SMART INTERNZ

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

KANNAN.T
21BLC1521

CERTIFICATE

Google Developers

N·E·A·T
प्रौद्योगिकी के लिए राष्ट्रीय शैक्षणिक सहयोग
National Educational Alliance for Technology

Smart Internz

Externship Certificate

This is to certify that **Kannan T** has successfully completed the externship program on **Artificial Intelligence and Machine Learning** **Powered by Google Developers** from **21 August 2023** to **21 November 2023** and fulfilled the project work requirements.

Certificate ID: Ext-AIML-2023-73148

December 16, 2023

Issued Date

SkillBridge Educational Services Pvt. Ltd. Hyderabad

Jayaprakash. Ch

Program Manager, SmartInternz

ABOUT SMART BRIDGE EXTERNSHIP

PROGRAM APPROVED BY AICTE
COURSES IN PARTNERSHIP WITH GOOGLE & IBM

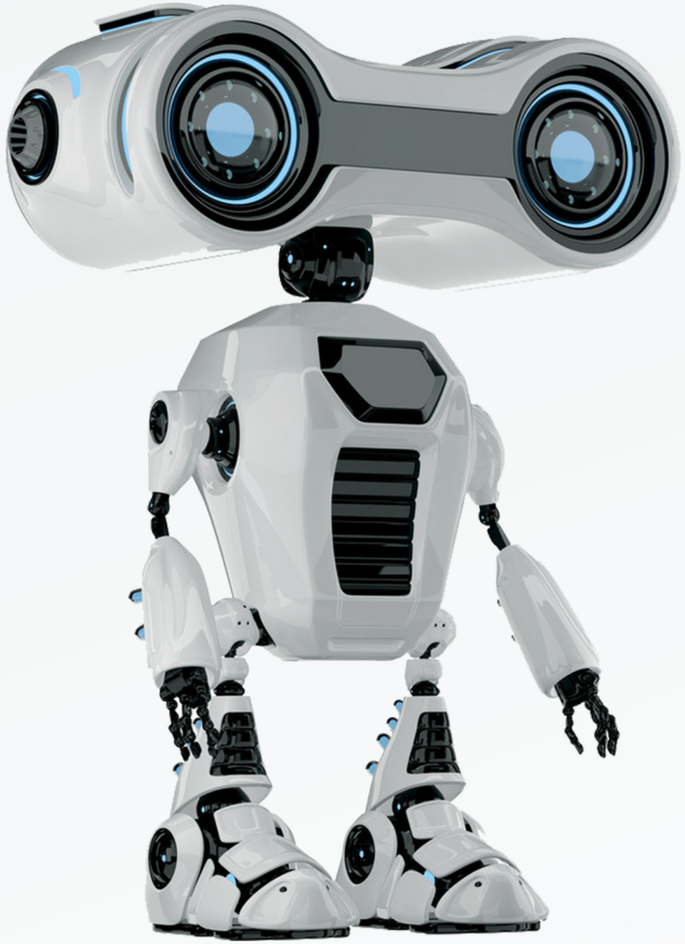
SMART BRIDGE

Founded in 2015, SmartBridge is an edtech company dedicated to building India's leading talent factory. Our vision is to cultivate a pool of job-ready talent that meets the evolving needs of the transforming industry, ensuring a readily available workforce for businesses. At SmartBridge, our cutting-edge ed-tech platform, "SmartInternz, " serves as a catalyst for fostering collaboration between academia and industry. By providing project-based, collaborative learning solutions intricately woven into the curriculum, it empowers students to cultivate the essential technical and professional skills required to become job-ready candidates. The platform's immersive learning journey equips students with the necessary expertise to excel in their chosen careers. Since the launch of SmartInternz in 2020, our talent development programs have successfully upskilled over 300,000 students and 30,000 faculty members in emerging technologies. Renowned companies such as IBM, Google, Salesforce, VMware, and others have placed their trust in our platform, providing an impressive 100,000 virtual internships to Indian students pursuing their graduation. SmartBridge is dedicated to a momentous talent mission: to provide "1 Million Virtual Internships" across a wide range of in-demand technologies. Our goal is twofold: to assist companies in finding job-ready talent and to play a pivotal role in building a thriving gig economy in India.

