edX   **Microsoft:** DAT210x Programming with Python for Data Science

📑 **Bookmarks**

🔖 **Bookmark**

▸ 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   ✎

Review Question 1

(1/1 point)
The purpose of the kernel function in support vector machines is to...

○ Intelligently change your feature space from linear to polynomial

○ Increase the dimensionality of your data set

◉ Compute how similar two samples are in the feature space provided to it ✔

○ Decrease the dimensionality of your data set, making it simpler to linearly separate your higher-dimensionality data

EXPLANATION

The kernel function doesn't exactly increase your dataset's dimensionality. Rather, it allows you to act as if you did, and moreover, it allows you to inspect how similar two samples are. Kernel functions are usually inner products, and at least the built-in ones in SciKit-Learn are almost

entirely dot products functions. These allow you to measure how similar (or dissimilar) two samples are by projecting one onto the other in their feature space.

You have used 2 of 2 submissions

Review Question 2

(1/1 point)
SciKit-Learn's SVC class takes in many parameters. Three of the parameters that end of contributing the most to your overall decision boundary shape are gamma, C and the kernel.

Assume you're using the rbf kernel and, not-surprisingly, you like the overall shape of your decision boundary. However you notice the decision boundary 'pockets', or has medium sized bubbles, which surround one class of your samples. Those pockets are simply too large.

How would you alter your parameters fix this?

○ Change rbf kernel to linear kernel, so that no bubbles form anymore

◉ Reduce the gamma coefficient to get rid of the bubbles ✔

○ Increase C to get rid of the bubbles

○ Increase the gamma coefficient to get rid of the bubbles

EXPLANATION

Either decreasing C or decreasing gamma would help with this. Gamma would have even more of an immediate effect, since it's the rbf coefficient.

You have used 1 of 2 submissions