What's covering me? - Prediction of forest cover type based on cartographic variables

Devin Austin and Duncan Tulimieri

University of Delaware

May 16, 2022

The Purpose

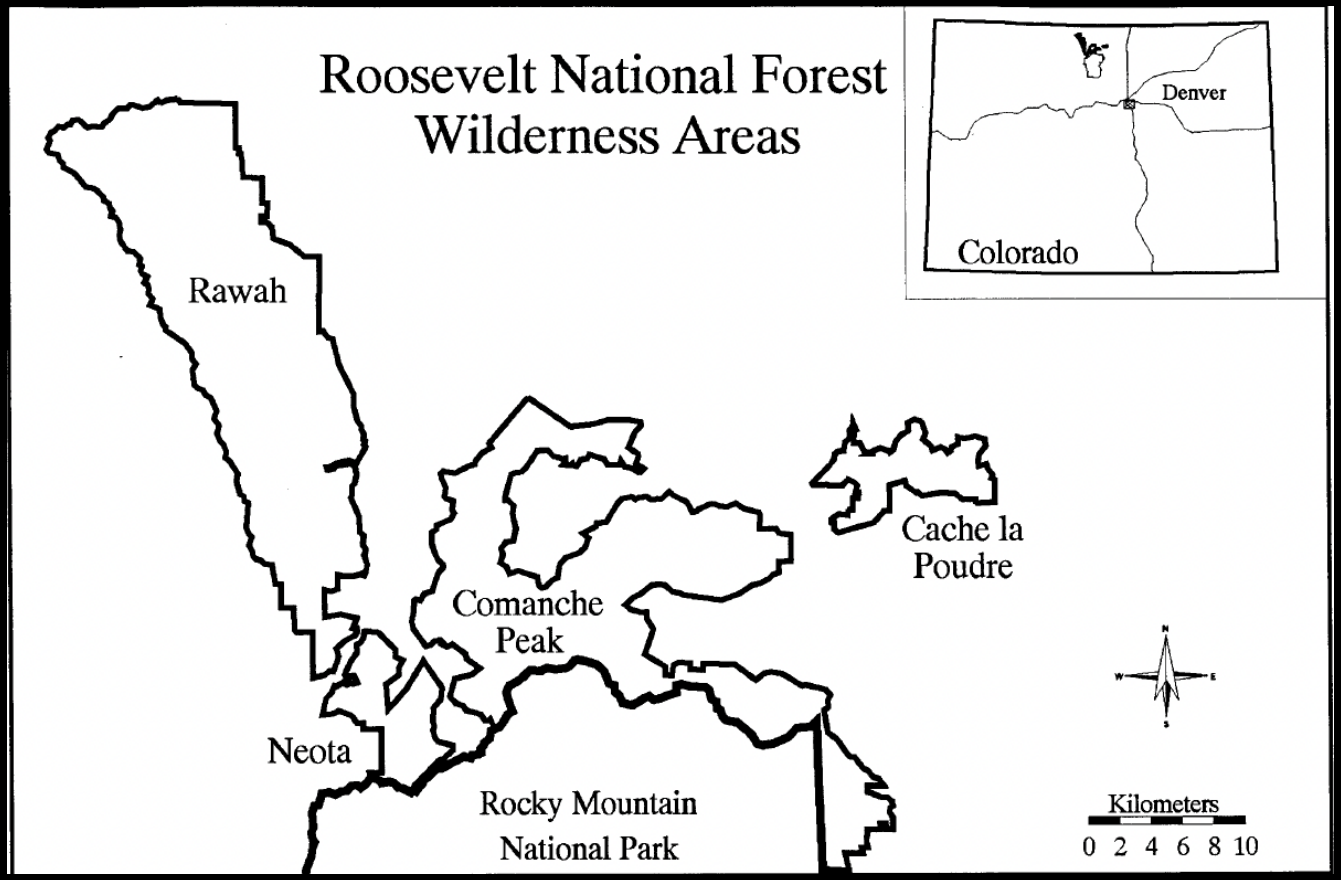
- Apply various machine learning techniques

The Purpose

- Apply various machine learning techniques
- Explore ensemble learning

The Data Set

- Cartographic data from four parks
  - Rawah, Comanche Peak, Neota, and Cache la Poudre



The Data Set

Spruce/Fir

Cottonwood/Willow

Krummholz

Lodgepole Pine

Aspen

Ponderosa Pine

Douglas Fir

Predictors of Forest Cover Type:

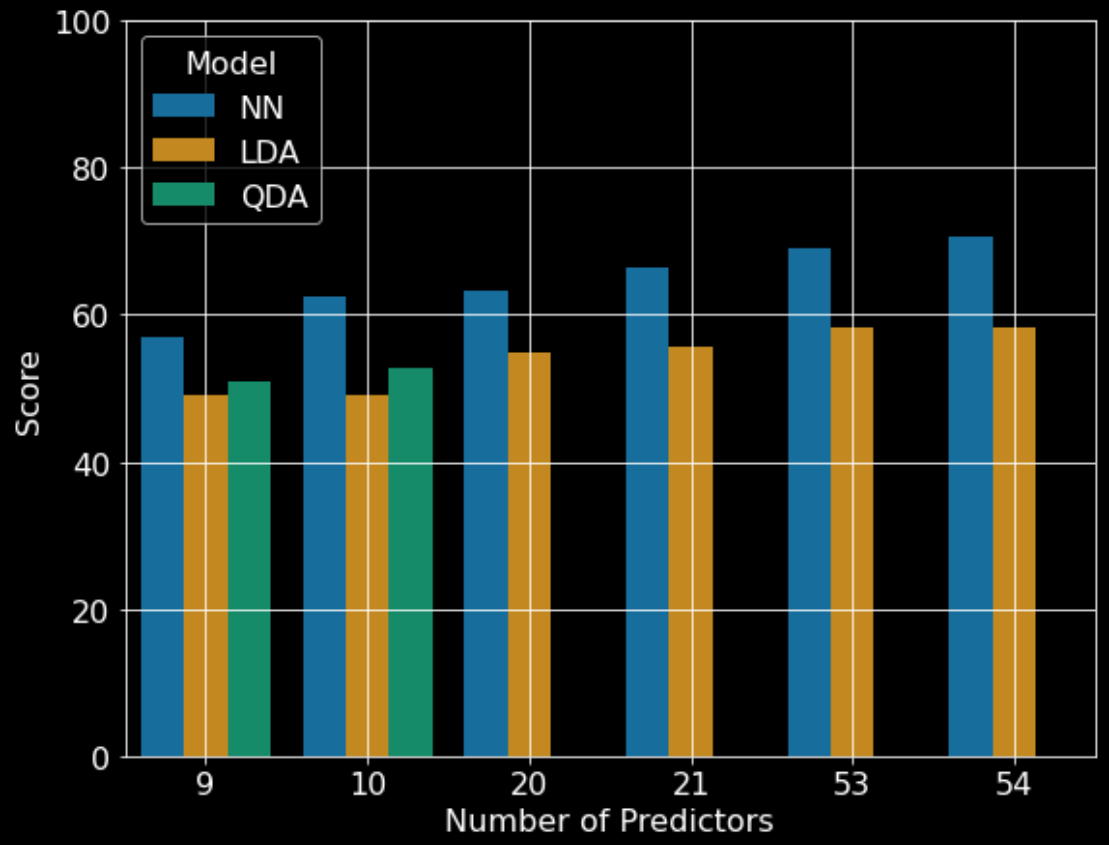
- Elevation in meters
- Aspect in degrees
- Slope in degrees
- Horz Dist water
- Vert Dist to water
- Horz Dist to roadway
- Hillshade index
- Horz Dist to wildfire
- Wilderness area
- Soil Type

