Random Forest

Megan Ruffley, Isaac Overcast

CompPhylo Oslo 2019

August 27, 2019

What is random forest?

- The forest is made up of decision trees
- Supervised machine learning
- Ensemble approach for both
  - Classification (discrete)
  - Regression (continuous)
- Random

Breimen L. (2001) Random Forests. *Machine Learning*, 45, 5-32.

The forest is made up of decision trees

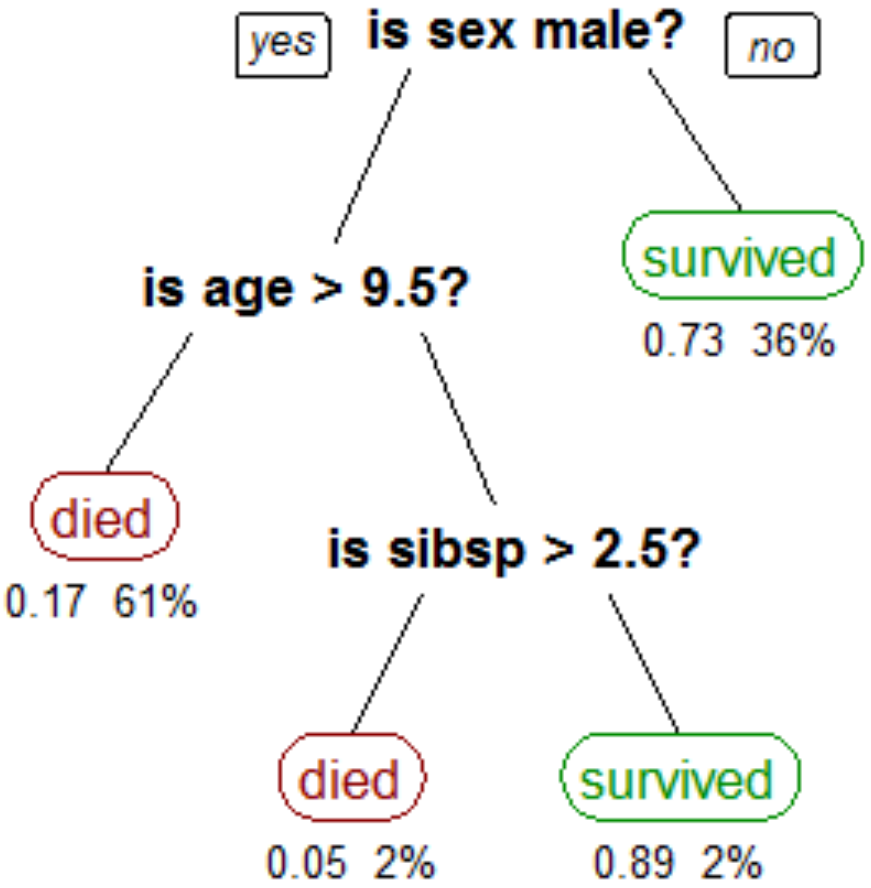
- There are two types of decision trees
  - Classification trees
  - Regression trees

- CART (classification and regression trees)

The forest is made up of decision trees

- There are two types of decision trees
  - Classification trees
  - Regression trees

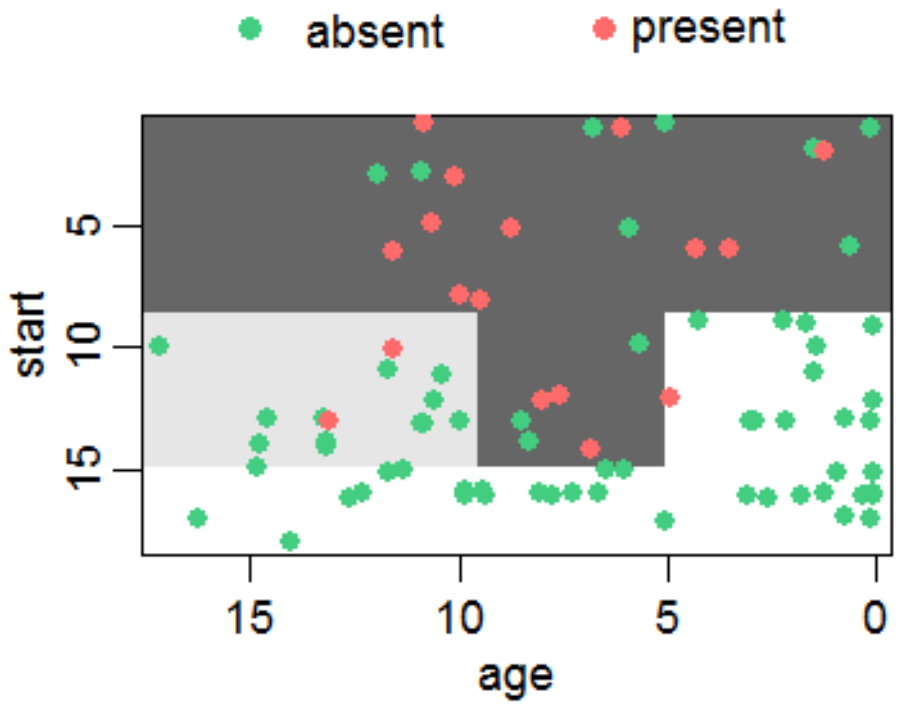
- CART (classification and regression trees)

