Google

#GeoForGood19

EE Classifiers and Regression

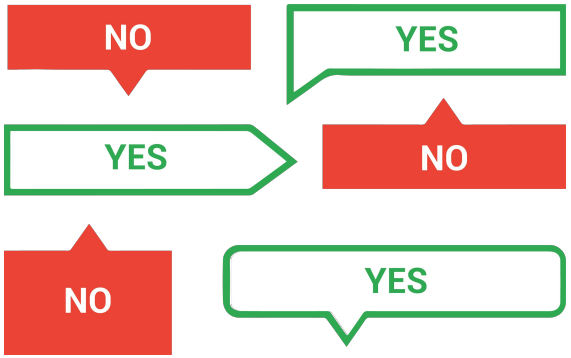
Arian Karbasi

Agenda

- What is Classification, Regression and Machine Learning?
- Supervised Classification example in EarthEngine
- Regression example in EarthEngine
- Unsupervised classification example in EarthEngine

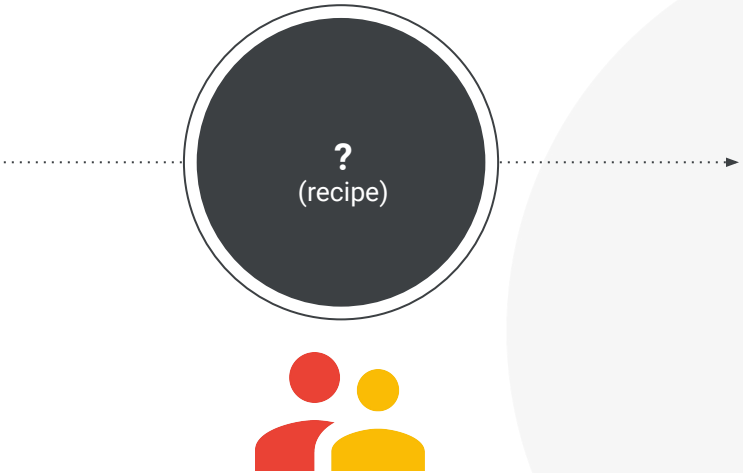
Input
BLACK BOX
Output

Machine learning is an approach to making lots of small decisions

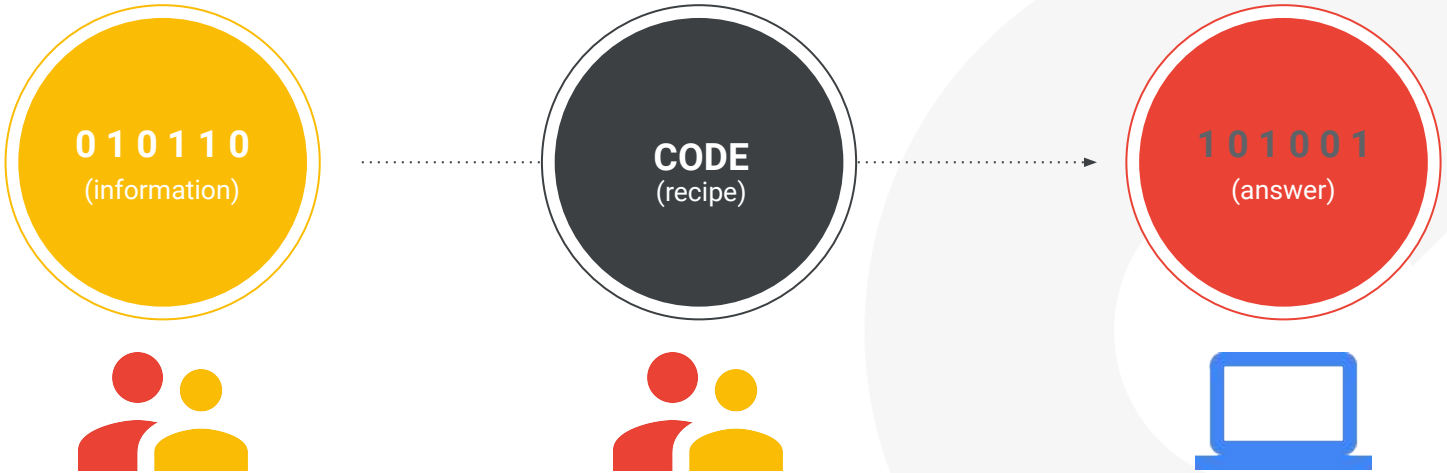


Machine learning vs traditional programming

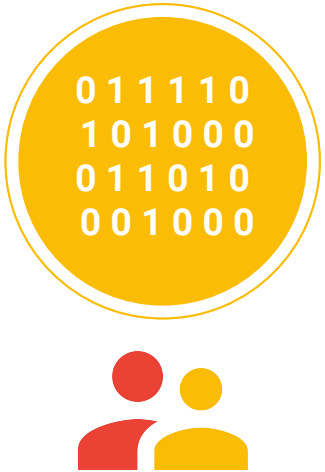
Traditional programming: Handcrafted recipe