CONTENT

MACHINE LEARNING

Within the discipline of artificial intelligence (AI), machine learning focuses on creating statistical models and algorithms that allow computers to carry out tasks without explicit instructions. Based on data, these models and algorithms are intended to learn from and generate predictions or judgements.



TYPES OF MACHINE LEARNING

Supervised Learning

In supervised learning, the algorithm learns from labeled data, where each example in the training set is associated with a target label or outcome.

Unsupervised Learning

In unsupervised learning, the algorithm learns from unlabeled data, where the algorithm tries to find hidden patterns or structures in the data.

Reinforcement Learning

In reinforcement learning, an agent learns to make decisions by interacting with an environment. The agent receives feedback in the form of rewards or penalties based on its actions and learns to optimize its behavior over time.

Ensemble Learning

Ensemble learning combines multiple base models to improve predictive performance. Common ensemble techniques include bagging (e.g., random forests), boosting (e.g., AdaBoost, gradient boosting machines), and stacking.

DIFFERENT ML ALGORITHMS

LINEAR REGRESSION

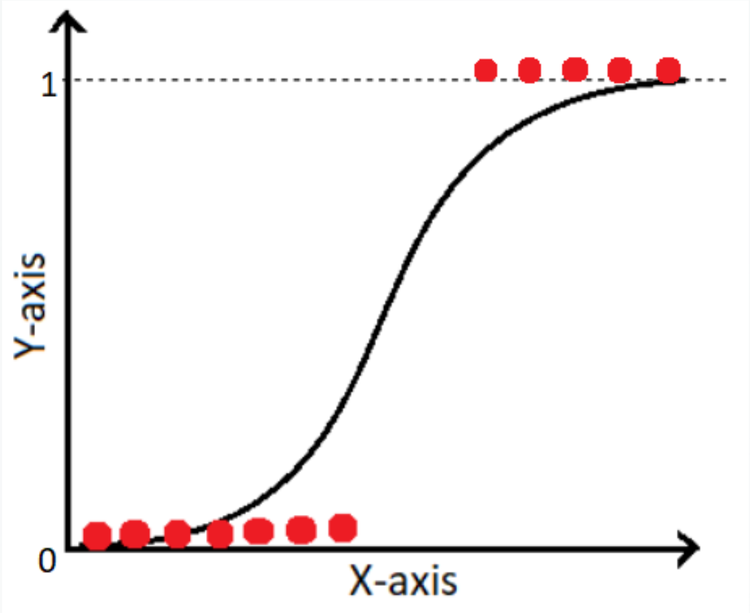
Types of Linear regression

- Simple Linear Regression
- Multiple Linear Regression

Linear regression is one of the simplest and most widely used statistical techniques for modeling the relationship between a dependent variable (target) and one or more independent variables (predictors). It assumes that the relationship between the variables is linear, meaning that a change in one variable is associated with a proportional change in the other variable(s).

