CMPE 407

March 15, 2022

Exercises

- Download iris.csv dataset and read by pandas library.

- Investigate the dataset. Observe what is the data types of features, how many data that datataset contains etc.

- Access to 10th data of dataset. What is the difference of using pandas.loc and pandas.iloc functions to get the data given index?

- Are all features convenient to be calculated by machine learning algorithms? If they are not, how can we cope with these issues? Try to fix it.

- Select the features that are appropriate for modelling. Prepare the data in order to feed to model.

Problems

- Download MallCustomer.csv dataset and read.

- Investigate the dataset. Observe what is the data types of features, how many data that datataset contains etc.

- Apply label encoding for the features that are not in processable.

- Select the features that are appropriate