Neural Network

Dataset

The dataset used in training the model is called the Seeds Dataset from UCI Machine Learning Repositories (source: https://doi.org/10.24432/C5H30K). It is a classification dataset with three classes. The dataset consists 210 instances in each of the 7 features namely "area", "perimeter", "compactness", "kernel_length", "kernel_width", "asymetry_coefficient", and "length_of_kernel_groove." Each of these features has numeric data with varying values each therefore it needs to be scaled first in order to be used in training the model. The target variable named "classification" has three values: 1, 2, and 3. These values were converted to categorical values in order to be used in the model.

Results

A Sequential model from keras library was created and trained using the processed dataset described above. It is composed of three layers: the input layer that is a Dense layer with 7 neurons, one hidden layer also a Dense layer but with 14 neurons, and the output layer that is a Dense layer with 3 neurons which serves as the 3 classes. The model was compiled and fitted with 'adam' optimizer and categorical_crossentropy for the loss. The training has 20 epoch and batch size of 2. The model showed a good performance with around 93% accuracy and 20% loss on the 20th epoch. When the model was evaluated, it displays an accuracy of around 98% and a loss of 15%.

The image below shows a confusion matrix between the predicted class of the model and its actual class.