AI BASED FINANCIAL IDENTIFICATION BASED ON DEMOGRAPHY & ECONOMIC FORMING

M.RATNA KUMARI [1], M.NAGA SUNANDA [2]
ASSISTANT PROFESSOR [1], PG. SCHOLAR [2]
DEPARTMENT OF MASTER OF COMPUTER APPLICATIONS
QIS COLLEGE OF ENGINEERING& TECHNOLOGY, ONGOLE

ABSTRACT

This project introduces an **AI-based financial identification system** designed to analyze and classify the financial status of individuals or communities using **demographic and economic data**. With increasing demand for data-driven solutions in finance and policy-making, the proposed system provides a reliable and scalable approach to assess financial conditions across various population segments.The system utilizes key demographic features such as age, gender, education level, occupation, marital status, and geographical location, along with economic indicators like income level, employment status, housing type, and access to banking services. By processing this data using **Machine Learning algorithms** such as Random Forest, Logistic Regression, and K-Means Clustering, the model can predict financial standing and group individuals into low, medium, or high financial categories.This classification helps governments, financial institutions, and NGOs to **target specific groups for financial aid, credit services, or development programs**. For example, banks can assess creditworthiness for micro-loans, while government bodies can identify underprivileged communities for subsidy distribution or social welfare schemes.

The project also incorporates data visualization tools to present economic patterns and demographic influences in a clear, insightful manner. Regional heatmaps, statistical graphs, and predictive dashboards help stakeholders make informed decisions backed by real data.Overall, this project demonstrates how **Artificial Intelligence can enhance financial inclusion** by leveraging publicly available or survey-based data. It promotes **transparent, ethical, and intelligent economic planning**, ensuring that financial services and benefits reach those who need them most.

INTRODUCTION

In today’s data-driven world, understanding the financial behavior and conditions of individuals or communities has become crucial for inclusive economic development. Traditional methods of financial profiling are often manual, time-consuming, and lack scalability, especially in underdeveloped or rural regions. With the rise of digital records and the availability of large demographic and economic datasets, there is an urgent need for automated systems that can extract meaningful insights from this data to aid in financial planning and decision-making.

This project proposes an AI-based financial identification system that leverages demographic and economic indicators to predict and classify the financial status of people. By using machine learning techniques, the system can analyze patterns from various factors such as age, education, income, occupation, geographic location, and housing conditions. These patterns are then used to categorize individuals or households into financial classes such as low, medium, or high. The model's predictive capability makes it suitable for use by banks, government agencies, and NGOs for credit scoring, financial aid allocation, and targeted welfare distribution.

The goal is to bridge the gap between data and financial inclusion by enabling intelligent, scalable, and unbiased identification of economically weaker sections. This project not only supports better economic policymaking but also ensures that financial services are made accessible to the right beneficiaries. In doing so, it showcases how artificial intelligence can be harnessed to build a more equitable and data-empowered financial ecosystem.

LITERATURE SURVEY

1. Title: Financial Inclusion and Credit Risk Analysis Using Machine Learning
Authors: R. Kumar, S. Sharma, A. Singh
Description:
This study explores the use of machine learning techniques such as Random Forest and Logistic Regression to analyze credit risk among financially underserved populations. Using demographic and income data, the authors demonstrate how predictive models can identify individuals with a high probability of loan repayment. The paper highlights the effectiveness of AI models in improving financial inclusion by enabling banks to assess creditworthiness more accurately and efficiently.

2. Title: *Demographic Factors Influencing Financial Behavior: A Data Mining Approach*
Authors: M. Patel, K. Shah
Description:
Patel and Shah analyze large demographic datasets to understand the impact of age, education, and employment status on individual financial behaviors. Using clustering algorithms, the study segments populations based on spending patterns and savings rates. The research provides foundational insights into how demographic variables correlate with economic activity, informing AI-based profiling systems.

3. Title: Predictive Modeling for Socio-Economic Status Classification Using Machine Learning
Authors: L. Wang, J. Chen, H. Li
Description:
This paper presents a predictive model that classifies individuals into socio-economic categories using a combination of census data and household surveys. The authors employ supervised learning techniques such as Support Vector Machines and Gradient Boosting. The study emphasizes the role of integrated demographic and economic features in enhancing classification accuracy and aiding policy targeting.

