



## NLP

### (Resume Analysis)

**By: Ebtesam Sultan - Nuha Aljohani - Hanaa Algarni**

---

#### **Question/Problem statement:**

Finding a job is considered one of the major activities for everyone. Typically, large companies do not have enough time to open each CV, so we are help them by use machine learning algorithms for the Resume Screening task. The purpose of this project is to use the Natural Language Processing to parse information from a resume PDF files, find the keywords, cluster them onto sectors based on their keywords, to helping human resources to fill jobs with the right person.

#### **Data Description:**

we extracted the dataset from Kaggle website, the dataset contained 2485 rows and 4 columns: (ID, Resume\_str, Resume\_html, Category), with data type (string, integer).

#### **Tools:**

**Programs:** Python and Jupiter Notebook.

**Libraries:** Pandas, NumPy, re, nltk, spacy, genism, sklearn.decomposition , sklearn.feature\_extraction.text , sklearn.model\_selection , pyLDAvis , pyLDAvis.sklearn, matplotlib.pyplot , warnings.

#### **MVP Goal:**

The goal of this project help HR employee to use machine learning algorithms for the Resume Screening task.