**Introduction:**

This project aims to analyze and predict the world ranking of universities using data from the Center for World University Rankings (CWUR). The dataset includes various features such as publications, quality of education, alumni employment, and more. We utilize a Random Forest Regressor to predict the world ranking based on these features and assess the importance of each feature.

### Data Description:

The dataset used in this project is cwurData.csv, which contains information about university rankings, publications, quality of education, alumni employment, and other related features.

### Data Preprocessing and Analysis:

### Loading the Data

### 

### Sorting and Descriptive Statistics

### 

### Grouping Data by Country and Calculating Average Publications

### 

### Median and Mode of Publications

### 

### Handling Missing Values and Duplicates

### 

### Unique Values in Columns

### 

### Data Visualization:

### Pie Chart for Year-wise Ranking in Times Data

### 

### Bar Plot for Top 20 Countries with Most Ranked Universities

### 

### Bar Plot for Top 10 Universities with Highest Alumni Employment Rates

### 

### Bar Plot for Top 10 Institutions in the USA Based on National Ranking

### 

### Histogram for Distribution of Publications

### 

### Random Forest Model for Predicting World Rank:

### Feature Selection and Train-Test Split

### 

### Training the Random Forest Model

### 

### Model Evaluation

### 

### Feature Importance Plot

### 

### Outputs:

### Pie Chart

### 

### Bar Plots

### 

### 

### 

### Histogram

### 

### Random Forest

### 

### Findings:

In this project, we applied a Random Forest Regressor to predict the world rankings of universities based on features like publications, quality of education, and alumni employment. The model achieved a Mean Squared Error (MSE) of 10120.67153977316, indicating the average squared difference between predicted and actual rankings. A lower MSE suggests better predictive accuracy. Additionally, through feature importance analysis, we identified that publications and alumni employment rates significantly influence university rankings. This project demonstrates the efficacy of Random Forests in leveraging multiple features to predict complex outcomes in educational data analysis.

### Conclusion:

The Random Forest model was trained to predict the world ranking of universities using features such as publications, quality of education, and alumni employment. The model's performance was evaluated using Mean Squared Error (MSE), which provides an understanding of the accuracy of the predictions.

The feature importance plot reveals that certain features have a greater influence on predicting the world ranking, helping us understand the factors that contribute significantly to a university's global standing.