# **Jordan University of Science and Technology**

**Project** – **Digit clustering** 

#### Instructions:

- You will be using Kaggle to build a model for tackling the problem of stroke detection. Your goal is to get the highest score.
- If you do not have Kaggle account, create yours using the link: SIGNUP
- You can access the competition through the following link: **COMPETITION LINK**
- This is a private competition; please do not disclose the link.
- You may use Kaggle Notebook instead of Colab.

### How do I start? The following steps require using python

- 1. Download the training and testing dataset files.
- Use clustering algorithms (for example, kmeans) to group the digits. You may need to change the default parameters. You may need to transform the attributes. https://scikit-learn.org/stable/modules/classes.html#module-sklearn.cluster
- 3. Create a data frame with two columns *ID* and *labels*. The *labels* is the predictions done in the previous step. Save the data frame in a csv file (for example *predictions.csv*).
- 4. Submit *predictions.csv* on kaggle and wait to see your score.
- 5. Take a look at the sample file (submission.csv) on Kaggle.

### What should I submit on elearning? Submit the following two files:

1. Submit a report showing your Kaggle score for each classification model and parameters you used. You may create a table like the following:

model	parameters	Kaggle score
kmeans	K=2	0.6
kmeans	init= Kmeans++	0.7
DBscan	##	0.9

2. Upload the code that achieved the maximum score on elearning and Github (you must create a Github account, if you do not have one, and create a repository). Submit the repository URL on elearning.

# What should do to get full mark?

- 1. Make sure your code is clean and free of bugs.
- 2. During the presentation, make sure you understand all the lines of your code.
- 3. Upload as many different submissions as you can. Keep in mind that the maximum number of submissions is 20 per day. Show the results of each submission in the report.
- 4. Get the highest score on the Leaderboard.

# What should I do to get Zero?

Copy someone's code or report. All copies will get zero.