Current Literature

- Blackard and Dean (1999) - Compare neural network and discriminant analyses.

Current Literature

- Blackard and Dean (1999) - Compare neural network and discriminant analyses.



Current Literature

- Oza and Russell (2001)
  - Used ensemble learning
  - Reports effective

The Idea

- Train multiple (weak) classifiers (using CV to determine parameters)

The Idea

- Train multiple (weak) classifiers (using CV to determine parameters)
  - Each model learns something about data, but not everything

The Idea

- Train multiple (weak) classifiers (using CV to determine parameters)
  - Each model learns something about data, but not everything
- Combine models via taking majority vote (i.e., ensemble)

The Idea

- Train multiple (weak) classifiers (using CV to determine parameters)
  - Each model learns something about data, but not everything
- Combine models via taking majority vote (i.e., ensemble)
- Models
  - K Nearest Neighbors (KNN)

The Idea

- Train multiple (weak) classifiers (using CV to determine parameters)
  - Each model learns something about data, but not everything
- Combine models via taking majority vote (i.e., ensemble)
- Models
  - K Nearest Neighbors (KNN)
  - Support Vector Machine (SVM)

The Idea

- Train multiple (weak) classifiers (using CV to determine parameters)
  - Each model learns something about data, but not everything
- Combine models via taking majority vote (i.e., ensemble)
- Models
  - K Nearest Neighbors (KNN)
  - Support Vector Machine (SVM)
  - Linear Discriminant Analysis (LDA)

The Idea

- Train multiple (weak) classifiers (using CV to determine parameters)
  - Each model learns something about data, but not everything
- Combine models via taking majority vote (i.e., ensemble)
- Models
  - K Nearest Neighbors (KNN)
  - Support Vector Machine (SVM)
  - Linear Discriminant Analysis (LDA)
  - Quadratic Discriminant Analysis (QDA)

The Idea

- Train multiple (weak) classifiers (using CV to determine parameters)
  - Each model learns something about data, but not everything
- Combine models via taking majority vote (i.e., ensemble)
- Models
  - K Nearest Neighbors (KNN)
  - Support Vector Machine (SVM)
  - Linear Discriminant Analysis (LDA)
  - Quadratic Discriminant Analysis (QDA)
  - Logistic Regression

The Models - KNN

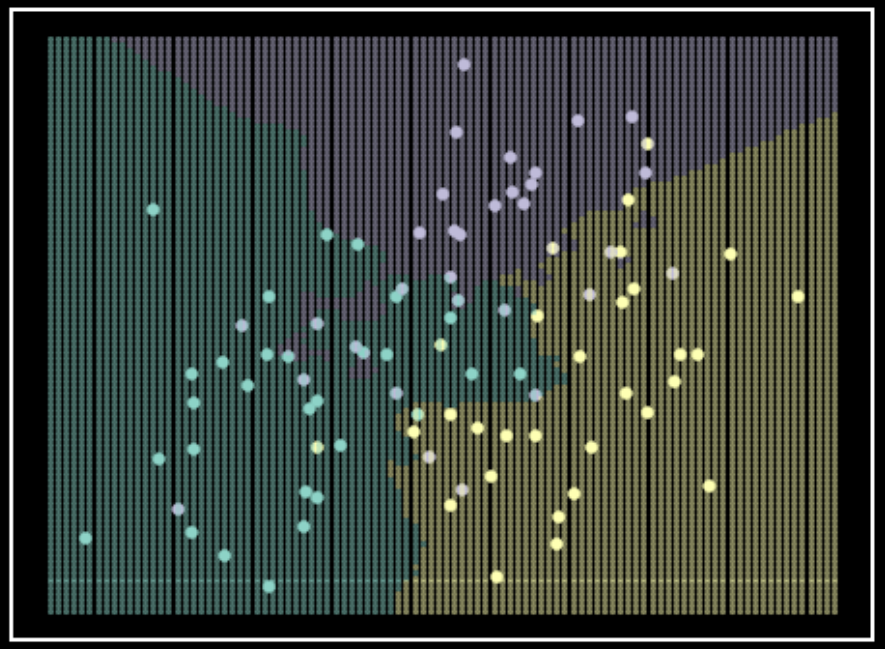
- K Nearest Neighbors (KNN)

The Models - KNN

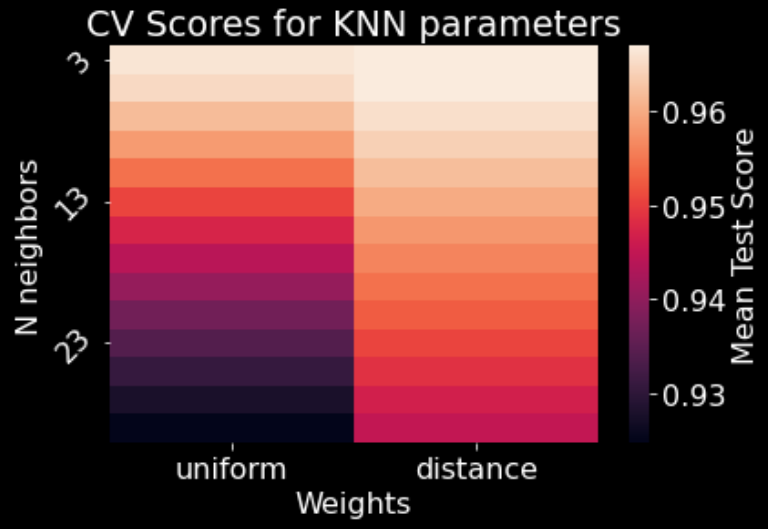
- K Nearest Neighbors (KNN)
  - Classify based on k neighbors, weighted by their distance

