

# Arabic Handwritten Recognition

## Anwar Aljohani

### I. Question/need:

- Human handwriting has unlimited imbalance from one person to another. Building a machine learning model that can classify these characters regardless of the handwriting difference
- With the current digital transformation a need of converting the physical storage documents to cloud storage. Scanning these documents and uploading them to the cloud storage may do the job, but many of them were handwritten. Here comes the benefits of this model which will help them to convert them from handwriting to text which by itself have other benefits like making it easier to search for important information

### II. Data:

The dataset for this project originates from kaggle kernels. Available for free and can be downloaded from [here](#).

All the datasets are CSV files representing the image pixels values and their corresponding label.

#### ❖ Description:

Arabic Letters Dataset is composed of 16,800 characters written by 60 participants, the age range is between 19 to 40 years, and 90% of participants are right-hand. Each participant wrote each character (from 'alef' to 'yeh') ten times. The images were scanned at the resolution of 300 dpi. Each block is segmented automatically using Matlab 2016a to determining the coordinates for each block. The dataset is partitioned into two sets: a training set of 13,440 characters to 480 images per class and a test set of 3,360 characters to 120 images per class. Writers of training set and test set are exclusive. Ordering of including writers to test set are randomized to make sure that writers of test set are not from a single institution to ensure variability of the test set.

### III. Tools:

- **Libraries**

- numpy
- pandas
- matplotlib
- Seaborn
- TensorFlow
- keras

- **Algorithm**

- Convolutional Neural Network (CNN)