

Peer-graded Assignment: Reflecting on Your Machine Learning Project

Submit by Aug 17, 12:29 PM IST

🔔 It looks like this is your first peer-graded assignment. [Learn more](#)



Submit your assignment soon

Even though your assignment is due on Aug 17, 12:29 PM IST, try to submit it 1 or 2 days early if you can. Submitting early gives you a better chance of getting the peer reviews you need in time.

Instructions

My submission

Discussions

Project Title *

Give your project a descriptive title

Describe your machine learning project, in terms of the following:

1. What problem were you trying to solve?
2. What data did you collect?
3. What classes did you use to classify your data, and why they were chosen?

I am trying to solve IRIS DATA-SET (It is the most common problem in Machine Learning)

Overview

The data set consists of 50 samples from each of three species of Iris (Iris setosa, Iris virginica and Iris versicolor). Four features were measured from each sample: the length and the width of the sepals and petals (in centimeters)

Column name : Sepal Length , Sepal Width , Petal Length , Petal Width

Based on the combination of these four features

1. I am trying to predict Class of Flower (Train Model to distinguish the species from each other.)
2. It is a multivariate data set ie (variation of Iris flowers of three related species) It can be found on UCI's Machine Learning

Repository

1. class 0: Iris setosa (ie , Setosa)

class 1: Iris virginica(ie, Virginica)

class 2: Iris versicolor (ie, Versicolor)

B I | 🔗 🔄 | ☰ ☷

This part of the question is about evaluating your project, in terms of the following:

1. How well did your classifier work? Were you happy with the results? What was the overall percentage accuracy on test data?
2. Describe what types of Images your models classifies well and which they classify badly.
3. Explain why you think it performed well or badly on the Images you described in the last part.
4. Were there problems with the classifier that you were able to solve? Describe your strategy for solving the problem

I used **K-Nearest Neighbor(KNN)** classification algorithm to predict the output.

Libraries used

1. numpy
2. pandas
3. matplotlib.pyplot
4. sklearn

1. Classifier works quite well, Yeah , I am happy with the results as the overall accuracy was 0.97777 (97.77%)

2. The model classify Setosa well , but faces some problem in other two classes(Virginica , Versicolor)

3. It performs well because it gives 97% accuracy on test set .

4. At the beginning I use n_neighbors = 1 ie (the value of knn) , test set =30% , training set=70% :

The model give Accuracy of 0.97777 (97.77%)

n_neighbors = 5 ie (the value of knn) on the same test and training set :

The model give Accuracy of 0.97777 (97.77%)

but on increasing test to 40% and decresaing training set to 60% :

The accuracy were as follows :

knn = 1 ----- accuracy : 0.95555 (95.55%)

knn = 5 ----- accuracy : 0.96667 (96.67%)



B *I* | |

☒ I, **Piyush Sambhi**, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.

[Learn more about Coursera's Honor Code](#)

This project has not been saved.

[Save draft](#)

[Preview](#)