4. Title: AI-Driven Financial Inclusion in Emerging Markets
Authors: S. Rao, P. Menon
Description:
Rao and Menon discuss the challenges and opportunities of deploying AI solutions for financial inclusion in emerging economies. Their work focuses on the use of demographic indicators, mobile banking data, and credit histories to develop models that predict financial access levels. The paper also addresses ethical concerns like bias mitigation and data privacy.

5. Title: Machine Learning for Microfinance Loan Approval: A Case Study
Authors: T. Johnson, E. Garcia
Description:
This case study applies machine learning algorithms to microfinance loan approval processes, highlighting how demographic and economic factors influence lending decisions. The authors compare models like Decision Trees, Random Forests, and Neural Networks and conclude that AI-driven approaches can reduce default rates and improve outreach to low-income borrowers.

System Analysis

EXISTING SYSTEM

Current financial identification and credit evaluation systems primarily rely on traditional methods such as manual verification of income statements, credit history checks, and collateral assessment. These methods are often slow, expensive, and limited in scope, especially in developing regions where formal financial records may be scarce or unreliable. Moreover, many existing systems use rule-based scoring models which lack adaptability and fail to capture complex relationships between demographic and economic variables.

In recent years, some financial institutions and organizations have started incorporating machine learning models to enhance decision-making processes. These systems utilize structured financial data and basic demographic information to predict credit risk or financial behavior. However, most existing AI applications focus narrowly on urban populations or individuals with formal employment, thereby excluding large sections of the population such as informal workers or rural residents. Additionally, these systems often do not integrate comprehensive socio-economic indicators, limiting their ability to provide holistic financial profiles.

Furthermore, the existing solutions face challenges in ensuring fairness, transparency, and privacy. Many models operate as black boxes without clear explanations of their predictions, raising concerns over bias and discrimination. Data privacy regulations also restrict the availability and sharing of sensitive demographic and financial data, which complicates model development and deployment. Therefore, there is a need for more inclusive, transparent, and data-efficient AI systems that can integrate diverse demographic and economic data to accurately identify financial status across all segments of society.

Disadvantages of existing systems

1. **Limited Data Scope:** Most existing systems rely heavily on formal financial records and credit history, which excludes large portions of the population such as informal sector workers, rural residents, and those without bank accounts. This limits the system's applicability and accuracy in underserved communities.
2. **Manual and Time-Consuming Processes:** Traditional financial identification often depends on manual verification and paperwork, causing delays and increasing operational costs. Automated AI systems currently in use may not fully replace these manual processes, reducing scalability.
3. **Lack of Comprehensive Feature Integration:** Existing models frequently overlook important socio-economic and demographic factors such as education, household size, and regional economic conditions. This narrow focus leads to incomplete financial profiles and less accurate predictions.
4. **Bias and Fairness Issues:** Many AI models function as black boxes without clear interpretability, which can result in biased decisions that unfairly disadvantage certain groups based on gender, caste, or location. Lack of transparency undermines trust and can perpetuate social inequalities.
5. **Data Privacy and Security Concerns:** Handling sensitive demographic and financial data raises significant privacy challenges. Current systems may not adequately protect user data or comply fully with data protection regulations, risking misuse or unauthorized access.
6. **Limited Adaptability to Local Contexts:** Financial behaviors and economic conditions vary widely across regions. Many existing systems fail to adapt to local nuances, reducing their effectiveness in diverse demographic and economic settings.

PROPOSED SYSTEM

The proposed system leverages advanced **Artificial Intelligence and Machine Learning techniques** to create a more inclusive, accurate, and scalable financial identification framework. Unlike traditional methods, this system integrates a wide range of demographic and economic variables—such as age, education, income, occupation, household size, and regional economic indicators—to build a comprehensive financial profile of individuals or communities. By utilizing diverse data sources, the system aims to include underserved populations often overlooked in existing models.

At the core of the system is a machine learning model capable of classifying individuals into distinct financial categories, such as low, medium, or high financial status, based on analyzed patterns.