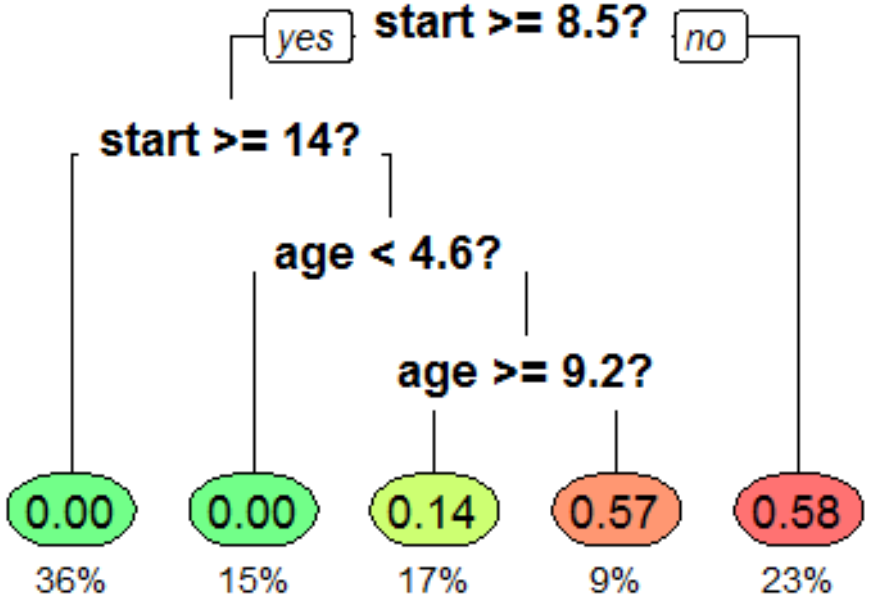


By Stephen Milborrow - CC BY-SA 3.

The forest is made up of decision trees

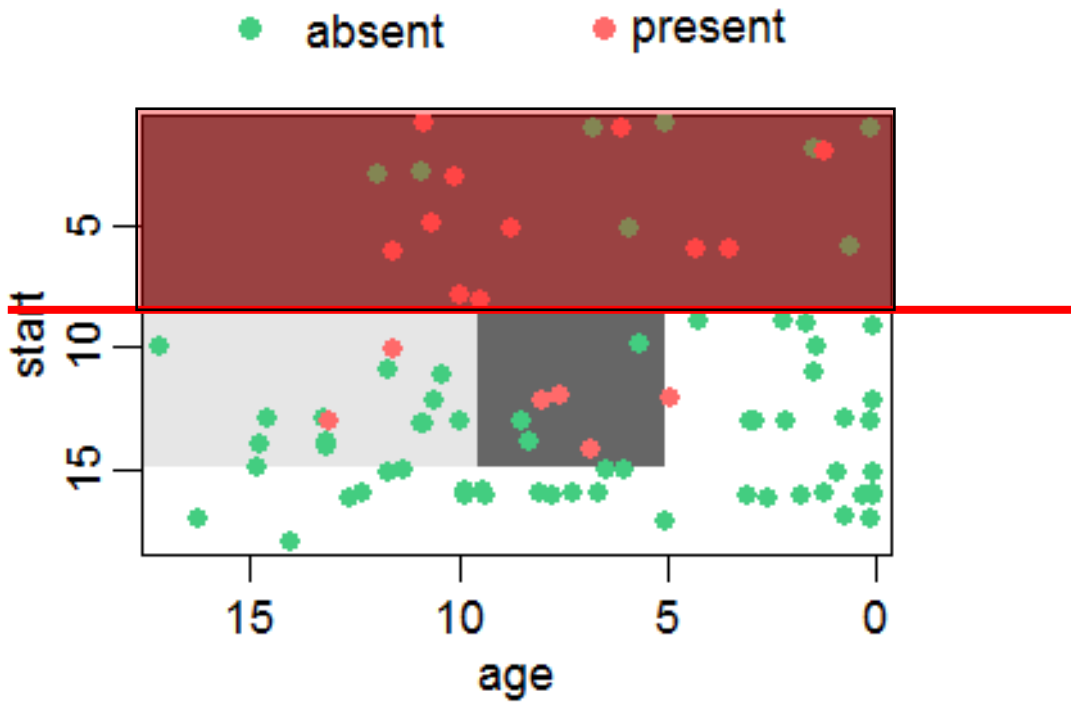
- There are two types of decision trees
  - Classification trees