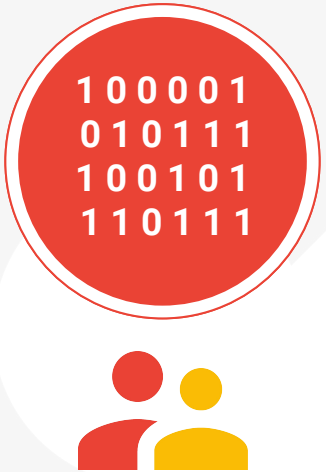


Traditional programming: Handcrafted recipe

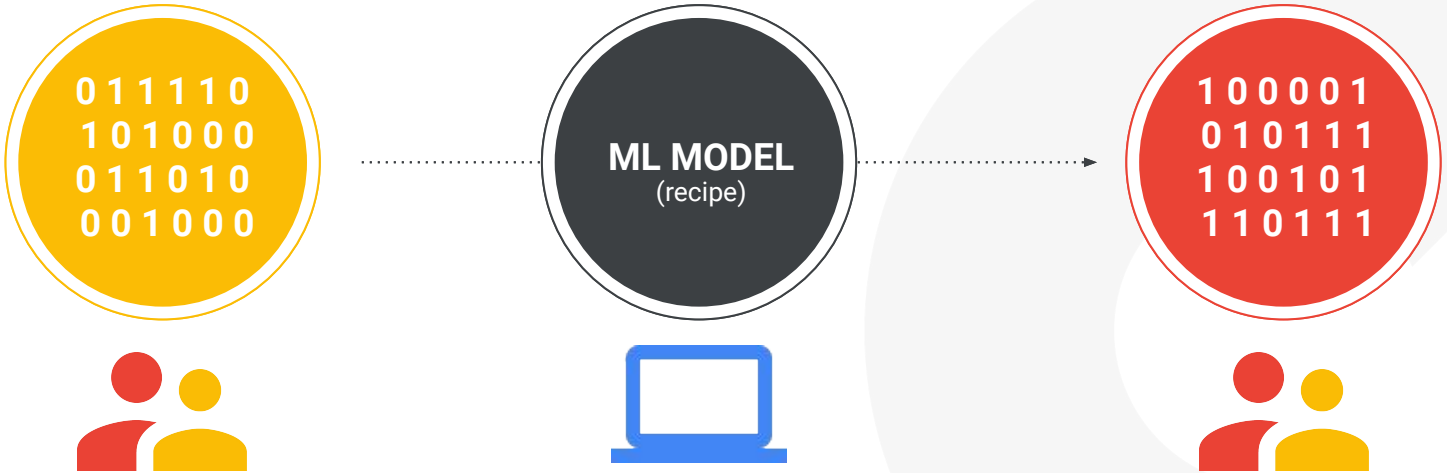