The Models - *K*NN

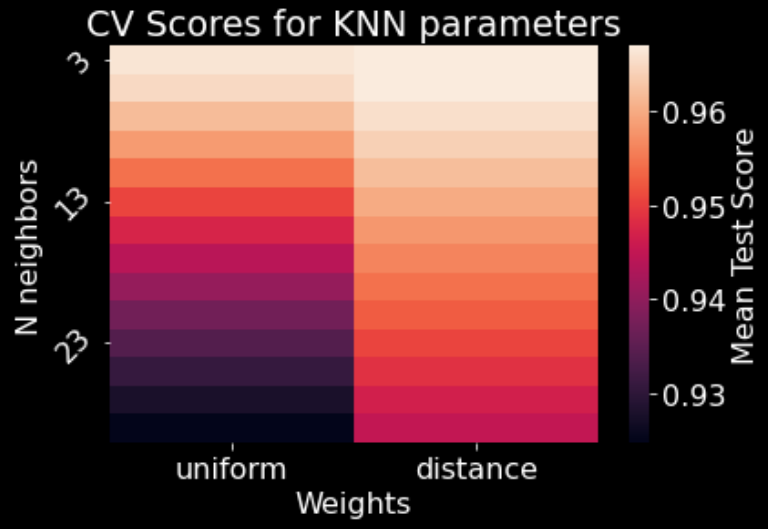
- *K* Nearest Neighbors (KNN)
  - Classify based on *k* neighbors, weighted by their distance

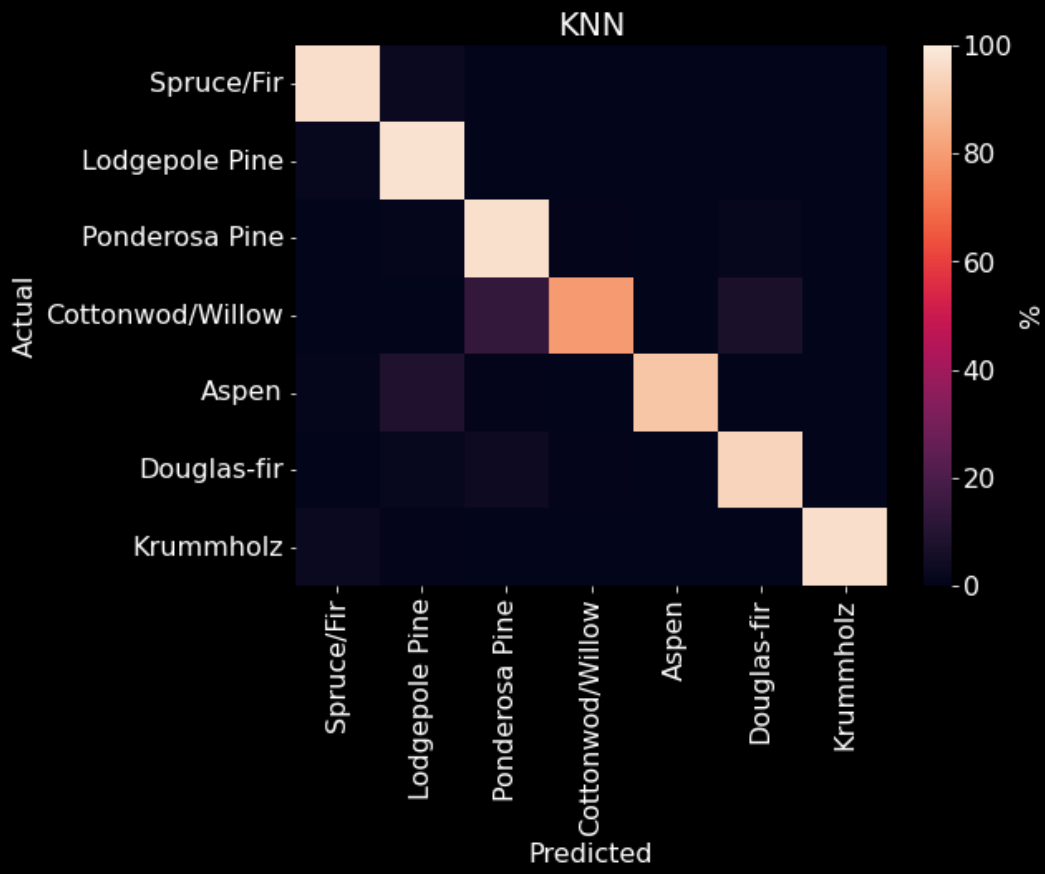


The Models - *K*NN



The Models - *K*NN





The Models - SVM

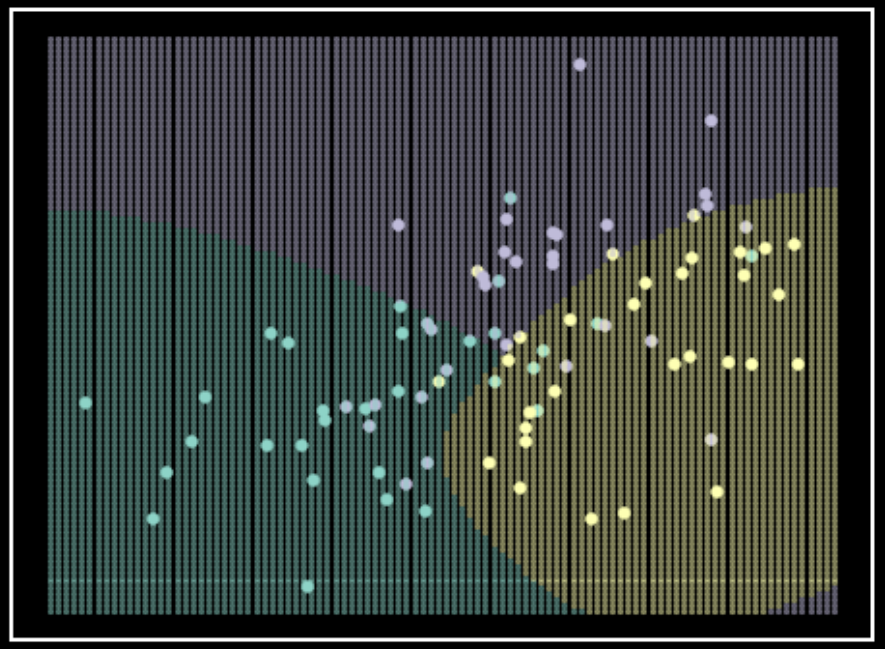
- Support Vector Machine (SVM)