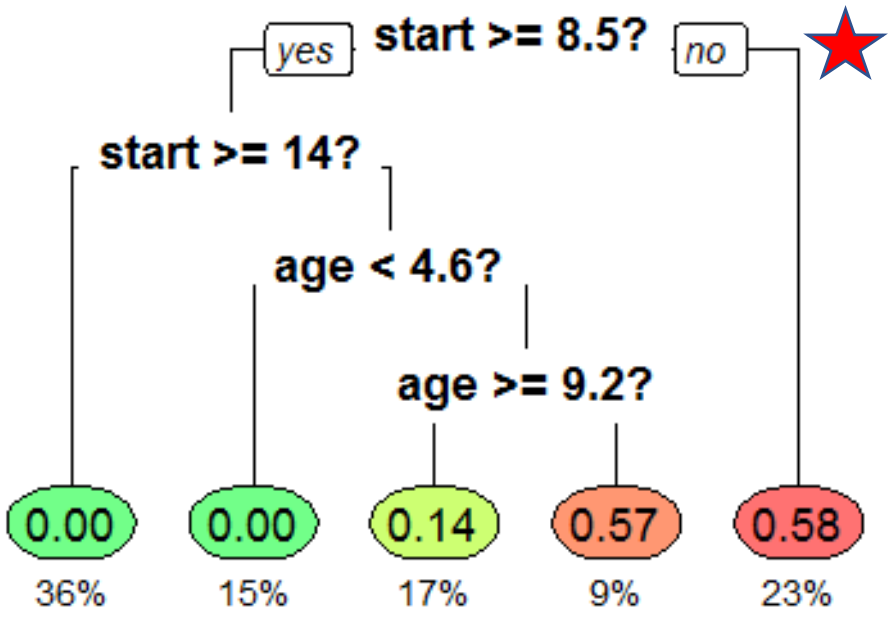




The forest is made up of decision trees

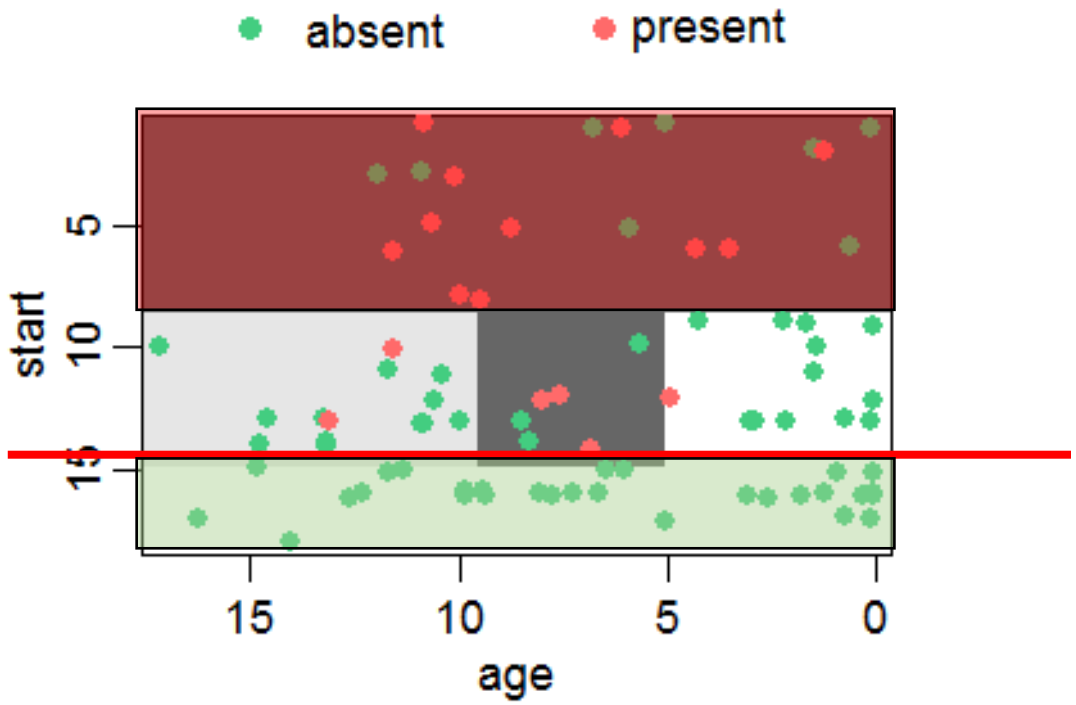
- There are two types of decision trees
  - Classification trees

