

Haonan Hu

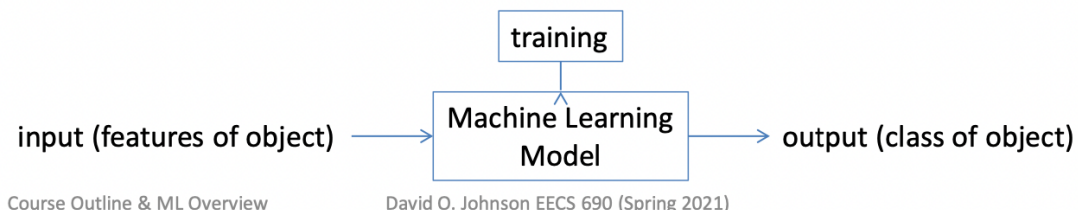
2/2/2021

2863545

In-Class Problem

- The “hello world” machine learning application is the Iris flower data set or Fisher's Iris data set.
 - This dataset is utilized to identify 3 species of the iris flower: Iris Setosa, Iris Versicolour, and Iris Virginica by the length and the width of the sepals and petals, in centimeters.
 - Assume we are going to design a machine learning application that identifies species of the iris flower using this dataset.
1. List the input features to the Machine Learning model.
 2. List the output classes of the object.
 3. Describe how we might train the Machine Learning model.
 4. Because no Machine Learning model can classify objects perfectly, describe how we might measure how accurate various ML models solved the problem.

- You may work together.
- You may ask me or the SI for help.
- Turn in your problem when you are done:
- PDF submitted to BlackBoard before 11:59 PM
- You may leave when you are done.



1. List the input features to the Machine Learning model.

Length, width, sepals and petals of flowers that we want to identify

2. List the output classes of the object.

Iris Setosa, Iris Versicolour, Iris Virginica

3. Describe how we might train the Machine Learning model.

Have a large set of identified Iris flower data, and get average length, width, petals and sepals of each species. Have another large set of unidentified Iris flower data and try to identify them through comparing each of the input with calculated average, get rid of the mistake identification and update average data if the identification is successful.

4. Because no Machine Learning model can classify objects perfectly, describe how we might measure how accurate various ML models solved the problem.

Saying that if we have 10000 (the more data you have the more accurate) identified Iris flowers data, and feed the data to the machine learning model, if

the model analyzed result is matched with flower's actual species, then count towards to successful counter, after all flowers are identified, calculate the accuracy based on how many flowers are successfully identified among the 10000 flowers. Then repeat the process several times and calculate the average accuracy.