**UNCLEAR LOAN REPAYMENT TRENDS IN YOUTH SACCO AND THE NEED TO EXPLORE BORROWER RELIABILITY ANALYTICS APPROACHES**

**SYSTEM PROPOSAL**

**SUBMITTED BY,**

**LIZPENCER ADHIAMBO OKELLO**

**DDA-01-0047/2024**

**A RESEARCH SYSTEM PROPOSAL SUBMITTED IN FULFILMENT FOR PROGRESSION TO THE SYSTEM PROJECT, TO THE DEPARTMENT OF ICT AND ENGINEERING OF ZETECH UNIVERSITY**

**AUG 2025**

# DECLARATION

The contents of this paper are a result of personal research, and all external sources from reading have been appropriately referenced. This paper does not contain any observable or nuanced AI-generated content. No version of this research paper has been plagiarized or submitted in this exact version to any other institution or agency. It is therefore an original, non-plagiarized, non-Al-generated document.

**Researcher’s Name**: LIZPENCER ADHIAMBO

**Admission Number**: DDA-01-0047/2024

**Sign:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

# ABSTRACT

The research conducted focuses on the challenge of unclear loan repayment trends among youth members of SACCOs, with special attention to Jipange Youth SACCO in Makueni County. It explores how borrower reliability analytics, including data analysis, scoring systems, and machine learning models, can help SACCOs identify reliable borrowers and make more informed, risk-based lending decisions.

This study does not involve new data collection but relies entirely on recent and credible literature published between 2020 and 2024. These sources include academic journals, SACCO reports, and practical case studies from digital lending platforms. The research compares traditional borrower assessment methods, such as personal interviews and guarantor records, with modern, data-driven approaches that offer greater accuracy and consistency in evaluating repayment behavior.

From the analysis, a simple and cost-effective Excel-based scoring tool is proposed. This tool will enable SACCO staff to assess repayment risk using existing borrower information, without the need for advanced software or technical expertise. It provides a practical solution that can be implemented within a short time frame and fits the current capacity of SACCO operations.

The paper is organized into three chapters. Chapter One presents the background, statement of the problem, research objectives, and scope. Chapter Two covers the literature review, while Chapter Three outlines the proposed system, including its objectives, features, and development timeline. The document is prepared using APA 7 formatting and reviewed by the ICT department standards.

**TABLE OF CONTENTS**

[DECLARATION i](#_Toc204683443)

[ABSTRACT ii](#_Toc204683444)

[LIST OF TABLES v](#_Toc204683445)

[DEFINITION OF KEY TERMS vi](#_Toc204683446)

[ABBREVIATIONS AND ACRONYMS vii](#_Toc204683447)

[CHAPTER ONE: RESEARCH OVERVIEW 1](#_Toc204683448)

[1.1Background information 1](#_Toc204683449)

[1.2 Statement of the Problem 1](#_Toc204683450)

[1.3 Study Justification 1](#_Toc204683451)

[1.4 Research Objectives 2](#_Toc204683452)

[1.4.1 General Objective 2](#_Toc204683453)

[1.4.2 Specific Objectives 2](#_Toc204683454)

[1.5 Scope of the Study 2](#_Toc204683455)

[1.6 Research Questions 2](#_Toc204683456)

[CHAPTER TWO: LITERATURE REVIEW 4](#_Toc204683457)

[2.1 Introduction to Literature 4](#_Toc204683458)

[2.2 The Review. 4](#_Toc204683459)

[2.2.1 Borrower Attributes That Influence Repayment Behavior 4](#_Toc204683460)

[2.2.2 Use of Data Analytics to Identify Reliable Borrowers 4](#_Toc204683461)

[2.2.3 Machine Learning Models in Borrower Risk Evaluation 5](#_Toc204683462)

[2.2.4 Role of Borrower Scoring in Reducing Financial Risk 5](#_Toc204683463)

[2.2.5 Importance of Relevant Financial Data in Credibility Assessment 5](#_Toc204683464)

[2.2.6 A Viable Borrower Reliability System for SACCOs 5](#_Toc204683465)

[2.2.7 Comparing Traditional Methods and Data-Driven Approaches in Borrower Evaluation 6](#_Toc204683466)

[2.2.8 Addressing Policy Gaps to Support Borrower Scoring Systems 6](#_Toc204683467)

[2.2.9 Impact of Repayment Patterns on SACCO Sustainability 6](#_Toc204683468)

[2.3 The Conclusion of Literature 6](#_Toc204683469)

[CHAPTER THREE: WORKPLAN 8](#_Toc204683470)