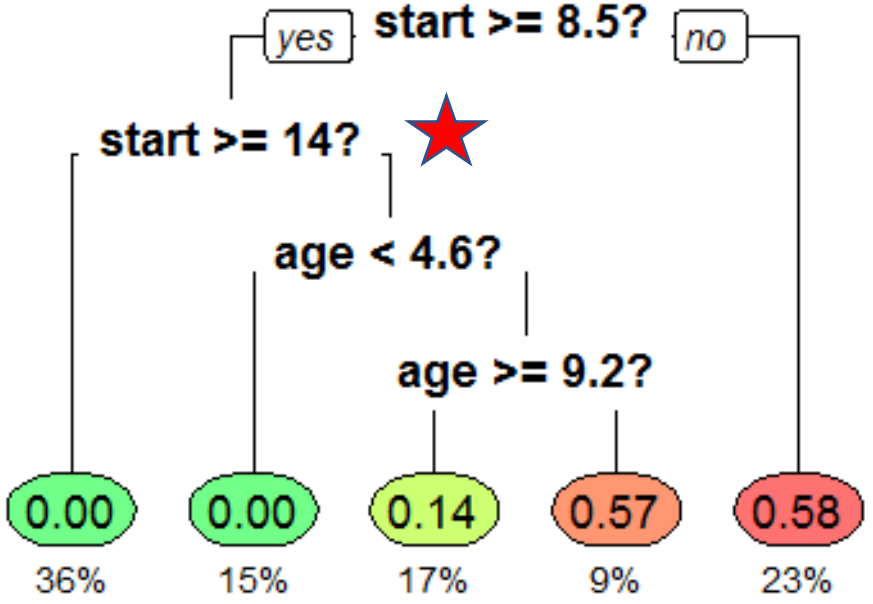




The forest is made up of decision trees

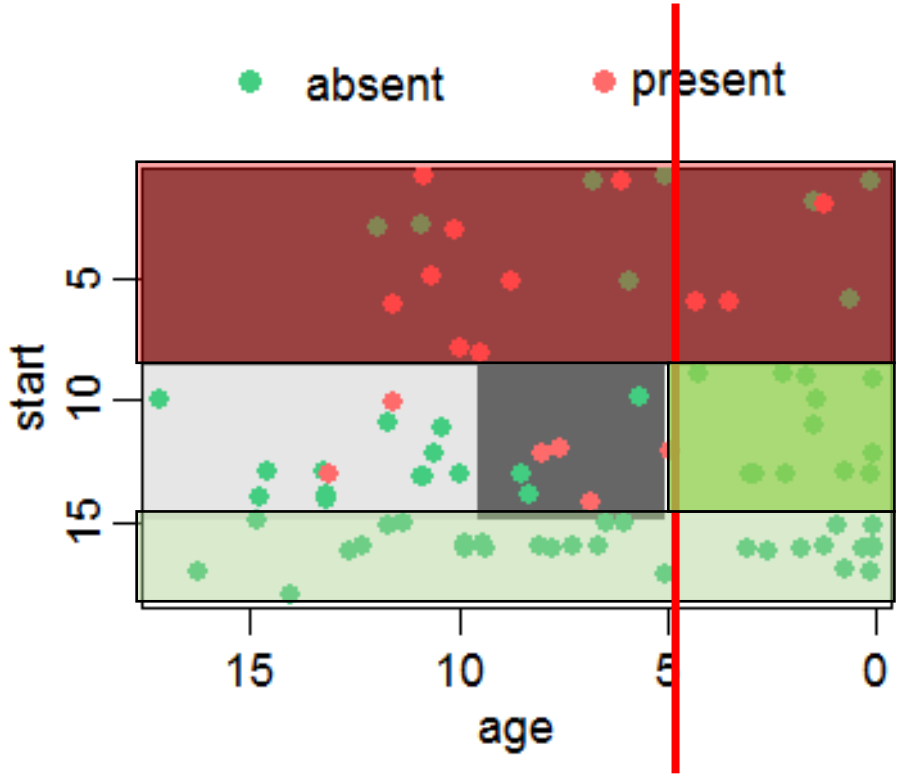
- There are two types of decision trees
  - Classification trees

