KNN and Model Persistence

TM Quest

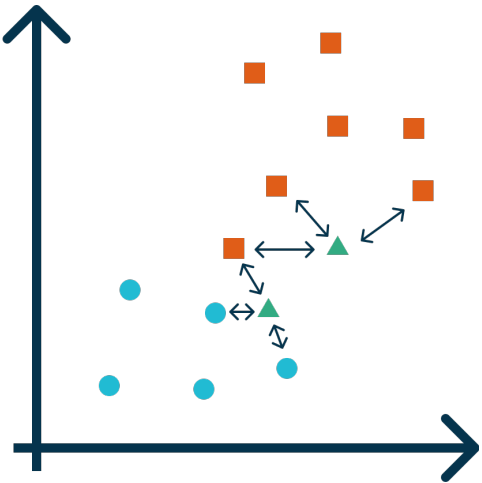
Overview

What Will we Learn in This Module?

- What is the K-nearest neighbors (KNN) model?
  - The idea behind KNN.
  - How to implement a KNN model in scikit-learn.
- What is model persistence?
  - How to save your model as a file.

K-Nearest Neighbors

Idea



3-Nearest Neighbors

- Two categories ● and ■.
- New points ▲.
- Then we look at the three nearest neighbors and choose the category with the most votes.
- We can choose how many neighbors we are going to look at.

Model Persistence

Motivation

What is Model Persistence?

Saving the model after training.

Why?

- Don’t need to train the model again.
  - Can take hours to train the model.
- The model does not change due to chance.

Model Persistence

Pickle

- Can save any Python object.
- Is included in Python's standard library.
- Can be faster for several smaller objects.

Joblib

- Can only save NumPy objects (like models).
- Need to be installed separately.
- Fast for large NumPy objects.

Warning

- Neither are secure.
- Don't load in data you don't trust.