LABORATORY 10: Embedded Artificial Intelligence

Part 1. Tenseorflow and keras

To work with artificial intelligence some APIs that will help us to ease the development. This APIs are **Tensorflow** and **Keras.**

An Application Programming Interface is a platform that will provide the user specific data according to the requests the user made. This means that an API will

What is an API? How does it work?

Which is the difference between tensorflow and keras?

To install tensorflow and keras as a standalone package you can visit the tensorflow main page:

https://www.tensorflow.org/install

Install Tensorflow GPU and CUDA on Linux Ubuntu, you can use the following video as a guide to complete the installation:

https://www.youtube.com/watch?v=dj-Jntz-74g&ab_channel=JeffHeaton

Part 2. ANN

We can implement artificial neural networks in python with help of different libraries, APIs and sometimes frameworks. For this, is important to know some of the most common tools used during programming:

- Numpy: It will help us to generate arrays math operations easily.
- Pandas: Helps us to manipulate customer data from different sources like excel files.
- Tensorflow – Keras: Main tool to generate ANN models, train them and make predictions.

To review a simple example of an ANN you can access to the following google collab page:
https://colab.research.google.com/drive/1elodGsgRqJ5U4W-uBi0ccFBLkgXauxdo?usp=sharing

What is an ANN network? Explain the elements it has.

What is a single layer Neural Network? What is multilayer NN?

When an ANN should be used? When is not recommended?

- **Exercise 1**

Given the following equation:

$$Y = 3X + 2$$

Generate an ANN to obtain an approximate result for the following values: X = 5, X = 3.3.

- **Exercise 2**

With the model generated in Exercise 1 modify the number of layers, neurons, and the type of activation functions to check if the results can be improved. What is overfitting? How it can be produced?

What is optimizer? Describe some of the existing optimizers.

What is loss function? Describe some of the existing loss functions.

- **Exercise 3**

Given the 'spam text data' dataset generate a model that can predict if a text is spam or not.

How does pandas work? What is it use for?

How can you use pandas to extract data from a dataset?

- **Exercise 4**

Generate an excel spreadsheet with the following columns:

- Student Name
- Student Age
- No. of Lab completed.
- Average score

Then using Pandas extract the information from the excel file and print it. Use Functional programming.

Part 3. Clustering

Clustering is a non-supervised artificial intelligence technique that will allow us to group different data representations according to some information relationship. Generally, this is done by groups of data that will help to associate different information.

You can check a simple example of clustering in the following google collab spreadsheet:

https://colab.research.google.com/drive/1oZkLnUbhZcyqSatXqNyESu7QIwCG79PX?usp=sharing

What is clustering? What is classification?

When do we use clustering and classification?

How do we use clustering with K-means?

Which is the difference between K-means and K-means ++?

- **Exercise 5**

A vet shop wants to determine if they should start selling "Vet shop cards" for their customers purchases. With this purpose thy have the following set of data:

Average purchase cost (bs)	Average # of purchases with credit card
750	3
1245	1
230	4
533	3
490	2

1000	**1**
190	**0**
900	**3**
600	**2**
50	**1**
1100	**0**
930	**4**
450	**3**
330	**2**
750	**0**

Use K-Means method to determine the best suited costumers to buy a "Vet shop card". Does it worth to generate this additional payment method? Why?

How did you choose the number of clusters? What happens if you try to use 1 more cluster? And with 1 less cluster?

Do you think is correct to use the amounts presented in their original values? What is feature scaling? Which are the differences between StandardScaler and MinMaxScaler?

Adjust the model to use one of the feature scaling techniques.

- **Exercise 6**
What is the elbow method? How can you use it?

Create a cluster algorithm using K-Means for the "Clustering" dataset provided. Also, determine the best number of clusters for the algorithm with the "Elbow Method" and compare the results with arbitrary numbers of clusters.

- From the "Clustering" dataset import all the values using pandas
- Select X, and Y columns and generate a new dataset.
- What is a nan value? How would you filter nan values?
What is PCA?

How does PCA works?

How can you implement it?

For the previous exercises analyze how PCA can improve the results obtained.

Reference Links:

- https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-6a6e67336aa1
- https://365datascience.com/tutorials/python-tutorials/pca-k-means/
- https://matplotlib.org/stable/tutorials/introductory/quick_start.html#sphx-glr-tutorials-introductory-quick-start-py
- https://realpython.com/k-means-clustering-python/
- https://machinelearningmastery.com/tutorial-first-neural-network-python-keras/