**Applied Machine Learning**

**Assignment 2**

|  |  |
| --- | --- |
| **Name** | YAP LI XEN |
| **Student ID** | P7414389 |
| **Lecturer** | MR. PAUL |

**PART B: DEEP LEARNING**

1. **How is your prediction task defined? And what is the meaning of the output variable?**

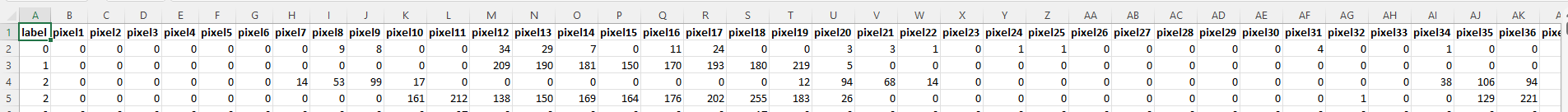
* The prediction task is to predict which category the image belongs to (to classify image)
* The output variables “Label” consists of 10 values (refer to screenshot below)
* This is multi-class variables

Table

Description automatically generated

1. **How do you represent your data as features?**

* Target: label
* Features: All columns except label. There are total of 784 columns (28 pixels \* 28 pixels)



1. **Did you process the features in any way?**

* Reshaping

Text

Description automatically generated

Graphical user interface, text, application

Description automatically generated

* Scaling



* Encoding

Graphical user interface, text, application

Description automatically generated

1. **Did you bring in any additional sources of data?**

* No. Not necessary

1. **How did you select which learning algorithms to use?**

* Convolutional Neural Network (CNN) is primarily used for image processing and recognized by many in the industry.

Table

Description automatically generated

1. **Did you try to tune the hyperparameters of the learning algorithm, and in that case how?**

* No

1. **How do you evaluate the quality of your system?**

* Score the trained model using test data
* We can see the Accuracy Score and Error Score from the trained model while predicting the test data





1. **How well does your system compare to a stupid baseline?**
2. **Can you say anything about the errors that the system makes? For a classification task, you may consider a confusion matrix.**
3. **Is it possible to say something about which features the model considers important? (Whether this is possible depends on the type of classifier you are using)**