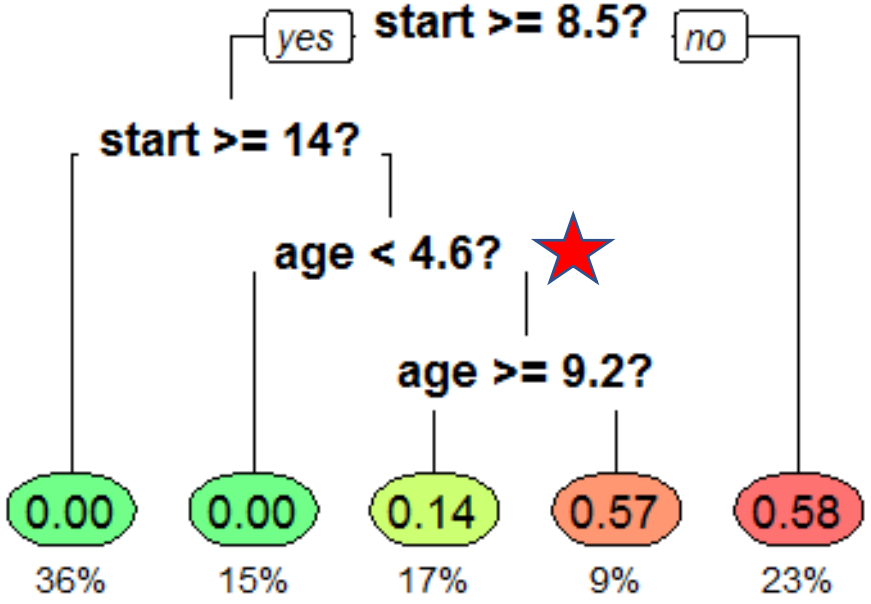




The forest is made up of decision trees

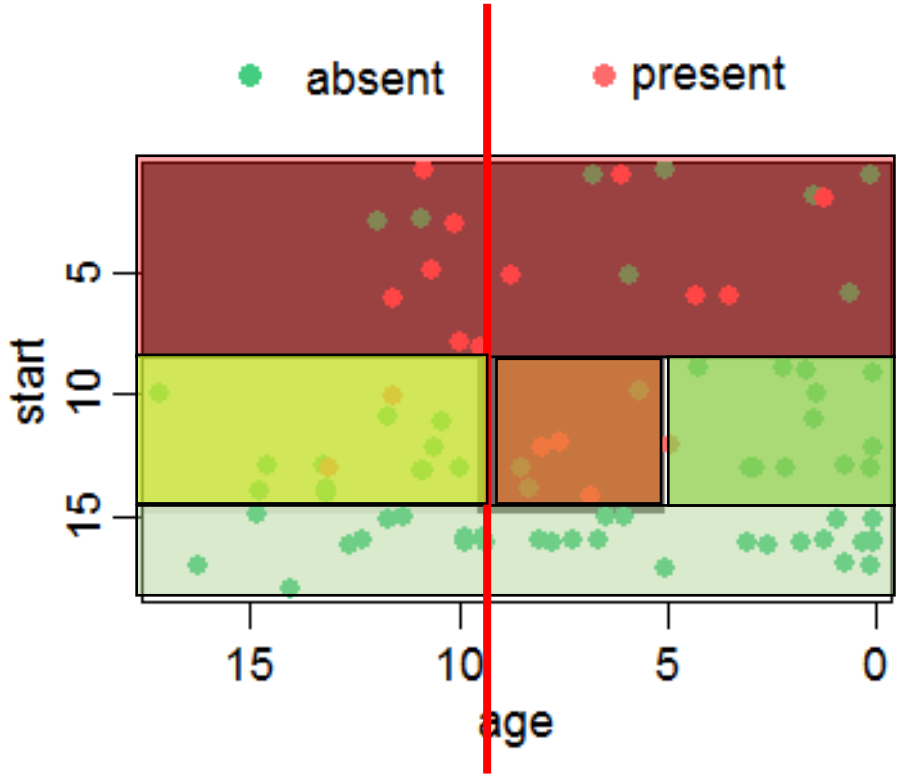
- There are two types of decision trees
  - Classification trees

