

**Power BI Dashboarding project proposal**

# **1. Executive Summary:**

This project **helps HR to negotiate their Salaries for new employee by building ML models as well as** developing interactive and insightful dashboards using Power BI. This interactive project enables HR to know the annual compensation based on their input like Age, Gender, Education, Job Title and Years of Experience.

# **2. Problem Statement:**

**Background:** Salary prediction models can provide valuable insights into the earning potential of various professions and help individuals and organizations make informed decisions.

**Objective:** The objective is to gain insight about the salaries of different job title. This analysis aims to classify the salary of the employee on the basis of age, gender, education level, job title and year of experience.

**Scope:**

1. Initial focus is on employees to overcome the loss of company
2. Improving students' motivation by focusing which job profiles pays highest compensation and planning accordingly to study ahead

**3. Data Sources:**

**Primary Data:** The dataset was obtained from multiple sources, including surveys, job posting sites, and other publicly available sources. A total of 6704 data points were collected. The dataset included five variables: age, experience, job role, education level and salary

**Secondary Data:** [Salary Data (kaggle.com)](https://www.kaggle.com/datasets/mohithsairamreddy/salary-data/data)

# **Methodology:**

**Data Sourcing: Data sourcing is the primary stage of find the relevant data for the model. For this model I’m using Online Salary Data from Kaggle.**

**Model Building**:

* In this step, I have explored the **EDA,** **data preprocessing** and then **Selecting algorithm**.
* EDA is an essential process where you try to explore the interdependencies in the data. We must also check for missing values and outliers before doing EDA.
* Selection of algorithm depends on the type of output column in the dataset, i.e. either categorical or continuous output. As the target value in our dataset is continuous, we have considered it as the regression problem.
* **Building** different linear regression models and comparing their accuracy scores. Also, once we finalize the model then tunning the model for highest possible accuracy.

**Data Integration:** Extracting and integrating data from various sources into Power BI. I will look for patterns and trends in the data also include relevant filters and detailed tooltips.

# **5. Expected Outcomes:**

* Interactive dashboards providing real-time insights into predicting salary for an employee.
* Person who wants to make a career switch to another job role that offers higher salary.
* The salary of the employees can be observed in a particular field according to their qualifications.

# **Tools and Technologies:**

1. Python for data extraction and transformation.
2. Power BI for dashboard development.

1. **Risk and Challenges:**
2. Incomplete data can affect the model accuracy.
3. Model complexity is important because complex model can overfit the model
4. Economic conditions and job markets change over time. Models trained on historical data may not accurately predict salaries in dynamic environments.

# **8. Conclusion:**

In this project I will be developing salary prediction system by using a linear regression. The result of the system is calculated by suitable algorithm by comparing it with another algorithms. I will try implementing different regression model to get better accuracy such as linear regression, ridge, lasso, support vector regression, decision tree regression and random forest regression. Also build on interactive dashboards that ensuring a user-friendly experience that promotes data-driven decision culture.