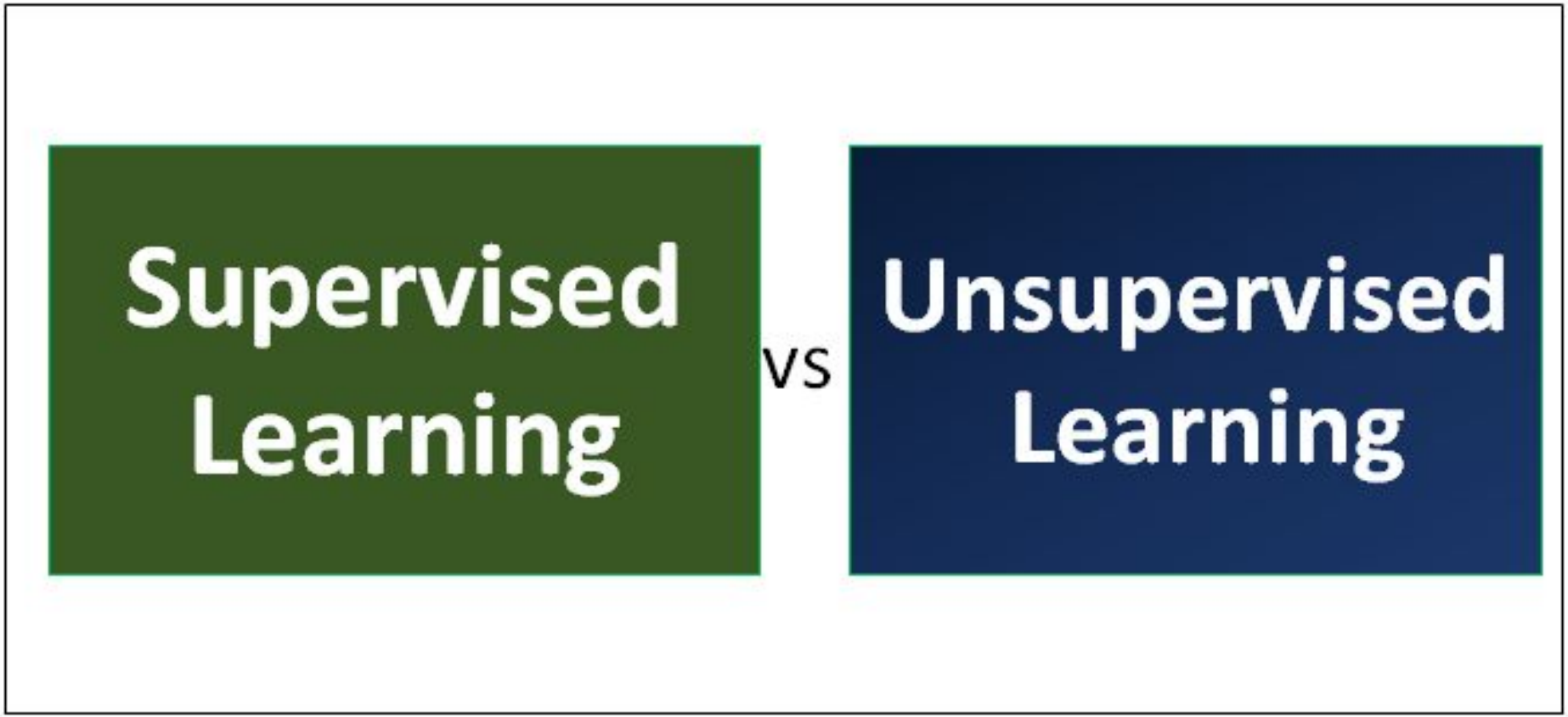
Machine learning: Recipe learned from data