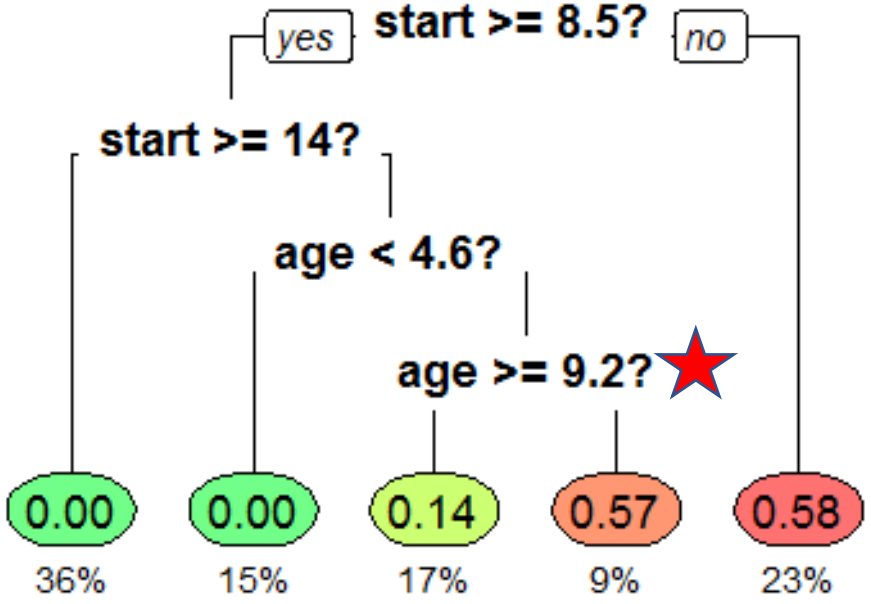




The forest is made up of decision trees

- There are two types of decision trees
  - Classification trees





The forest is made up of decision trees

- There are two types of decision trees
  - Regression trees
  - These are a little bit more complicated. We will get into them later.

What is random forest?

- The forest is made up of decision trees
- Supervised machine learning
- Ensemble approach for both
  - Classification (discrete)
  - Regression (continuous)
- Random

Breimen L. (2001) Random Forests. *Machine Learning*, 45, 5-32.

Supervised machine learning

- Trains a function that, given a sample of data and desired outputs, best approximates the relationship between input and output observable in the data.
- Required prior knowledge of what the output should be
- Two main types of supervised learning….

Unsupervised learning, on the other hand, is untrained and infers the natural structure present within a set of data points.

Supervised machine learning

- Trains a function that, given a sample of data and desired outputs, best approximates the relationship between input and output observable in the data.
- Required prior knowledge of what the output should be
- Two main types of supervised learning….
  - ***Classification***
  - ***Regression***

Unsupervised learning, on the other hand, is untrained and infers the natural structure present within a set of data points.

Supervised machine learning

- Two main types of supervised learning….
  - ***Classification***
  - ***Regression***

- Common algorithms include random forests, neural networks, logistic regression, and support vector machines.

Unsupervised learning, on the other hand, is untrained and infers the natural structure present within a set of data points.

