



Leading the Human Capital Revolution with Artificial  
Intelligence

Technical Assessment Activity

for the Role of:

**Data Scientist**

# 1. Problem

Mohammad, a recruiter at KABI based in KSA, has a list of jobs that need to be published. He is seeking assistance in establishing the appropriate salary range for each position based on job descriptions, taking into consideration factors such as skills, work experience, education, and more.

Mohammad is struggling to determine the optimal salary range. As a Data Scientist, you will assist him in determining the most suitable salary range for each position.

# 2. Request

You have been tasked with creating a solution to accurately predict salary ranges for positions in the Kingdom of Saudi Arabia (KSA), taking into account factors such as skills, work experience, education, and more.

Your assignment involves designing a solution that accomplishes the following:

## Part I: Research

1. Conduct a review of current and recent research in the field of position salary prediction and provide a concise summary of the prevalent findings and recent development.

## Part II: Data Acquisition and Integration

2. Develop a solution capable of extracting data from diverse sources, each with distinct data structures. Keep in mind that some data sources may not be consistently available. Among these sources, you will be working with the "[positions\\_salaries.csv](#)" file, as well as scraping data from <https://www.glassdoor.com/Salaries/index.htm>.
3. Your solution should efficiently manage and store data suitable for training a Salary Predictor AI model.
4. Outline the essential steps for data preprocessing and cleaning.



## Part III: Exploratory Data Analysis (EDA)

5. Conduct an in-depth Exploratory Data Analysis (EDA) addressing the following inquiries:
  - ☐ Determine the average salary increase rate when transitioning between any two sectors (e.g., from healthcare to real estate).
  - ☐ Generate and interpret the correlation matrix to assess the impact factors on position salaries (Skills, Gender, Experience Level, Education, Sector, etc.).
6. Incorporate any noteworthy observations or insights gleaned from your data analysis.

## Part IV: AI Model Training & Evaluation

7. Detail the process of training and evaluating the AI model for position salary prediction. Include information on feature selection, algorithm choice, handling biases, model evaluation metrics, hyperparameter tuning, validation techniques, and any challenges faced during the process.
8. Provide insights into how missing data was managed, any feature engineering performed, and discuss how the model results would be effectively communicated to non-technical stakeholders.



## Guidelines

- ❖ Please allocate sufficient time for crafting a well-considered solution. We evaluate your work based on both the design of your solution and the results achieved.
- ❖ You are expected to provide a detailed write-up that encompasses your methodology for both recommendations and analysis. Additionally, include relevant visualizations and summaries to effectively communicate your findings. Should you need to make any assumptions, please explicitly state them.
- ❖ Feel free to contact us anytime via email to clarify any inquiries you might have.
- ❖ The deadline for submitting this task is set for 7 days from the date of receiving the assessment activity. If you encounter any circumstances that impede your ability to meet this deadline, kindly inform us, and we will be happy to arrange an alternative due date.