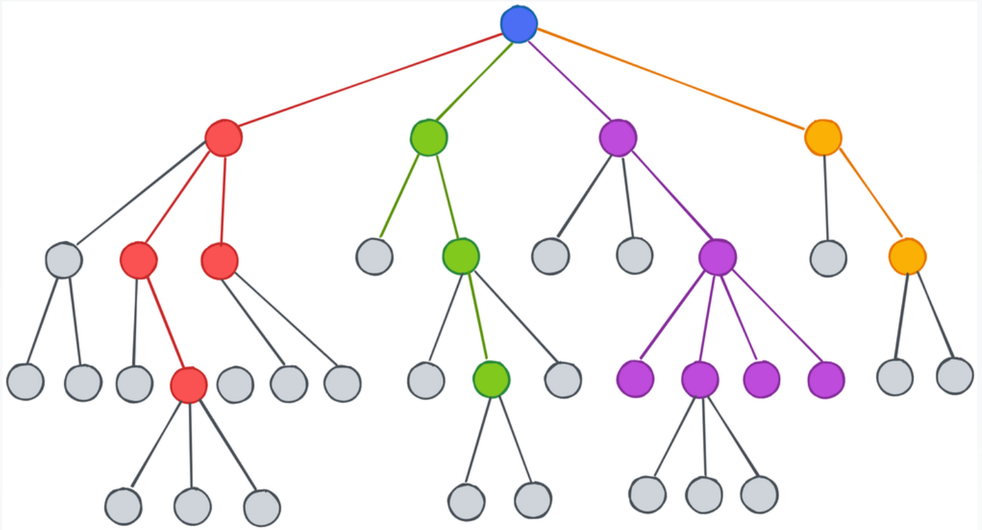
LOGISTIC REGRESSION

Logistic regression is a type of regression analysis used for predicting the probability of a categorical dependent variable. It is commonly used for binary classification problems, where the outcome variable has only two possible outcomes (e.g., yes/no, true/false, 0/1).



DECISION TREE

A decision tree is a popular supervised learning algorithm used for both classification and regression tasks. It's a tree-like structure where each internal node represents a feature, each branch represents a decision based on that feature, and each leaf node represents the outcome or predicted value. Decision trees are intuitive and easy to understand, making them widely used in various fields.



RANDOM FOREST

Random Forest is an ensemble learning method that combines multiple decision trees to create a more robust and accurate model. It is widely used for both classification and regression tasks due to its high performance, scalability, and versatility.

Ensemble Learning:

- Random Forest belongs to the ensemble learning family, which combines multiple base models to improve predictive performance.
- It builds multiple decision trees during training and aggregates their predictions to make the final prediction.

Bagging (Bootstrap Aggregating):

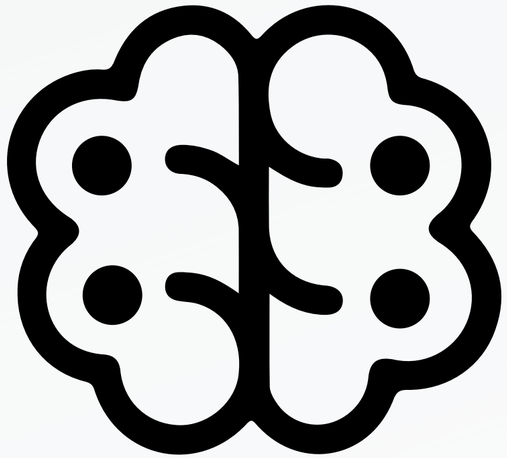
- Random Forest employs a technique called bagging, which involves training each decision tree on a random subset of the training data.
- During the training process, each tree is built using a bootstrap sample of the training data, where samples are drawn randomly with replacement.

Decision Tree Construction:

- Each decision tree in the Random Forest is constructed using a subset of the training data and a subset of the features.
- The trees are typically grown to their maximum depth without pruning, which helps capture complex relationships in the data.

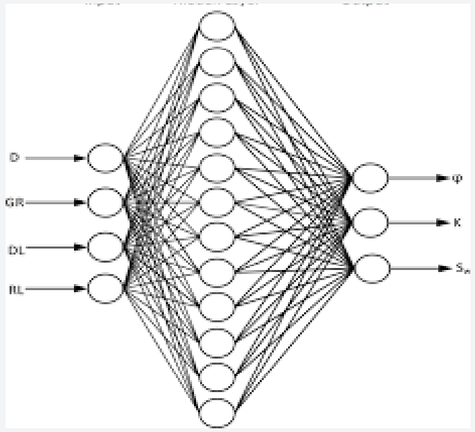
DEEP LEARNING

Deep learning is a subset of machine learning that deals with artificial neural networks (ANNs) containing multiple layers. These neural networks are composed of interconnected nodes, or artificial neurons, that process input data and gradually extract higher-level features as the data moves through the layers.