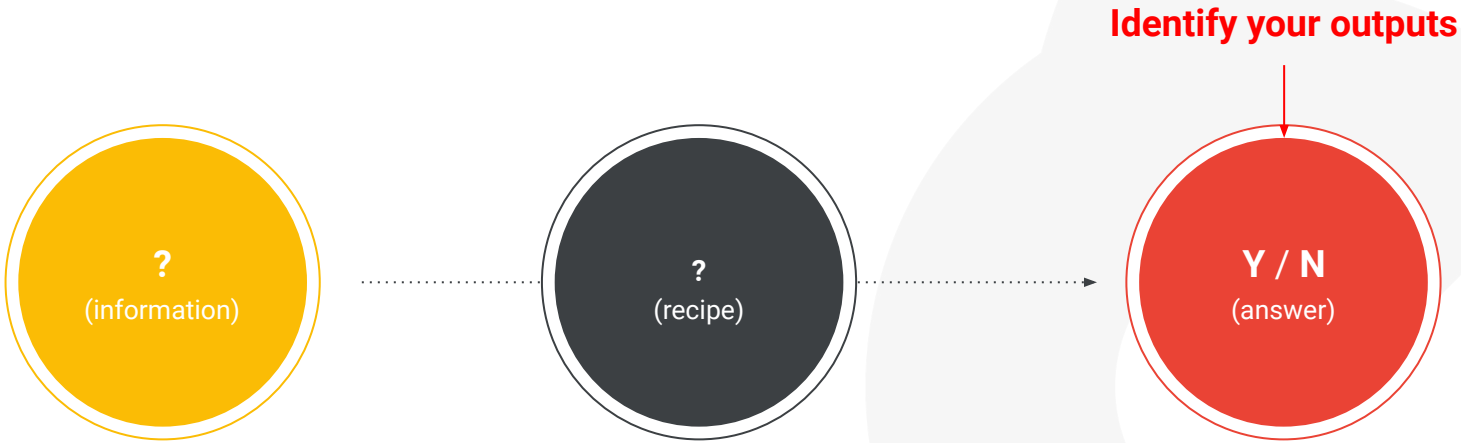


Machine learning: Recipe learned from data

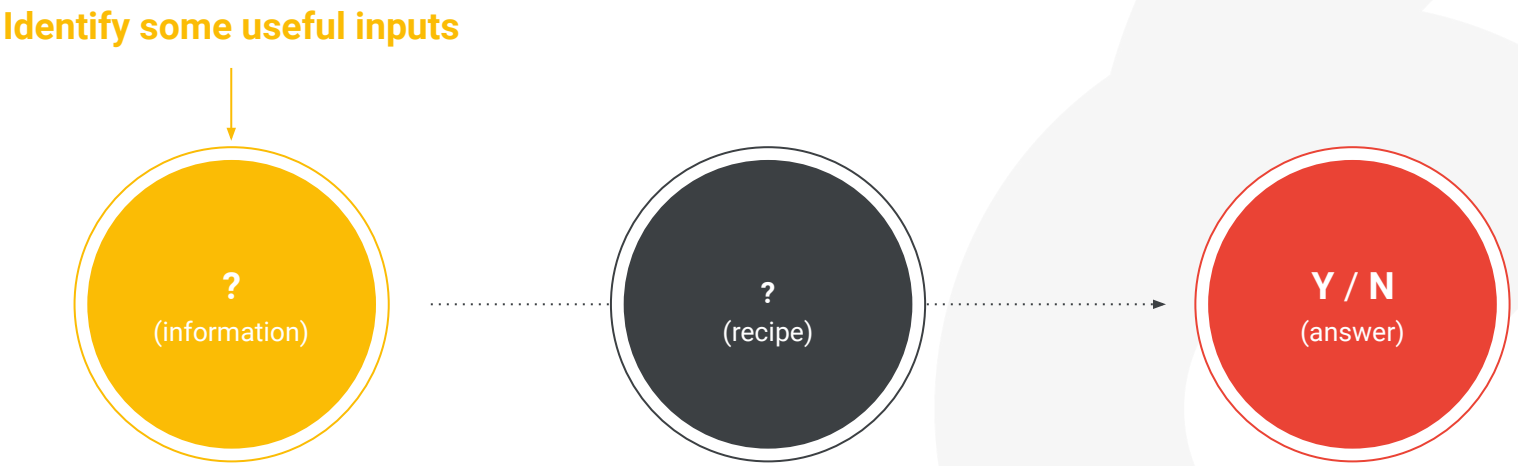



Supervised Learning
vs
Unsupervised Learning

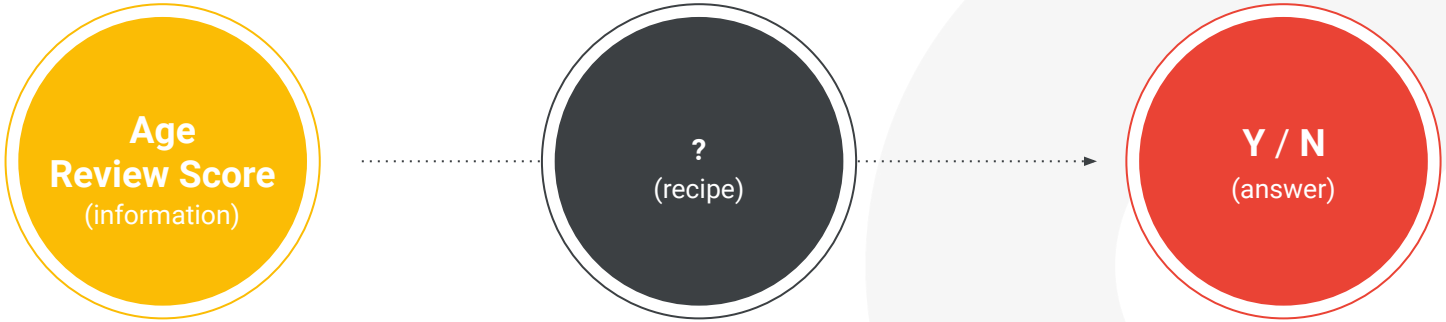
Define our outputs



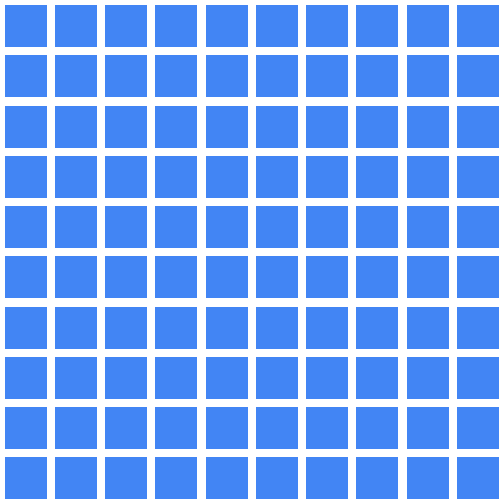
Get some data to learn from



Get some data to learn from

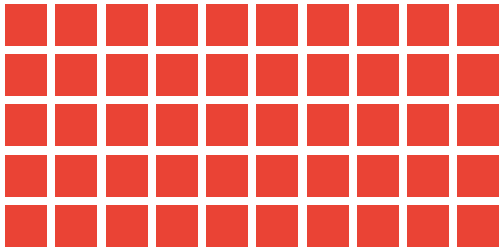