The Models - SVM

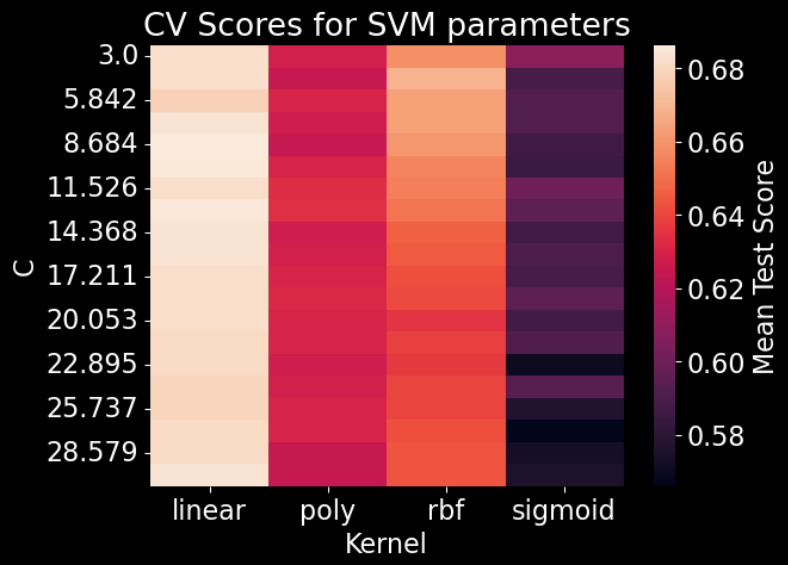
- Support Vector Machine (SVM)
  - Find a hyperplane that maximizes the distance between the closest points of each class.

The Models - SVM

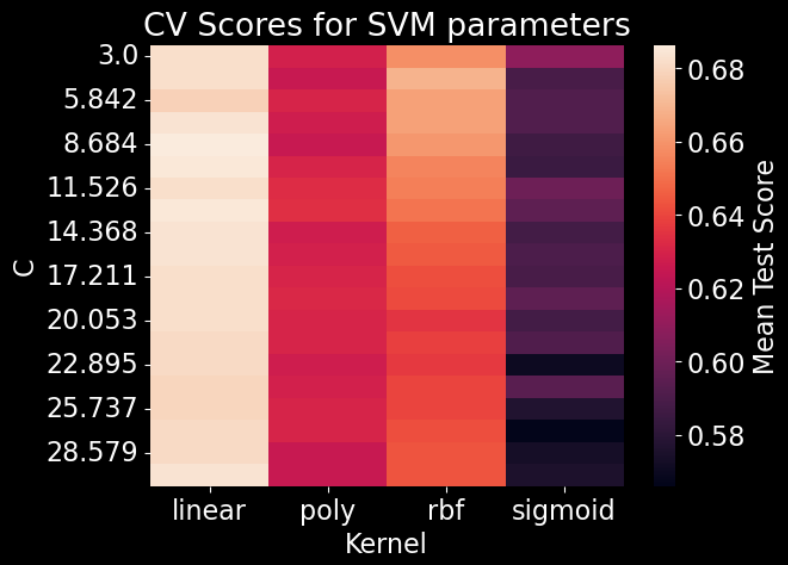
- Support Vector Machine (SVM)
  - Find a hyperplane that maximizes the distance between the closest points of each class.

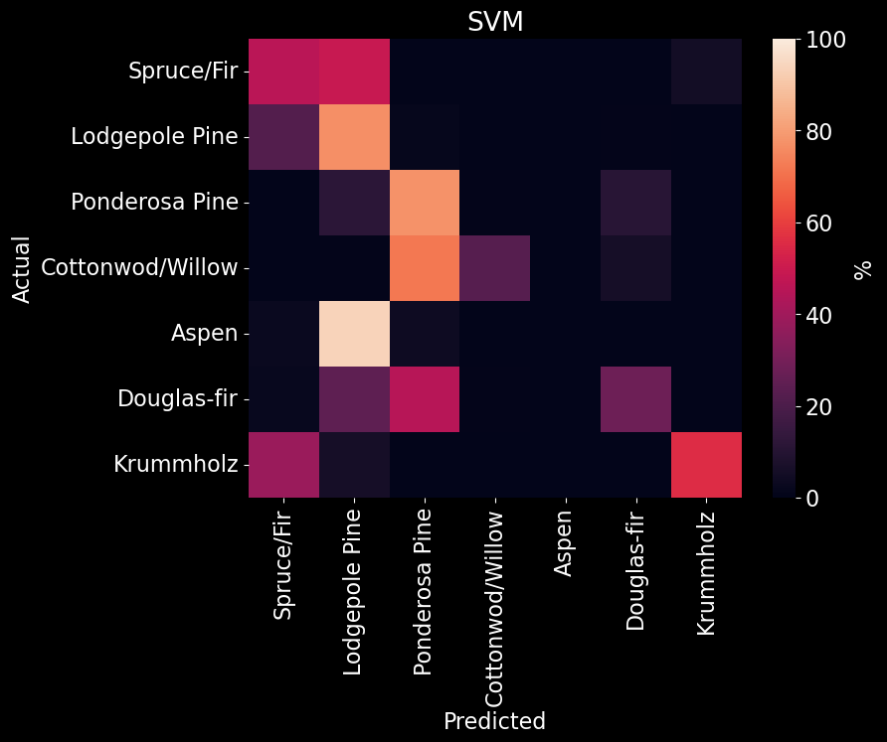


The Models - SVM



The Models - SVM





The Models - LDA

- Linear Discriminant Analysis (LDA)

The Models - LDA

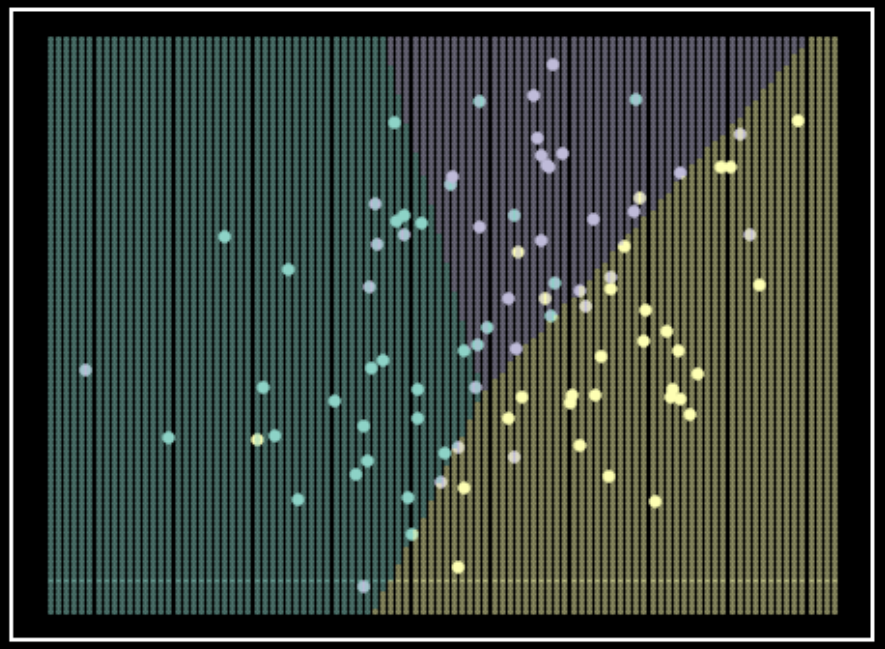
- Linear Discriminant Analysis (LDA)
  - Models each class as a multivariate Gaussian