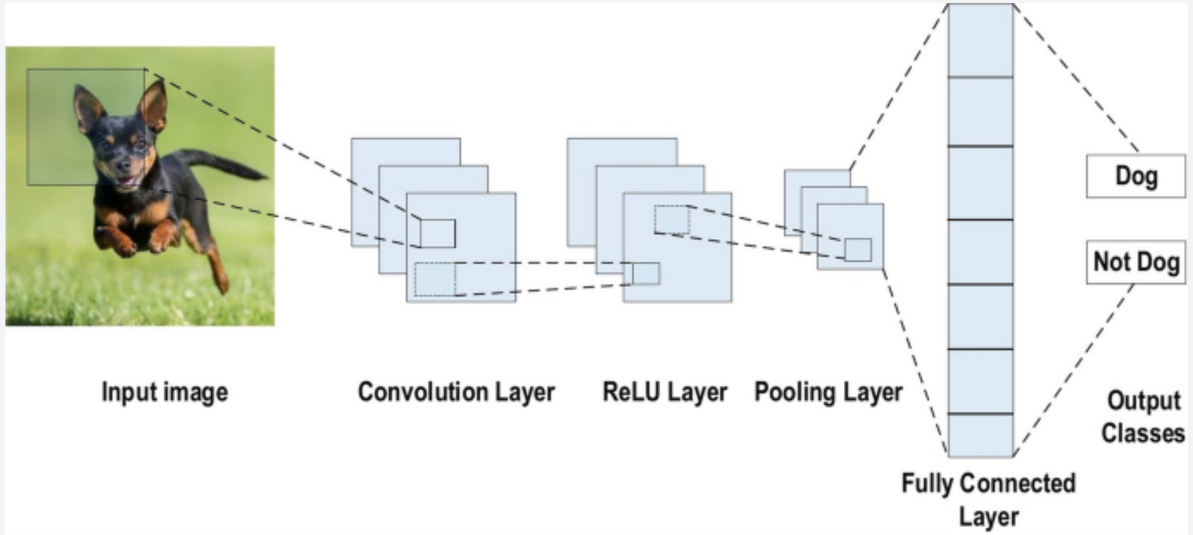
ARTIFICIAL NEURAL NETWORKS (ANNS)

- ANNs are computational models inspired by the structure and function of the human brain. They consist of interconnected layers of artificial neurons, with each neuron performing a simple computation on its input and passing the result to the next layer.
- The input layer receives the raw data, such as images, text, or numerical features. The output layer produces the final prediction or output.
- Intermediate layers, called hidden layers, extract and transform the features of the input data through a series of weighted connections and activation functions.



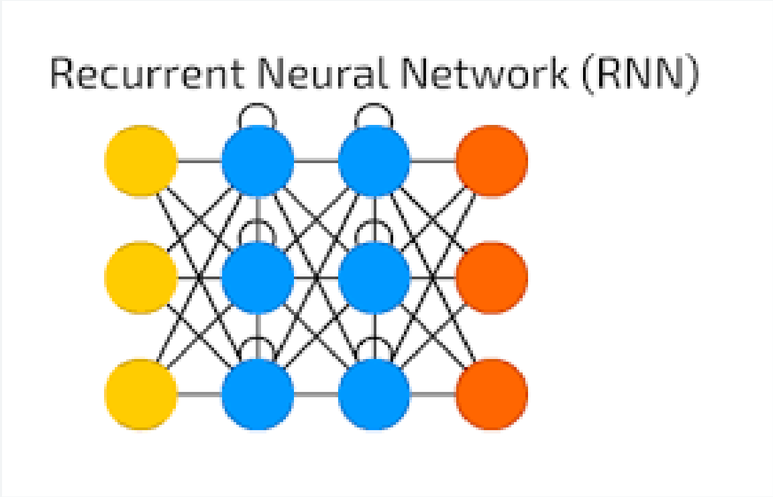
CONVOLUTIONAL NEURAL NETWORKS (CNNS)

- CNNs are a type of deep neural network designed for processing structured grid-like data, such as images.
- They consist of convolutional layers, pooling layers, and fully connected layers. Convolutional layers apply convolution operations to the input data, extracting spatial features and patterns.
- CNNs have been highly successful in computer vision tasks such as image classification, object detection, and image segmentation.



RECURRENT NEURAL NETWORKS (RNNS)

- RNNs are a type of neural network designed for processing sequential data, such as time series, text, and speech.
- Unlike feedforward neural networks, RNNs have connections that form directed cycles, allowing them to capture temporal dependencies and context information.
- Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) are popular variants of RNNs that address the vanishing gradient problem and improve the model's ability to capture long-range dependencies.