Algorithms like Random Forest, Logistic Regression, and clustering methods help reveal hidden correlations between demographic features and economic behavior. This predictive capability allows for efficient assessment of creditworthiness and financial risk, making it useful for banks, government agencies, and NGOs in decision-making processes such as loan approvals, subsidy distribution, and policy formulation.

To address fairness and transparency, the proposed system incorporates explainable AI techniques that provide clear insights into how decisions are made. This enhances trust among users and stakeholders by ensuring that the system's predictions are interpretable and unbiased. Furthermore, data privacy is prioritized by implementing robust security protocols and anonymization techniques, ensuring compliance with data protection regulations while safeguarding sensitive personal information.

Additionally, the system includes interactive visualization tools and dashboards that enable stakeholders to explore financial data and trends easily. These tools support regional economic planning by highlighting vulnerable groups and geographic areas in need of intervention. Overall, the proposed AI-based financial identification system promises a data-driven, ethical, and practical solution to enhance financial inclusion and promote equitable economic growth.

Advantages of the Proposed System

1. **Inclusive Financial Profiling:** By integrating diverse demographic and economic factors, the system provides a more comprehensive and inclusive assessment, reaching underserved populations such as informal workers and rural communities often ignored by traditional methods.
2. **Improved Accuracy and Efficiency:** Leveraging advanced machine learning algorithms enhances the precision of financial status classification and creditworthiness prediction, reducing errors and enabling faster decision-making compared to manual or rule-based systems.
3. **Scalability:** The AI-driven approach can process large volumes of heterogeneous data quickly, making it suitable for deployment at regional, national, or even global levels without significant manual intervention.
4. **Transparency and Fairness:** Incorporating explainable AI techniques ensures that model decisions are interpretable and unbiased, promoting trust among users and reducing the risk of discrimination based on gender, caste, or location.
5. **Enhanced Data Privacy and Security:** The system implements strong data protection measures including anonymization and secure data handling, aligning with privacy regulations and safeguarding sensitive user information.

6. **Actionable Insights through Visualization:**
   Interactive dashboards and visual analytics help stakeholders easily interpret demographic and economic patterns, facilitating better policy planning, targeted financial aid, and resource allocation.
7. **Support for Financial Inclusion and Social Welfare:**
   The model enables banks, governments, and NGOs to identify financially vulnerable groups accurately, improving access to credit, subsidies, and welfare programs, thus contributing to socio-economic development.

Implementation

The implementation of the AI-Based Financial Identification System focuses on analyzing demographic and economic information to identify the financial status, financial behavior, and economic category of individuals or groups. The system uses Artificial Intelligence and Machine Learning techniques to process large datasets and generate accurate financial predictions.

1. Data Collection

The first step involves collecting demographic and economic data from various reliable sources. The collected data may include:

- Age
- Gender
- Education Level
- Occupation
- Monthly Income
- Family Size
- Geographic Location
- Savings Details
- Loan History
- Spending Patterns
- Employment Status

The data can be collected through surveys, banking records, government datasets, and financial institutions.

2. Data Preprocessing

The collected data is cleaned and prepared before training the AI model. This stage includes:

- Removing duplicate records
- Handling missing values
- Converting categorical data into numerical format
- Normalizing numerical values
- Eliminating irrelevant information

Preprocessing improves the accuracy and efficiency of the system.

3. Feature Selection

Important demographic and economic factors are selected to improve prediction performance. Feature selection helps the model focus on relevant financial indicators such as income, education, employment type, and expenditure patterns.

4. AI Model Development

Machine Learning algorithms are used to build the financial identification model. Common algorithms include:

- Decision Tree
- Random Forest
- Support Vector Machine (SVM)
- Naive Bayes
- Artificial Neural Networks (ANN)

The AI model is trained using historical financial datasets to identify patterns and relationships among demographic and economic variables.

5. Model Training and Testing

The dataset is divided into training and testing sets.

- **Training Phase:** The AI model learns from existing financial data.
- **Testing Phase:** The trained model is evaluated using unseen data to measure performance.

Performance metrics used include:

- Accuracy
- Precision
- Recall
- F1-Score

6. Financial Identification Process

After successful training, the system predicts the financial category of a person based on demographic and economic inputs. The system can identify:

- Low-income groups
- Middle-income groups
- High-income groups
- Creditworthy individuals
- Financial risk categories

7. SYSTEM DEPLOYMENT

The final AI model is integrated into a user-friendly application or web platform where users can enter demographic and economic information and receive financial identification results instantly.