The Models - LDA

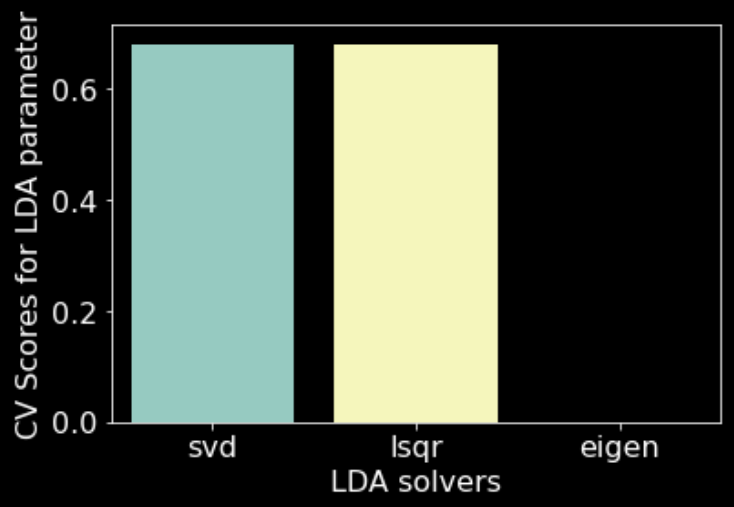
- Linear Discriminant Analysis (LDA)
  - Models each class as a multivariate Gaussian
  - Assumes shared covariance matrix between classes

The Models - LDA

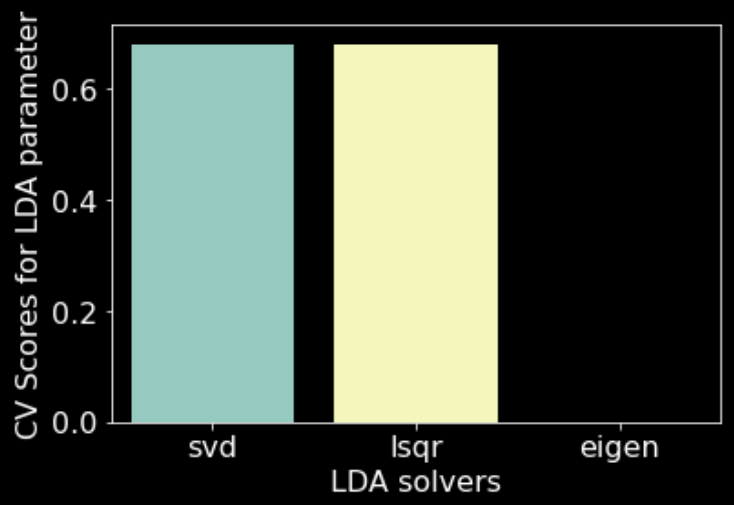
- Linear Discriminant Analysis (LDA)
  - Models each class as a multivariate Gaussian
  - Assumes shared covariance matrix between classes

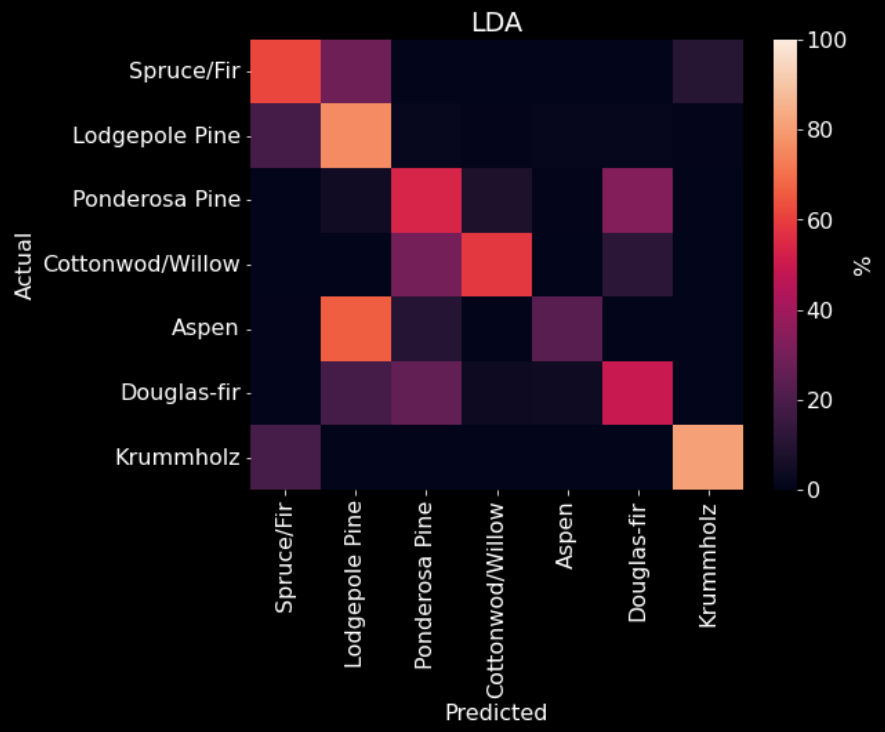


The Models - LDA



The Models - LDA





The Models - QDA

- Quadratic Discriminant Analysis (QDA)

