2026 Winter High Schools Students Online Research

Project Topic

AI for Loan Approval: Exploratory Data Analysis, Model Building, Training, and Implementation

Loan Approval Background

Loan approval is a critical function in the financial services industry, directly affecting both applicants and lending institutions. Traditionally, approval decisions depend on manual review processes that can be inconsistent, biased, and time-consuming. With the growth of Artificial Intelligence (AI) and data-driven methods, automated loan approval systems are increasingly becoming a reality. These systems can allow faster decision-making, improved accuracy, reduced operational costs, and fairer assessments.

Applicants are typically evaluated using details such as income, employment history, credit history, and loan intent. Using these features, AI can help financial institutions improve efficiency and provide a consistent, transparent decision-making process.

Research Project Description

This research program is designed to introduce high school students to the fundamentals of data science, visualization, AI model building, training, and evaluation for predictive analytics. Over four weeks, students will develop a solid foundation in data analytics and machine learning concepts while applying them to a practical project: predicting loan approval outcomes. Students will explore end-to-end workflows, from problem formulation through data gathering, preprocessing, exploratory data analysis (EDA), and building predictive models.

Students will also utilize Python-based tools (Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn) to visualize and model the data. The program culminates in a comprehensive research project where students will present a predictive model supported by their analysis of loan approval data. The overall goal of the project is to create an interpretable and fair AI-based system for loan approval prediction.

Prerequisites

- **Grade Level:** 10th grade and above.
- **Interest:** A keen interest in data science, finance, and AI applications.
- **Skills:** Basic understanding of data and statistics such as mean, median, and mode.
- **Technology:** Access to a computer with internet connection, and basic Python coding knowledge.

Project Outcomes

By the end of the program, students will be able to:

1. **Review Data Analytics Fundamentals**: Understand different types of data analysis methods.

2. **Use AI Tools**: Explore AI research tools, including ChatGPT, to support predictive modeling.

3. **Conduct Descriptive Analytics Research**: Gather and process loan-related datasets to perform descriptive and diagnostic analysis.

4. **Explore and Visualize Data**: Use Python libraries to analyze and visualize features like income, employment length, and credit score.

5. **Develop Predictive Models**: Build, train, and validate regression and classification models for loan approval.

6. **Enhance Research Skills**: Gain experience in data collection, analysis, and interpretation.

7. **Improve Communication Skills**: Present findings through written reports and oral presentations.

8. **Gain Practical Experience**: Work on a hands-on capstone project applying data science and AI to a real-world financial problem.

Tentative Schedule

Week 1: Introduction to Loan Approval Research

- Overview of financial lending and loan approval processes.
- Assessing applicant features: income, credit history, employment length, loan intent.
- Literature review and case studies of AI in finance.
- Gathering datasets and performing initial analysis.
- Hands-on practice with Excel and Python.

Week 2: Data Processing

- Understanding data preprocessing and cleaning techniques.
- Introduction to AI tools and ChatGPT.
- Reading and preparing data using Python (Pandas, NumPy, Scikit-learn).
- Visualizations with Matplotlib, Seaborn, Tableau, or CODAP.
- Creating scatter plots, histograms, and correlation heatmaps.

Week 3: Model Development

- Introduction to predictive modeling for classification.
- Building Logistic Regression, Decision Tree, and Random Forest models.
- Training, testing, and evaluating models.
- Comparing model accuracy, precision, recall, and AUC-ROC.

Week 4: Project Development and Presentation

- Finalizing the loan approval predictive model and research report.
- Preparing presentations of findings.
- Peer review, feedback, and refinement of deliverables.

Deliverables

- **Research Report:** A detailed report on the chosen project, including background, data sources, methodology, analysis, and findings.

- **Loan Approval Proposal:** A comprehensive study plan for building a predictive model for loan approval.

- **Presentation:** A PowerPoint presentation summarizing the research findings and project story.

References

- https://scikit-learn.org/
- https://pandas.pydata.org/
- https://numpy.org/
- https://matplotlib.org/
- https://seaborn.pydata.org/
- https://www.kaggle.com/datasets
- https://about.bankofamerica.com/en/making-an-impact/artificial-intelligence-in-finance
- https://www.ibm.com/topics/ai-in-finance
- chatgpt.com

Thursday, Friday 8-9 PM