Split your data



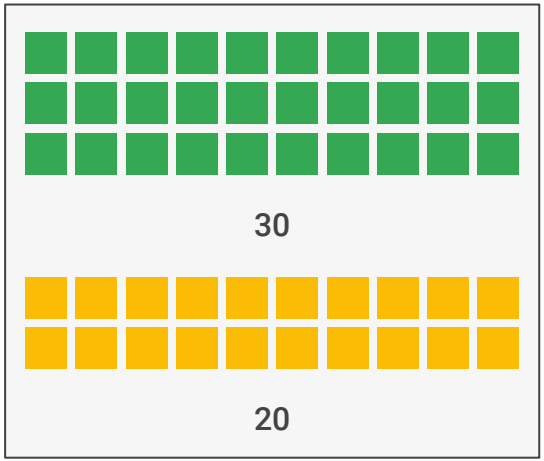
Training set

Fitting the model



Validation set

Soft interim evals

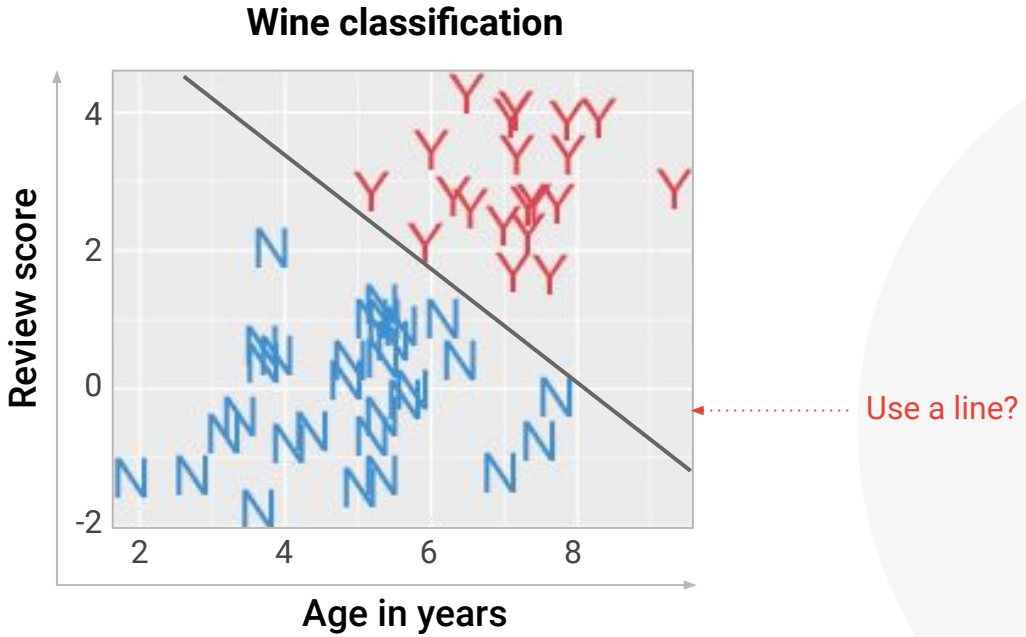


Test set

Rigorous final eval

Data inputs

Use a line?


Wine classification
Review score
4
2
0
-2
2
4
6
8
Age in years
Use a line?

Train your solution

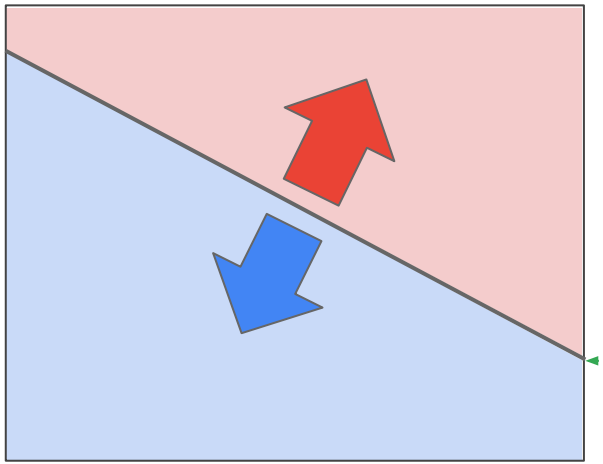
Algorithm selection:

Support Vector Machine

Decision Tree

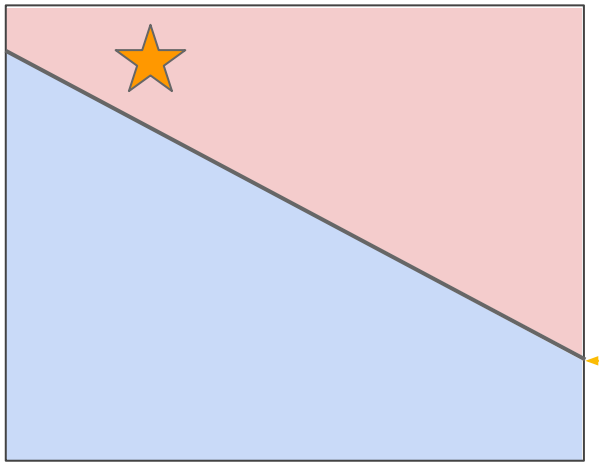
Neural Network

Train your solution



Get model (recipe)

New Data



Decision Tree Training

Algorithm selected:

Decision Tree

Run ML algorithm

Decision Tree Training

Algorithm selected:

Decision Tree

Run ML algorithm

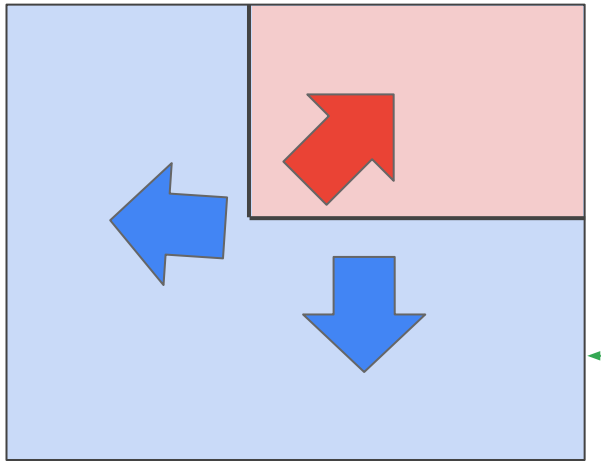
Decision Tree Training

Algorithm selected:

Decision Tree

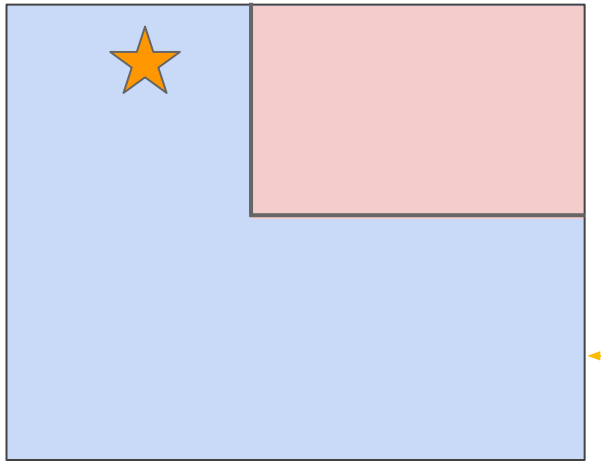
Run ML algorithm

Decision Tree Model



Get model (recipe)

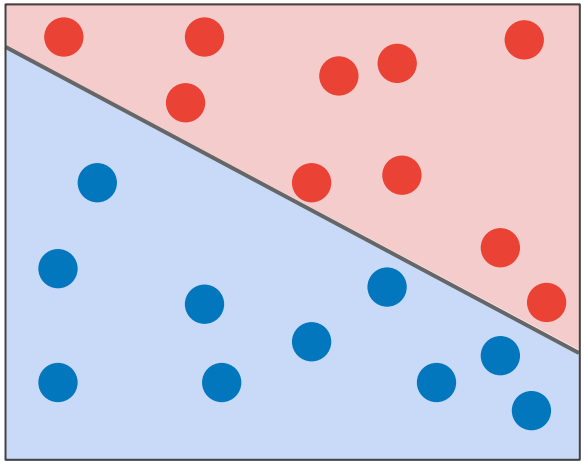
Validate your solution



New data? Use recipe!

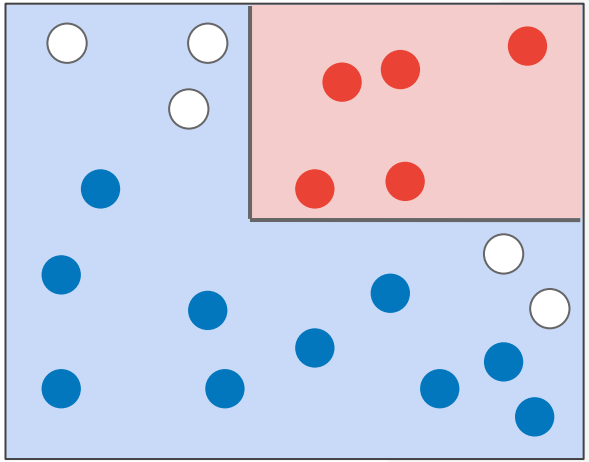
Get model performance

Support vector machine



Test Accuracy: **100%**

Decision tree



Test Accuracy: **75%**

EE Classifier Limitations

- Good for about 100mb of training data, so with about 100 bands that gives about 200k points
- Feature collection **cannot** be sparse
- Generally small number of classes

Follow along!

- https://code.earthengine.google.com/?accept_repo=users/akarbasi/simple_classifier

Common EE Errors

- Out of memory error
  - Usually happens when tile is too large. When you sample from an image you're sampling from a tile. In Earth Engine each tile 1000x1000 so with a lot of bands that can be too much memory for us to handle.

- Computation timed out
  - Can happens if classifying the table/image takes too long

- Request too large
  - Too large of a training dataset or the model created is too large.

Data inputs - Clustering

Data inputs - Clustering

Links

- ML Overview: bit.ly/quaesita_emperor
  - By Cassie Kozyrkov
- Classifier examples:
  - http://code.earthengine.google.com/?accept_repo=users/akarbasi/simple_classifier