The Models - QDA

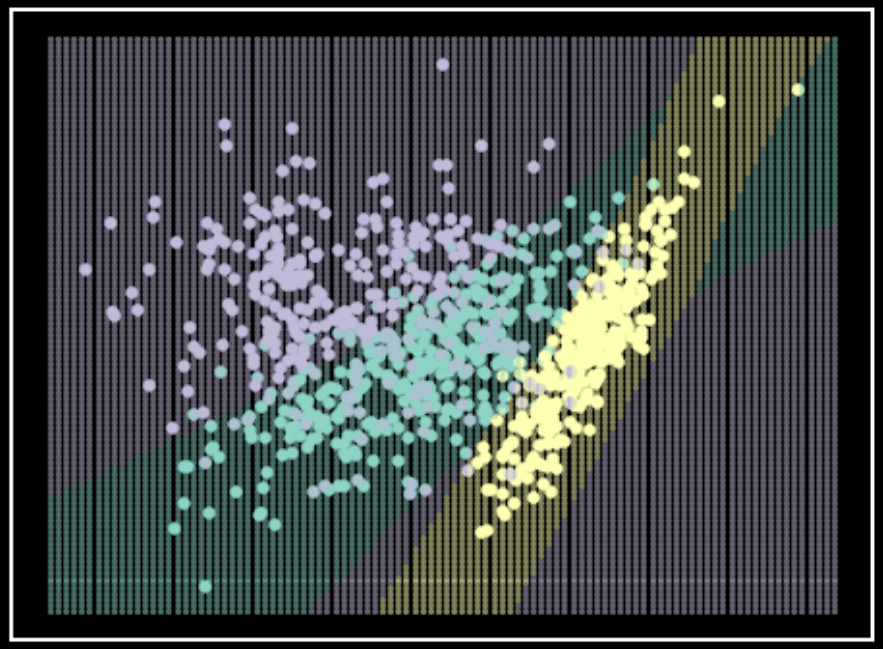
- Quadratic Discriminant Analysis (QDA)
  - Models each class as a multivariate Gaussian

The Models - QDA

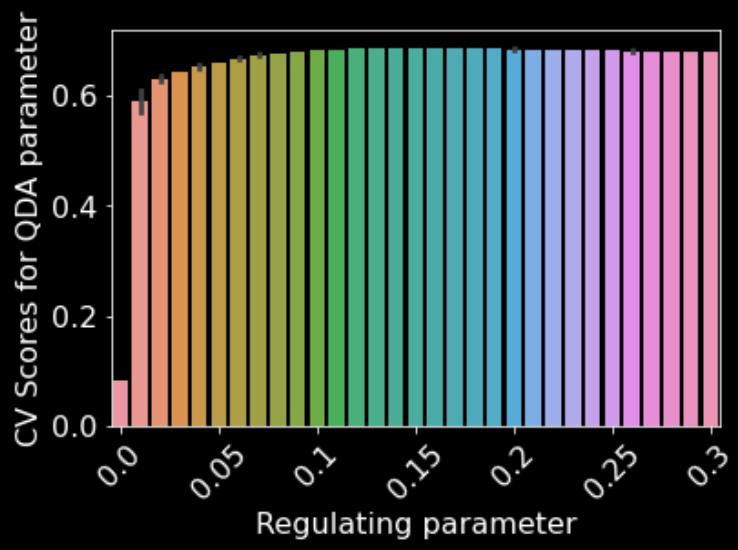
- Quadratic Discriminant Analysis (QDA)
  - Models each class as a multivariate Gaussian
  - Assumes each class has it's own covariance matrix

The Models - QDA

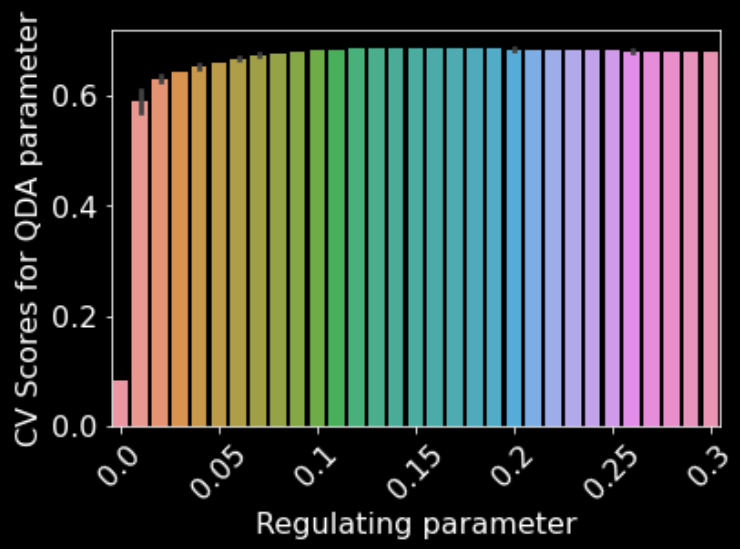
- Quadratic Discriminant Analysis (QDA)
  - Models each class as a multivariate Gaussian
  - Assumes each class has it's own covariance matrix