Supervised machine learning

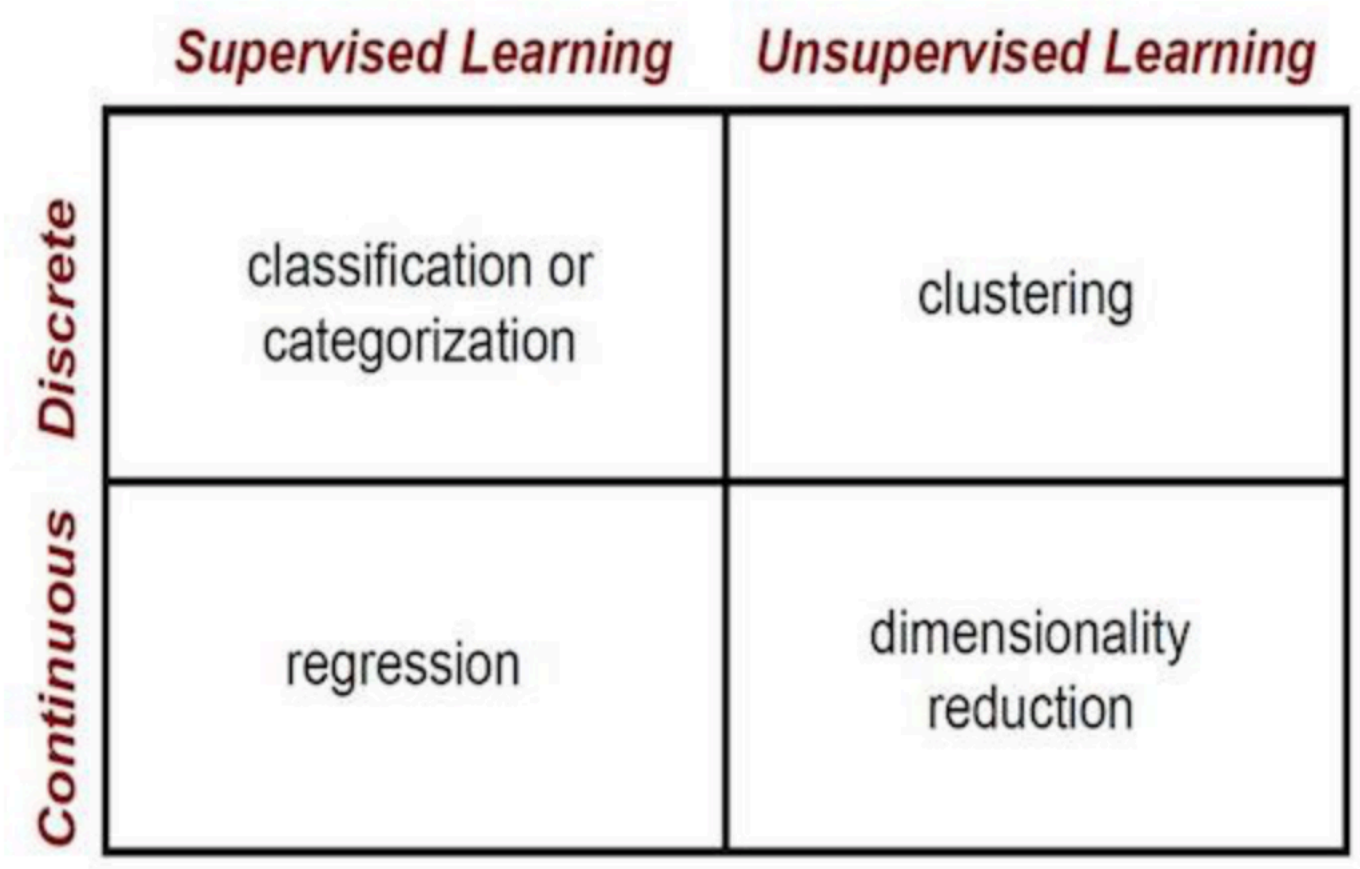
- Two main types of supervised learning….
  - ***Classification***
  - ***Regression***

- Common algorithms include random forests, neural networks, logistic regression, and support vector machines.

Unsupervised learning, on the other hand, is untrained and infers the natural structure present within a set of data points.

- Mainly for clustering and dimensionality reduction.

Supervised machine learning



What is random forest?

- The forest is made up of decision trees
- Supervised machine learning
- Random
- Ensemble approach

Breimen L. (2001) Random Forests. *Machine Learning*, 45, 5-32.

What is random forest?

- The forest is made up of decision trees
- Supervised machine learning
- Random
- Bagging
- Out of bag error

Breimen L. (2001) Random Forests. *Machine Learning*, 45, 5-32.

What is random forest?

- The forest is made up of decision trees
- Supervised machine learning
- Random
- Ensemble approach