



Bookmarks

- ▶ Start Here
- ▶ 1. The Big Picture
- ▶ 2. Data And Features
- ▶ 3. Exploring Data
- ▶ 4. Transforming Data
- ▶ 5. Data Modeling

▼ 6. Data Modeling II

Lecture: SVC

Quiz



Lab: SVC

Lab



Lecture: Decision Trees

Quiz



Lab: Decision Trees

Lab



6. Data Modeling II > Lab: SVC > Assignment 1

Assignment 1

🔖 Bookmark this page

Lab Assignment 1

In this lab, you'll get started with support vector classifier by revisiting UCI's wheat-seeds dataset. First, you will benchmark how long it takes to train and predict with SVC relative to how long K-Neighbors took to train and test, and then you'll compare the decision boundary plot produced by the two.

1. Start by reviewing the starter code located in Module6/**assignment1.py** and of course the wheat dataset in Module6/Datasets/**wheat.data**. Be sure to read through **everything**, so that you know exactly what is being asked of you...
2. Add in the KNN and SVC classifiers
3. Load up your dataset properly and prune it
4. Run the convenience benchmark and plotting functions
5. Then, answer the following questions.

Lab Question 1

1 point possible (graded)

In this lab, you run KNeighbors and SVC in high dimensional feature space, 5000 times! Ensure you don't have any spyware, movies, games, video editing, or other hardware intensive apps running while you execute the timed runs.

Lecture: Random Forest

Quiz

**Lab: Random Forest**

Lab

**Dive Deeper**

► 7. Evaluating Data

► Course Wrap-up

After the high dimensional classification and benchmarking finishes, the lab plots for you a matrix of 2D, feature vs feature plots against one another. This way, you can see how each pair of features stack against the other. Since a square matrix is produced, there are two copies of each plot, although the axes are swapped, and that is controllable by the FAST_DRAW parameter.

Which of the four tests produced the highest scoring result?

☐ KNeighbors High Dimensionality Score☐ KNeighbors "Best" 2D Score☐ SVC High Dimensionality Score☐ SVC "Best" 2D Score**Submit**

You have used 0 of 2 attempts

Lab Question 2

1 point possible (graded)

Change iterations from 5000 to 200,000 and then re-run your lab with the default parameters.

Ensure you don't have anything else processor intensive running on your machine for the duration of time it takes to do its number crunching (should be 1-10 minutes on modern laptops).

Which of these four computations was **fastest** to calculate?

Which of these four computations was **fastest** to calculate?

☐ KNeighbors Training Time

☐ KNeighbors Testing Time

☐ SVC Training Time

☐ SVC Testing Time

Submit

You have used 0 of 2 attempts

© All Rights Reserved



© 2016 edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open EdX logos are registered trademarks or trademarks of edX Inc.

POWERED BY
OPENedX®



