

Machine Learning Exercises - III

September 30, 2021

Kannan Singaravelu

kannan.singaravelu@fitchlearning.com

* * *

Exercise 5

- a. What is Gini Impurity?
- b. When will the nodes record its highest and lowest impurity measure?
- c. Plot the Gini Impurity Measure for two classes.

Exercise 6

- a. What is Anomaly Detection? List down the popular learning algorithms used for anomaly detection.
- b. What is the Isolation Forest algorithm? How do Isolation Forests work?
- c. For this exercise, use the credit card dataset from <https://www.kaggle.com/mlg-ulb/creditcardfraud>. The dataset contains thirty predictor variables and one binary response variable called class. Build a model to detect credit card fraud using the Isolation Forest Anomaly Detection algorithm.

[Kannan Singaravelu](#)

Certificate in Quantitative Finance, Jan 2021