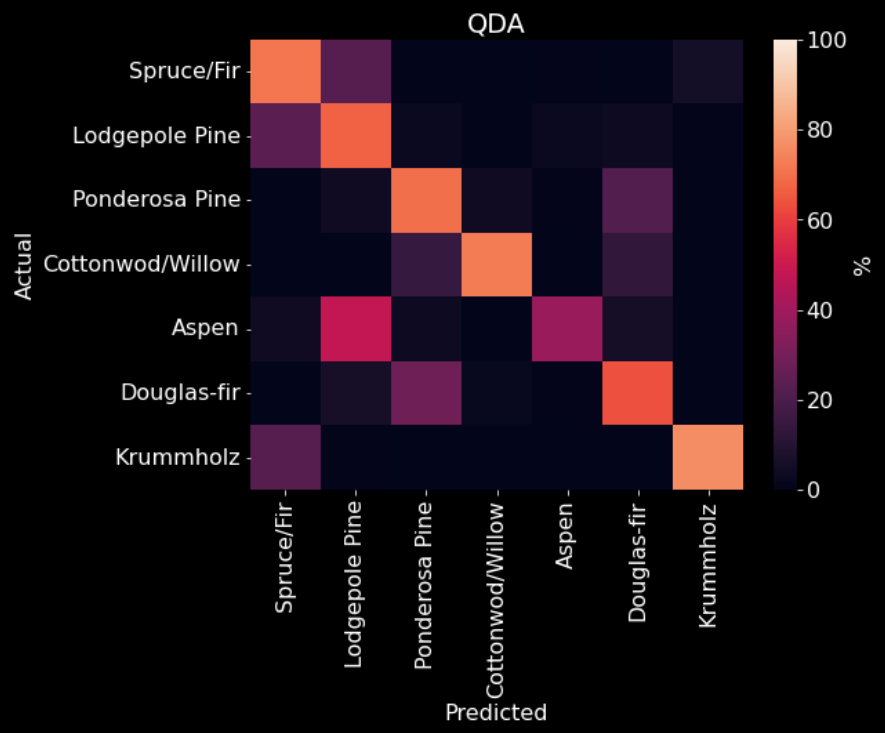


The Models - QDA



The Models - QDA





The Models - Logistic Regression

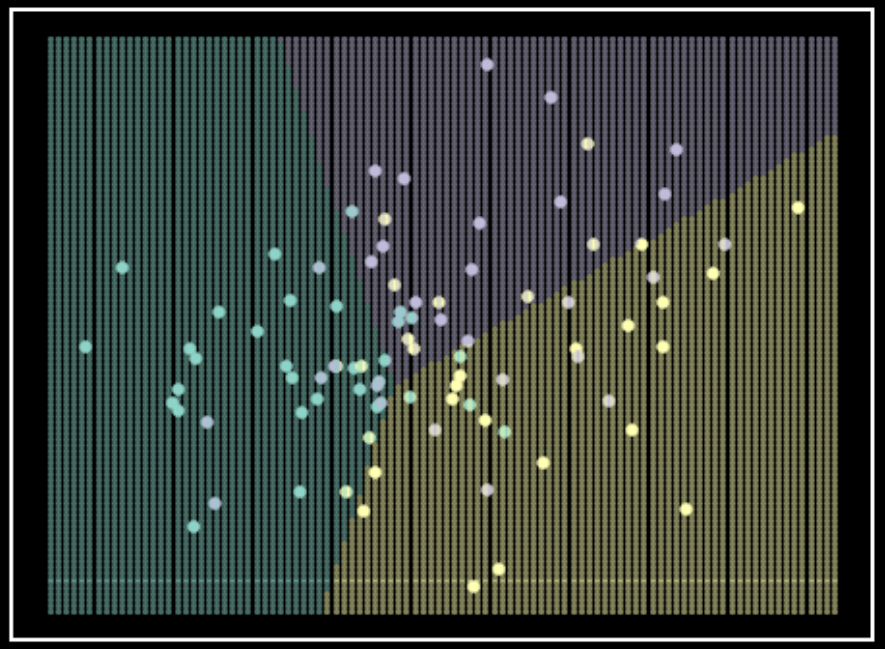
- Logistic Regression

The Models - Logistic Regression

- Logistic Regression
  - Uses sigmoid curve to classify

The Models - Logistic Regression

- Logistic Regression
  - Uses sigmoid curve to classify



The Models - Logistic Regression

- Parameters chosen via CV:

The Models - Logistic Regression

- Parameters chosen via CV:
  - penalty: l1, **<u>l2</u>**, elastic net, none

The Models - Logistic Regression

- Parameters chosen via CV:
  - penalty: l1, **<u>l2</u>**, elastic net, none
  - C: 5 linearly spaced values from 0.01 to 1, **<u>0.7525</u>**

The Models - Logistic Regression

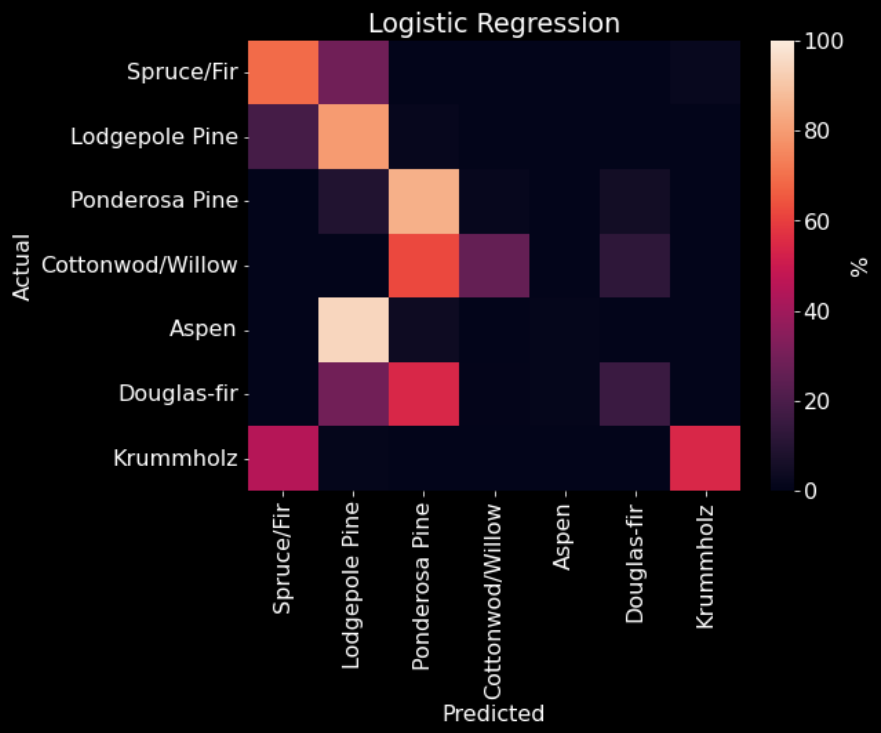
- Parameters chosen via CV:
  - penalty: l1, **<u>l2</u>**, elastic net, none
  - C: 5 linearly spaced values from 0.01 to 1, **<u>0.7525</u>**
  - intercept: True, **<u>False</u>**

The Models - Logistic Regression

- Parameters chosen via CV:
  - penalty: l1, **<u>l2</u>**, elastic net, none
  - C: 5 linearly spaced values from 0.01 to 1, **<u>0.7525</u>**
  - intercept: True, **<u>False</u>**
  - l1 ratio: 5 linearly spaced values from 0 to 1, **<u>0</u>**

The Models - Logistic Regression

- Parameters chosen via CV:
  - penalty: l1, **<u>l2</u>**, elastic net, none
  - C: 5 linearly spaced values from 0.01 to 1, **<u>0.7525</u>**
  - intercept: True, **<u>False</u>**
  - l1 ratio: 5 linearly spaced values from 0 to 1, **<u>0</u>**