PROJECT

EMPLOYEE ATTRITION RATE AND ANALYSIS

INTRODUCTION

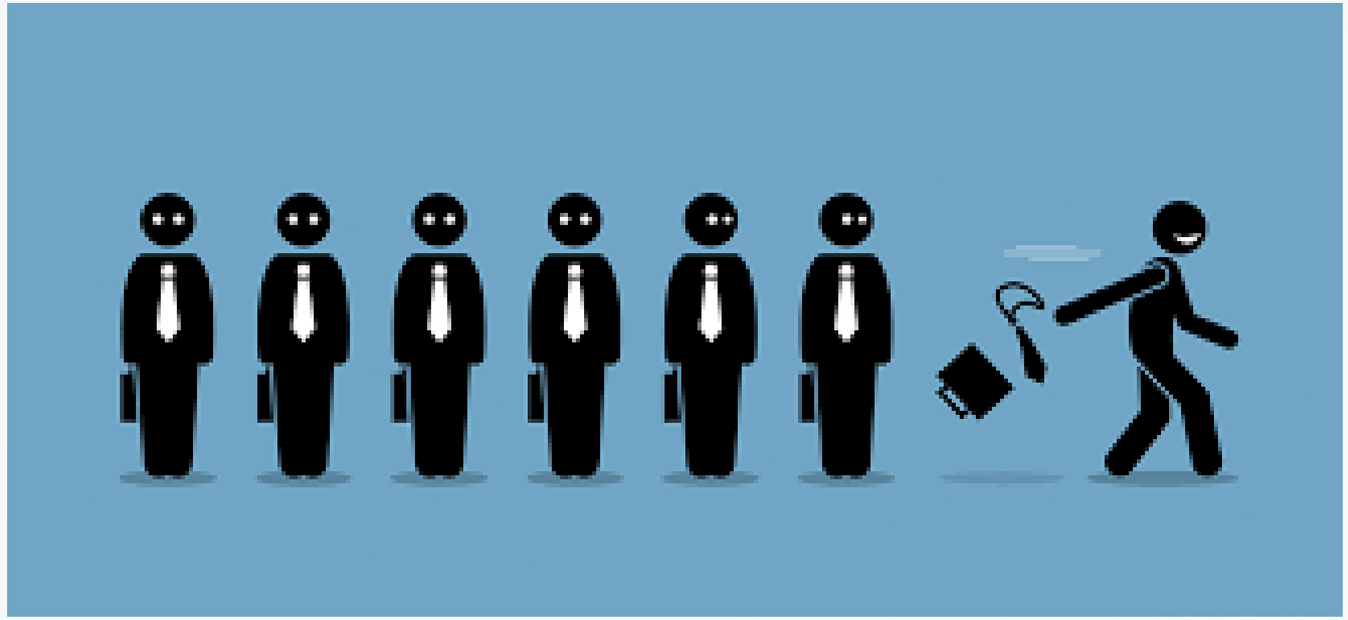
The aim of the project is to analyze and predict employee attrition within a company using machine learning techniques. The goal is to identify factors that contribute to employee turnover and develop predictive models to forecast attrition

