## Artificial Intelligence and

## Machine Learning

Project Report

Semester-IV (Batch-2022)

Loan Eligibility Prediction

A red and white sign

Description automatically generated with low confidence

**Supervised By: Submitted By:**

Rajeev Bhardwaj Name:Sneha Chaudhary

Roll no:2210990857

Group:G-13

**Department of Computer Science and Engineering**

## Chitkara University Institute of Engineering & Technology,

## Chitkara University, Punjab

# Abstract

Title: Predicting Loan Eligibility Using Machine Learning: A Comprehensive Study

Abstract:

In recent years, the use of machine learning (ML) algorithms in financial institutions has gained considerable attention due to their potential to streamline processes, mitigate risks, and enhance decision-making accuracy. One critical application of ML in the financial sector is predicting loan eligibility, which involves assessing the likelihood of an individual being approved for a loan based on various factors. This study presents a comprehensive analysis of loan eligibility prediction using ML techniques.

The primary objective of this research is to develop robust ML models capable of accurately predicting loan eligibility while considering factors such as income, credit score, employment status, debt-to-income ratio, and other relevant features. To achieve this goal, a diverse dataset comprising historical loan application data is collected and preprocessed. Data preprocessing involves handling missing values, feature scaling, and encoding categorical variables to prepare the dataset for training ML models.

Several ML algorithms, including logistic regression, decision trees, random forests, support vector machines, and gradient boosting machines, are implemented and evaluated for their performance in predicting loan eligibility. Model evaluation metrics such as accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC) are utilized to compare the effectiveness of different algorithms.

Furthermore, feature importance analysis is conducted to identify the most influential factors affecting loan eligibility decisions. This analysis provides valuable insights into the key determinants considered by the ML models and aids in improving model interpretability.

The study also explores the impact of data imbalance on model performance and investigates techniques such as oversampling, undersampling, and synthetic data generation to address this challenge effectively.

Additionally, the deployment of the developed ML models into real-world lending scenarios is discussed, highlighting the potential benefits for financial institutions in automating loan approval processes, reducing manual effort, and minimizing the risk of default.

Overall, this research contributes to advancing the understanding of ML-driven loan eligibility prediction and provides practical insights for implementing effective and reliable loan approval systems in the financial industry.