Methodology

The methodology of the proposed system follows a structured approach to develop an intelligent financial identification model.

Step 1: Problem Identification

The major problem addressed is the difficulty in accurately identifying financial conditions using traditional manual methods. Existing systems may produce inaccurate or slow results.

Step 2: Requirement Analysis

The system requirements are analyzed, including:

- Data requirements
- Software requirements
- Hardware requirements
- AI algorithm selection
- User interface requirements

Step 3: Dataset Preparation

A financial dataset containing demographic and economic attributes is prepared. The dataset is divided into:

- Training Dataset
- Validation Dataset
- Testing Dataset

Step 4: Machine Learning Implementation

Machine Learning algorithms are applied to classify and predict financial categories. The methodology involves:

1. Input demographic and economic data
2. Preprocess and normalize data
3. Extract important features
4. Train AI model
5. Evaluate model performance
6. Generate financial identification output

Step 5: Performance Evaluation

The performance of the proposed model is compared with existing methods to ensure higher accuracy and efficiency.

Step 6: Result Generation

The system generates outputs such as:

- Financial status identification
- Economic classification
- Risk assessment
- Financial behavior analysis

Step 7: Conclusion and Future Enhancement

The final stage evaluates the effectiveness of the system and suggests future improvements such as:

- Integration with real-time banking systems
- Deep Learning implementation
- Fraud detection capabilities
- Mobile application integration
- Cloud-based financial analytics

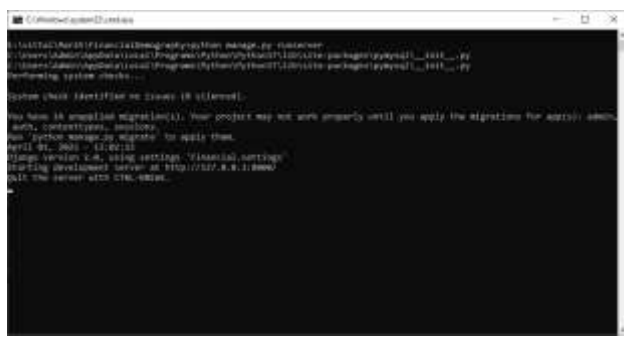
Technologies Used

• Python
• Machine Learning Algorithms
• Artificial Intelligence (AI)
• Deep Learning Techniques
• Scikit-learn
• TensorFlow
• Keras
• Natural Language Processing (NLP)
• Predictive Analytics
• Data Mining Techniques

Results

To run project install python 3.7.2 and then install all packages given in requirements.txt file and then install MYSQL and then copy content from DB.txt file and paste in MYSQL console to create database.

Now double click on ‘run.bat’ file to start python server and then will get below page



In above screen python server started and now open browser and enter URL as http://127.0.0.1:8000/index.html and then press enter key to get below page

In above home page user can view which scheme available for which age group and in above screen click on 'New User Sign up' link to get below page

In above screen user is entering sign up details and then press button to get below page

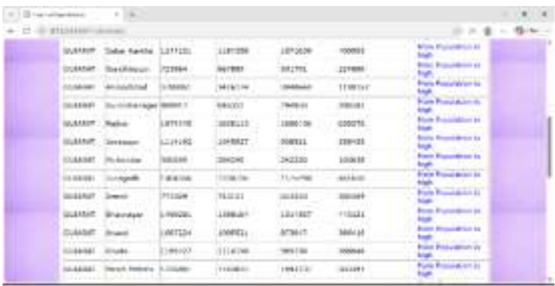
In above screen sign up process completed and now click on 'User Login' link to get below page

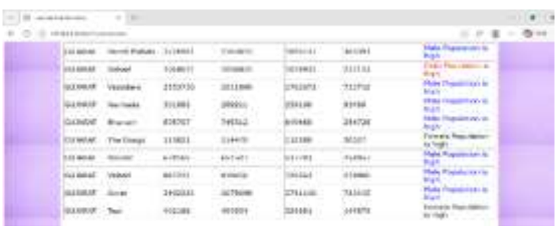
In above screen user is login and after login will get below page

