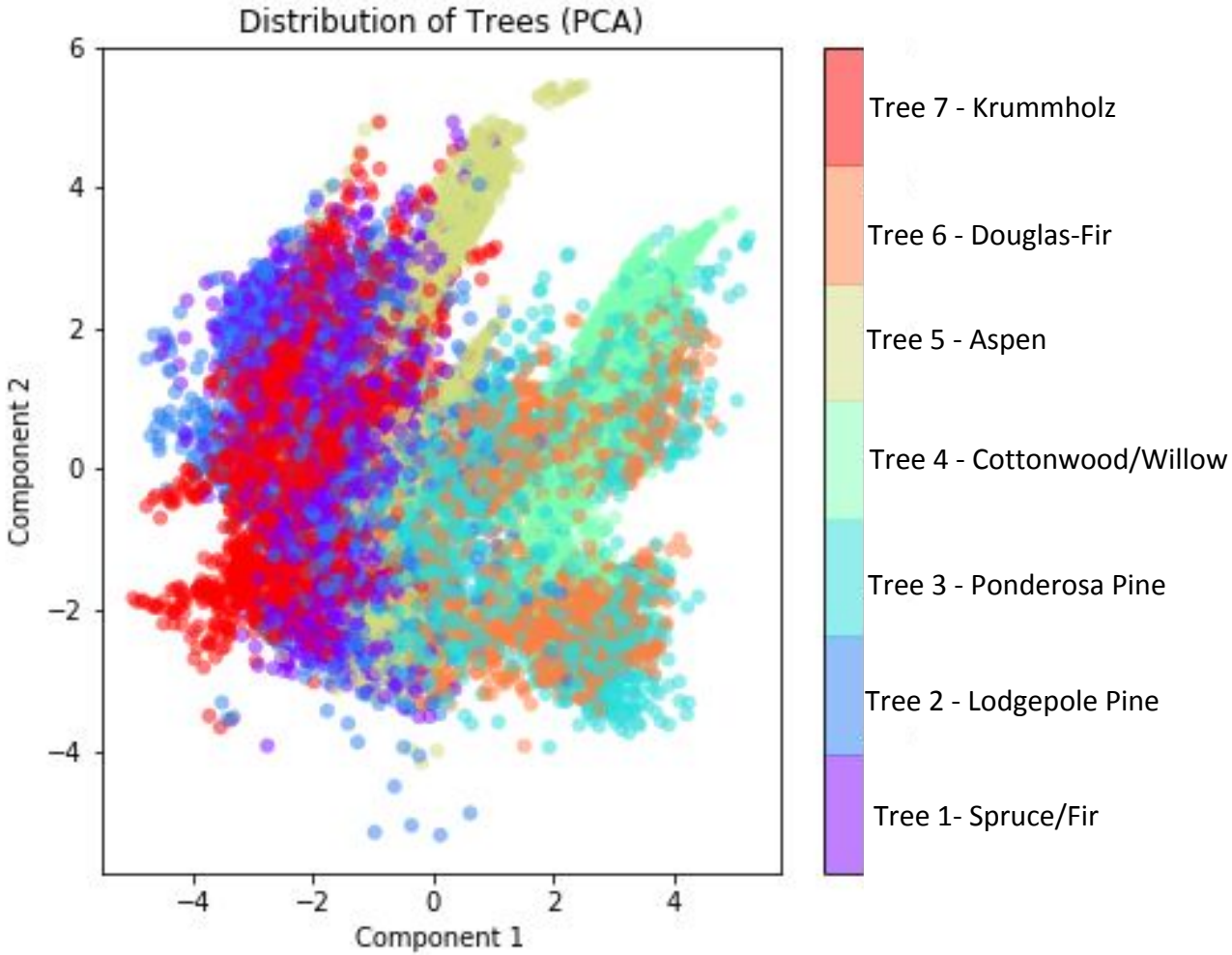
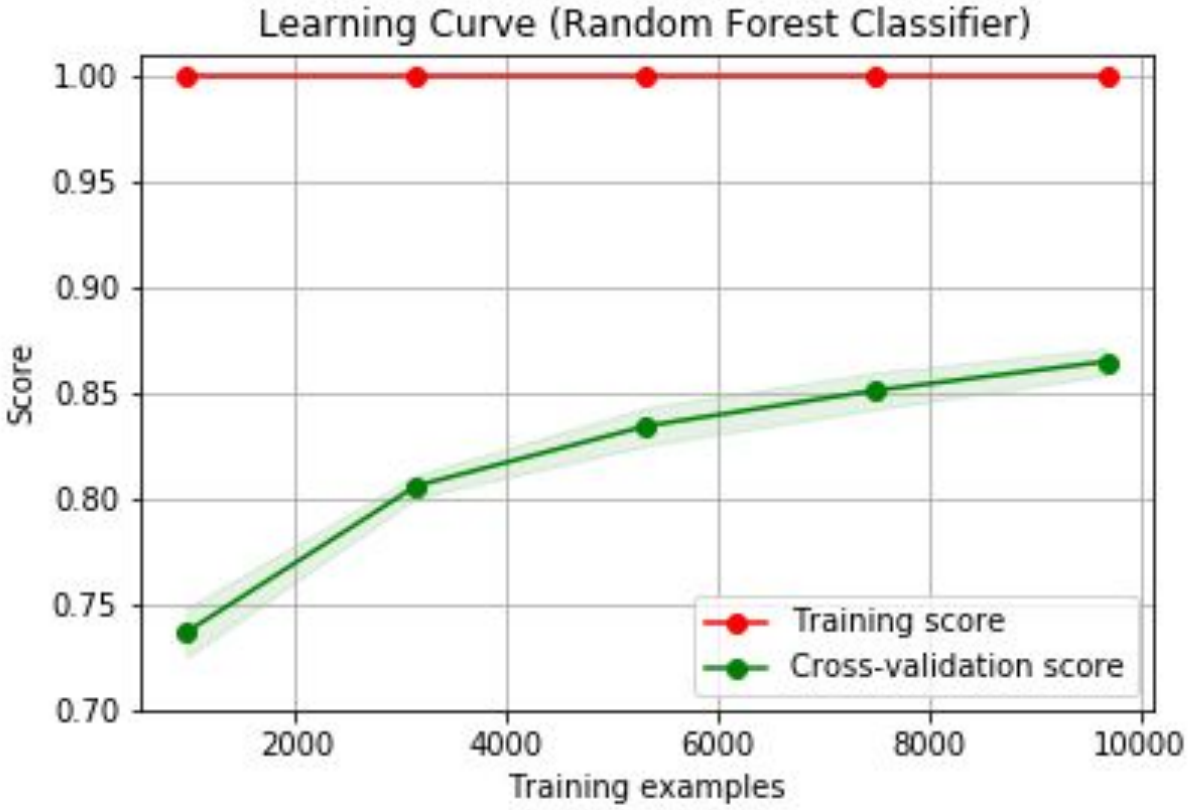
Contact:

Email: benlin100@gmail.com

LinkedIN: linkedin.com/in/benjaminlin24/

GitHub: github.com/benlin100

Appendix


Distribution of Trees (PCA)
Component 2
Component 1
6
4
2
0
−2
−4
−4
−2
0
2
4
Tree 7 - Krummholz
Tree 6 - Douglas-Fir
Tree 5 - Aspen
Tree 4 - Cottonwood/Willow
Tree 3 - Ponderosa Pine
Tree 2 - Lodgepole Pine
Tree 1- Spruce/Fir

Model Analysis