**Project Name:** Employee Attrition Prediction

**Github Link:** https://github.com/projectsforstudents2022/Employee\_Attrition\_Prediction.git

**Why was this project created?**

The implementation of artificial intelligence has an impact on a company's decision-making processes in a number of organizational sectors. Human resources have received more attention in recent years as a result of the fact that employees' quality and abilities are a growth factor and a genuine competitive advantage for businesses. In fact, artificial intelligence is now beginning to influence business decisions regarding their employees after becoming more widely used in the sales and marketing sectors, with the intention of basing HR management decisions on the analysis of objective data rather than subjective considerations.

**What problem is it solving?**

The major goal of this project is to create a model that can aid in predicting whether or not an employee will stay with the organization. Measurement of employee satisfaction and assessment effectiveness inside the organization is key to lowering the attrition rate of staff members.

**Entire explanation of project**

* **PROPOSED APPROACH**

IBM Analytics provides the HRM dataset that was used in this study [32]. This dataset uses data from the United States and has 35 features connected to 1500 observations. Every element is connected to the personal and professional lives of the employees. One of the most crucial components of machine learning is data preparation, which is typically difficult and takes a lot of time. In reality, it has been shown that, generally speaking, this process takes 60% as much time and effort as a data science project. In order to observe the properties of all variables, we now constructed the dataset's descriptive statistics.

In this instance, a variety of predictive models, including those based on decision trees, the Bayesian approach, logistic regression, and SVM, were applied. Finding the most effective classifier for the studied problem is the aim. Because of this, each classifier needs to be trained on the feature set, and the classifier that produces the best classification results is used for prediction. To get more dependable and accurate indicators of faults, it's crucial to conduct independent evaluation tests and employ many observations when assessing a model's performance. In order to improve our error estimation, we implemented the two strategies listed below.

Algorithm for creating next word prediction model :

**Step 1:** Import Libraries & Load Dataset

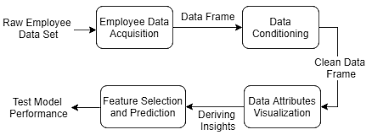
**Step 2:** Data Preprocessing

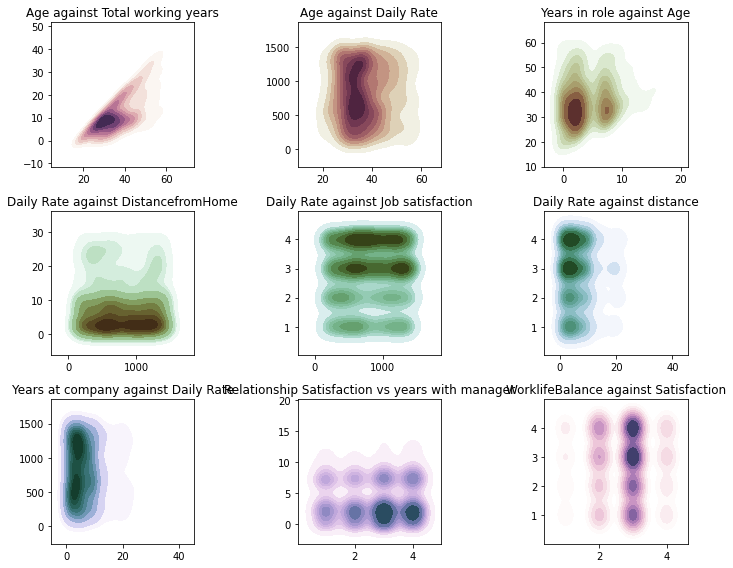
**Step 3:** Data Analysis

**Step 4:** Build Random Forest Classifier

**Step 5:** Train Model

**Step 6:** Testing & Visualization

* **DATA FLOW DIAGRAM**
* **RESULT**



* **CONCLUSION**

We have successfully learned how to use random forest to analyze staff attrition. A corporation can quickly identify the areas that need to be attended to in order to improve employee comfort at work and sustain the company's human resource power with just a few codes and a suitable data collection. We got an accuracy of 85.3%.