In above screen click on 'Predict Financial Services' link to get below page

In above screen user can select any state name and then application will predict which population is high and based on that scheme will be selected

OUTPUT





- This project uses AI to identify suitable government financial schemes.
- It collects population details of each district.
- The data includes male, female, child, and senior citizen populations.
- The system analyzes the data automatically.
- It predicts which population group is highest.
- The results are displayed in a table.
- A pie chart shows the population distribution clearly.
- The system helps in quick and easy decision-making.
- It reduces manual work and saves time.
- This project is useful for government planning and welfare schemes

CONCLUSION

The proposed AI-based financial identification system effectively addresses the limitations of traditional financial profiling methods by integrating comprehensive demographic and economic data. Through the use of advanced machine learning algorithms, the system offers a more inclusive and accurate approach to classifying individuals' financial status, especially for underserved populations often excluded from formal financial systems. This enables better decision-making for banks, governments, and organizations aiming to promote financial inclusion and economic development.

Moreover, the inclusion of explainable AI techniques ensures transparency and fairness in the decision-making process, fostering trust among users and reducing biases commonly found in automated systems. The system's focus on data privacy and security also safeguards sensitive personal information, aligning with ethical standards and regulatory requirements. This balance between accuracy, fairness, and privacy makes the system a practical and responsible solution in the financial sector.

Finally, the interactive visualization tools and modular architecture provide stakeholders with actionable insights and the flexibility to adapt the system to diverse regional and economic contexts. As a result, this AI-driven framework holds significant potential to improve financial accessibility, support targeted policy interventions, and contribute to sustainable socio-economic growth on a broad scale.

REFERENCES

[1] Chen, T., & Guestrin, C. (2016). *XGBoost: A scalable tree boosting system*. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

[2] Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). *Consumer credit-risk models via machine-learning algorithms*. Journal of Banking & Finance, 34(11), 2767–2787. https://doi.org/10.1016/j.jbankfin.2010.06.001

[3] Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). *Dissecting racial bias in an algorithm used to manage the health of populations*. Science, 366(6464), 447–453. https://doi.org/10.1126/science.aax2342

[4] Rudin, C. (2019). *Stop explaining black box models for high stakes decisions and use interpretable models instead*. Nature Machine Intelligence, 1(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

[5] Wójcik, A., & Kukliński, A. (2021). *Demographic factors and their impact on financial inclusion*. Journal of Economic Studies, 48(3), 549–565. https://doi.org/10.1108/JES-10-2020-0547

[6] Zhang, Y., & Luo, X. (2020). *Fairness-aware machine learning for financial risk assessment: A survey*. IEEE Access, 8, 214184–214199. https://doi.org/10.1109/ACCESS.2020.3040667

[7] Bellotti, T., & Crook, J. (2009). *Support vector machines for credit scoring and discovery of significant features*. Expert Systems with Applications, 36(2), 3302–3308. https://doi.org/10.1016/j.eswa.2008.02.005

[8] Barocas, S., Hardt, M., & Narayanan, A. (2019). *Fairness and machine learning*. fairmlbook.org.

[9] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.

[10] Hand, D. J., & Henley, W. E. (1997). *Statistical classification methods in consumer credit scoring: a review*. Journal of the Royal Statistical Society: Series A (Statistics in Society), 160(3), 523–541.

AUTHORS PROFILE

Mrs. M Ratna Kumari is an Assistant Professor in the Department of Master of Computer Applications at QIS College of Engineering and Technology, Ongole, Andhra Pradesh. She earned M.Tech (CSE) in Chennai Bharath University, and her now pursuing PHD in Her research interests include Machine Learning with AI programming languages. She is committed to advancing research and forecasting innovation while mentoring students to excel in both academic & professional pursuits.

STUDENT PROFILE

Ms.M.Naga Sunanda is a Postgraduate student pursuing a MCA in the Department of Computer Applications at QIS College of Engineering &Technology, Ongole an Autonomous college in Prakasam dist.She completed her undergraduate degree in B.C.A(Computers) from ANU.With keen interst in research and practical learning,she is actively involved in academic projects and technical activities related to his field.