**Microsoft Student Accelerator Program Australia 2020**

**Western Sydney University MICT**

**Hiroki Seino** [**20008068@student.westernsydney.edu.au**](mailto:20008068@student.westernsydney.edu.au)

**MSA 2020 - AI & Advanced Analytics - Coding Approach**

**Part-1**

Microsoft Learn Learning Path for machine learning

[**https://docs.microsoft.com/en-gb/users/hirokiseino-9025/**](https://docs.microsoft.com/en-gb/users/hirokiseino-9025/)

**Part-2**

**The Challenge**

**◆Build a machine learning model to classify the species of an iris flower based on its size of petal and sepal. Random forests make many different decision trees. The correct answer is the label with the highest probability from the results predicted by all decision trees. Individual trees may be overfitting, but aggregating many results has the effect of suppressing overfitting.**

◆**Data Set Characteristics:**

Number of Instances: 150 (50 in each of three classes below)

Number of Attributes: 4 numeric, predictive attributes and the class

Attribute Information:

- sepal length in cm

- sepal width in cm

- petal length in cm

- petal width in cm

- class:

- ‘Iris-Setosa’

- ‘Iris-Versicolour’

- ‘Iris-Virginica’

**A close up of a map

Description automatically generated**

**◆It can be said that the dataset is relatively easy to classify as the same varieties are solidified regardless of the feature pairs.**

**◆After using fit method to learn,** **the important thing for machine learning is that when data is given, it can be predicted, and its prediction accuracy is high.**