Purpose

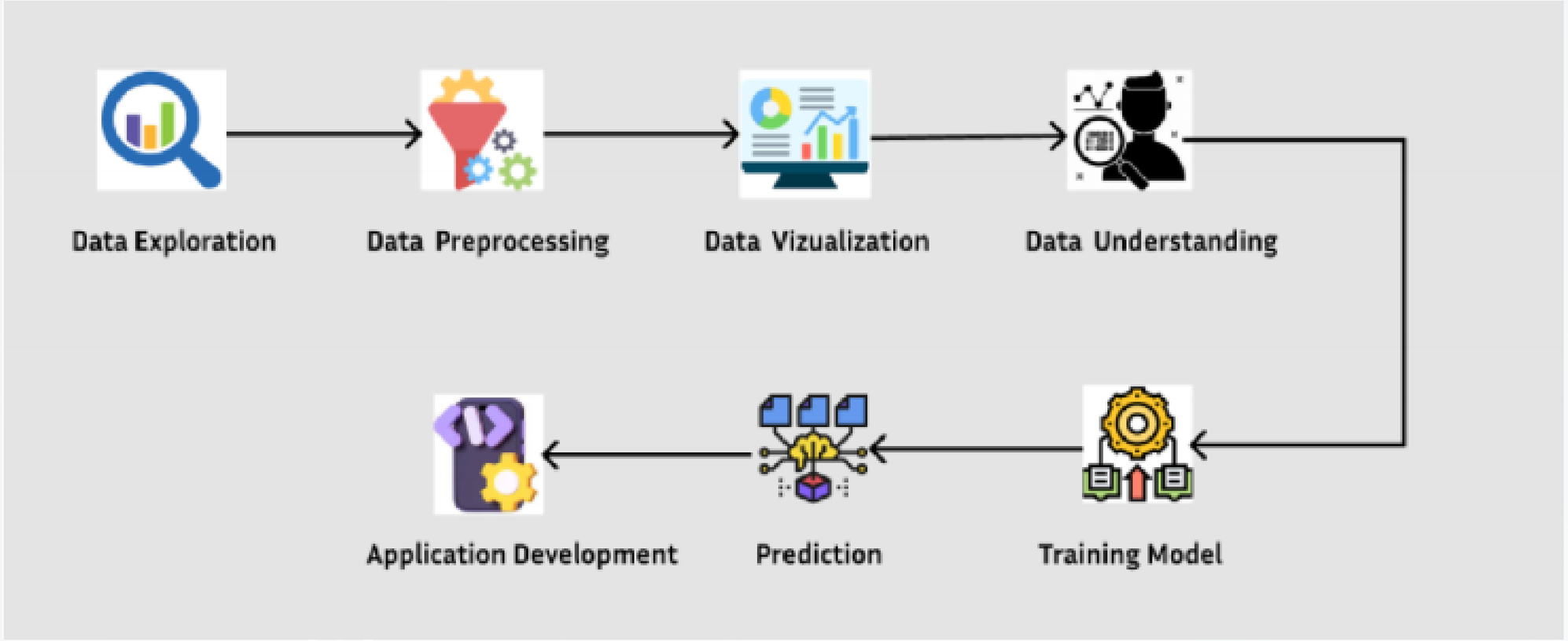
The purpose of analyzing employee attrition rate using machine learning techniques is to gain insights into the factors that contribute to employee turnover within an organization. By leveraging machine learning algorithms, businesses can identify patterns and relationships in large datasets to predict and understand why employees are leaving. This analysis can help organizations take proactive measures to retain valuable employees, improve employee satisfaction, and reduce turnover rates.

Problem Statement Definition

Many well-established companies operating in a competitive industry, have been facing persistent challenges related to employee turnover. High employee attrition rates have resulted in increased recruitment and training costs, disruption of workflow, and a negative impact on overall productivity



Solution Architecture



MODEL USED

LOGISTIC REGRESSION

In the followin project we have used Logistic Regression for the prediction of Employee Attrition.

Efficiency

Logistic regression is computationally efficient and can be trained quickly, especially on large datasets compared to more complex models like neural networks. It is well-suited for scenarios where computational resources are limited.

Low Complexity

Logistic regression is a simple and linear model, making it less prone to overfitting, especially when the number of features is small relative to the number of observations. Regularization techniques such as L1 or L2 regularization can be applied to further reduce overfitting.

Interpretability

Logistic regression produces interpretable results, as the coefficients of the model represent the relationship between the independent variables and the log-odds of the outcome. This makes it easy to understand the impact of each feature on the probability of the outcome.

OUTCOME OF PREDICTION

```
print(confusion_matrix(Y_test,prediction))
```

```
[[361  10]
 [ 51  19]]
```

```
print(classification_report(Y_test,prediction))
```

```
              precision    recall  f1-score   support

           0       0.88      0.97      0.92       371
           1       0.66      0.27      0.38        70

    accuracy                           0.86       441
   macro avg       0.77      0.62      0.65       441
weighted avg       0.84      0.86      0.84       441
```

CONCLUSION

Overall, machine learning offers tremendous potential in addressing the problem of employee attrition, but it should be used as a tool to augment human decision-making rather than replacing it entirely.

FUTURE SCOPE

Machine learning algorithms can be used to build predictive models that analyze various employee attributes, such as job satisfaction, performance metrics, and engagement levels, to forecast the likelihood of attrition. These models can help organizations identify high-risk employees or departments and take proactive measures to retain them.

THANK YOU