The Models - Ensemble

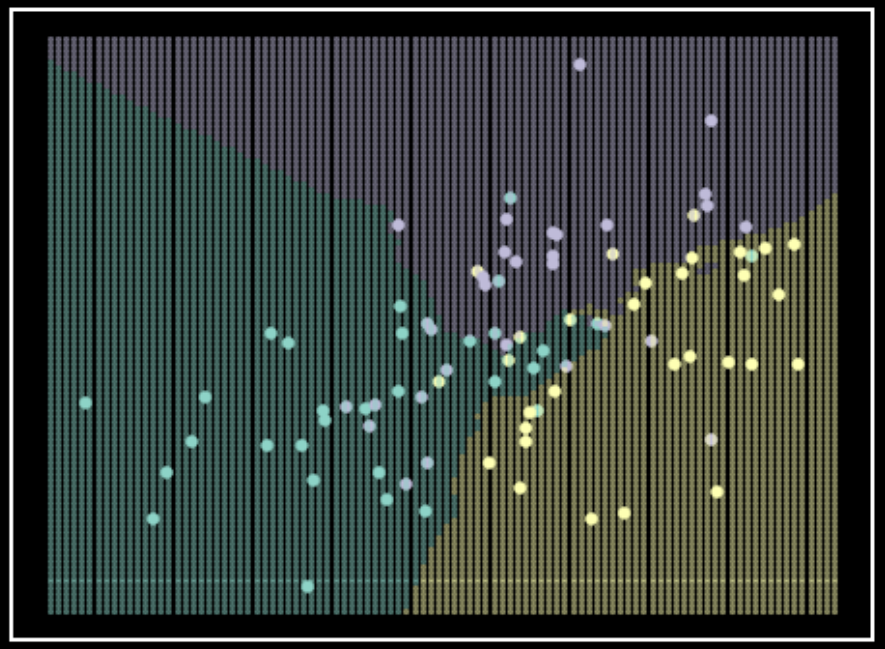
- Ensemble

The Models - Ensemble

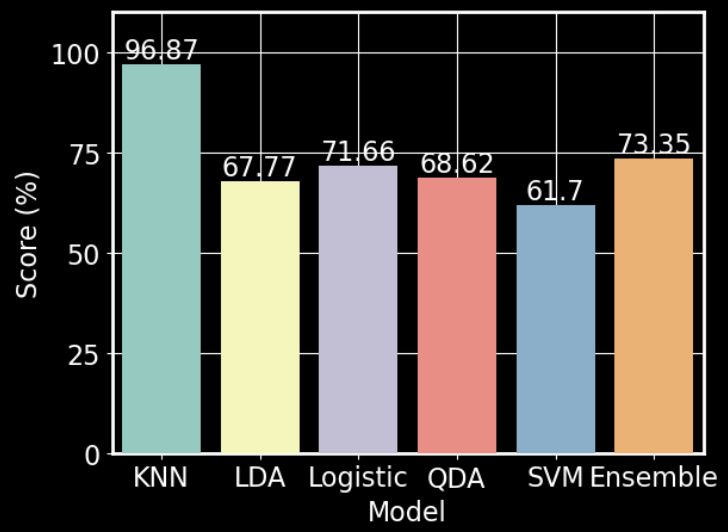
- Ensemble
  - Classifies based on majority vote of several models

The Models - Ensemble

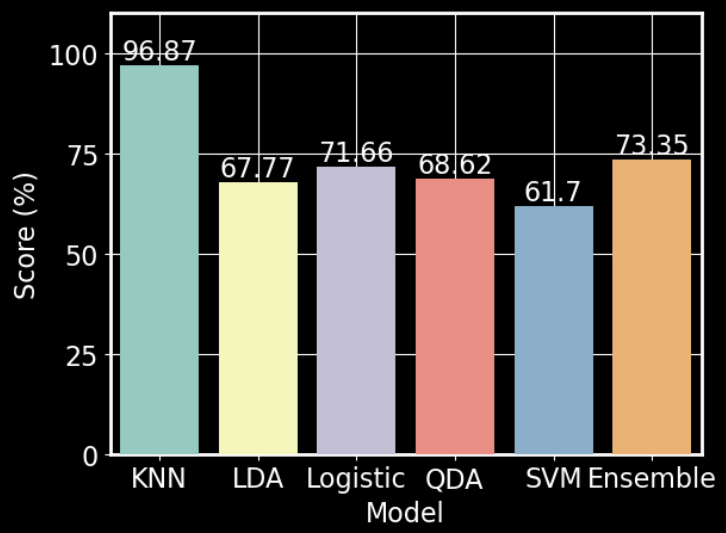
- Ensemble
  - Classifies based on majority vote of several models

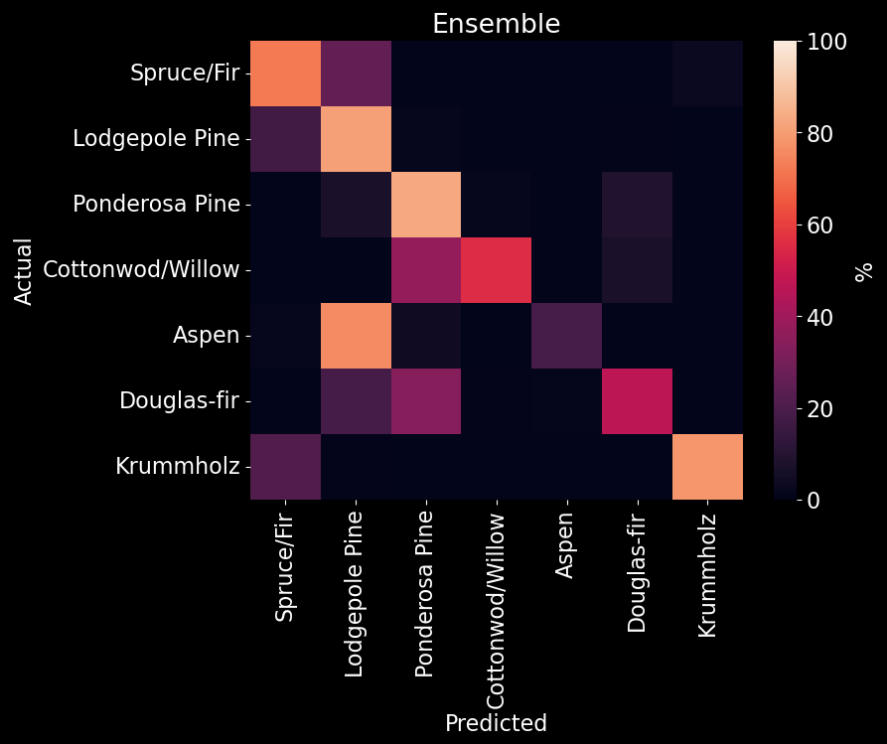


The Models - Ensemble



The Models - Ensemble





Future Directions

- Examine agreeability between models

Future Directions

- Examine agreeability between models
- Other models (deep learning, linear regression, naive Bayes, random forest, etc.)

Future Directions

- Examine agreeability between models
- Other models (deep learning, linear regression, naive Bayes, random forest, etc.)
- Train and test with more balanced data sets

Future Directions

- Examine agreeability between models
- Other models (deep learning, linear regression, naive Bayes, random forest, etc.)
- Train and test with more balanced data sets
- Ensemble prediction as weighted majority vote

Future Directions

- Examine agreeability between models
- Other models (deep learning, linear regression, naive Bayes, random forest, etc.)
- Train and test with more balanced data sets
- Ensemble prediction as weighted majority vote

Figure: Link to website.