[3.1 Introduction 8](#_Toc204683471)

[3.2 Feasibility Assessment 8](#_Toc204683472)

[3.3 Project Justification 9](#_Toc204683473)

[3.4 System Objectives 9](#_Toc204683474)

[3.4.1 General Objectives 9](#_Toc204683475)

[3.4.2 Specific Objectives 9](#_Toc204683476)

[3.5 Functional Requirements 10](#_Toc204683477)

[3.6 Project Schedule Breakdown (Gantt Chart) 11](#_Toc204683478)

[REFERENCES 12](#_Toc204683479)

# LIST OF TABLES

Table 3.2 Feasibility Assessment Table

Table 3.5 Functional Requirements Table

Table 3.6 Project Schedule Gantt Chart

# DEFINITION OF KEY TERMS

**Borrower Reliability Analytics-**A method of analyzing data related to borrowers past behaviors to predict their likelihood of repaying future loans.

**Credit Assessment**-Process of evaluating a borrower's trustworthiness using structured information before approving a loan.

**Failure to repay a loan according to the agreed**-upon terms and conditions results in a financial loss for the lending institution.

**Financial Stability**-The ability of a financial institution, such as SACCO, to maintain its operations without suffering significant losses, ensuring it can meet its repayment obligations.

**The Loan disbursement process** involves providing funds to a borrower, which can be in the form of loans or credit.

**Loan Officer**-The SACCO staff member who inputs applicant information

**Repayment-**The patterns and habits exhibited by borrowers in repaying their loans, which can vary based on individual circumstances and economic conditions.

**SACCO (Savings and Credit Cooperative Organization**)-A member-owned financial institution that provides savings and loan services to its members.

**Unclear Loan Repayment**-The situation where repayment outcomes are inconsistent or unpredictable, making it hard for SACCO staff to identify reliable borrowers.

# ABBREVIATIONS AND ACRONYMS

**AI**-Artificial Intelligent

**CBK** -Central Bank of Kenya

**CSV**-Comma Separated Value

**EDA**-Explanatory Data Analysis

**KUSCCO-** Kenya Union of Savings and Credit Cooperatives

**ML**-Machine Learning

**MS-**Microsoft

**SACCO**-Savings and Credit Cooperative Organization

**SQL-**Structured Query Language

# CHAPTER ONE: RESEARCH OVERVIEW

## 1.1Background information

Jipange YouthSACCO *in Makueni County* provides loans to young people running small businesses. However, while some members repay on time, others delay or completely default, creating repayment inconsistencies. SACCO’s treasurer admits that it is difficult to identify in advance which borrowers are most reliable. This is a common challenge faced by many youth-focused SACCOs in developing countries, where formal credit histories are rare, and lending is often based on trust, guarantors, or basic member records.

With youth borrowers forming a *growing segment of SACCO clientele*, the financial future of institutions like Jipange depends on how well they can assess risk and promote responsible borrowing. Traditional evaluation methods are limited, especially when irregular income patterns, lack of business training, or weak follow-up systems influence borrower behavior. On the other hand, *advances in data analytics* offer an opportunity to analyze past loan records, identify repayment patterns, and develop tools that more accurately predict borrower reliability. This study examines the opportunity, informed by Jipange SACCO’s experience, to develop practical solutions that enhance loan decision-making and foster sustainable youth entrepreneurship.

## 1.2 Statement of the Problem

Jipange Youth SACCO faces a growing problem of *unclear loan repayment trends* among its youth borrowers. Despite a steady increase in loan disbursement, the repayment behavior of members aged 18–35 remains inconsistent. Some repay early, others delay, while some default completely. The current manual and judgment-based borrower assessment system makes it hard for SACCO to anticipate and manage default risk. This uncertainty puts SACCO’s *financial stability* at risk. Without a reliable way to analyze borrower behavior and identify trustworthy clients, SACCO cannot effectively grow its youth lending program or ensure repayment sustainability. There is a *pressing need to explore borrower reliability analytics* to better understand repayment trends and support smarter lending decisions.

## 1.3 Study Justification

This study is important because it focuses on improving financial decision-making in a SACCO that serves young entrepreneurs, a group with high potential but also high risk. By reading into the repayment challenges Jipange SACCO is facing and exploring how real-life digital solutions work, *including their software, hardware, data, and skills requirements,* the researcher aims to gain a deeper understanding of what makes some borrowers more reliable than others. By comparing multiple solutions and understanding how long each takes to build, what it costs, and what it needs in terms of technical expertise, the study will support an informed judgment on the most practical approach to implement. This insight will be crucial in selecting the best possible project to develop in the next semester.

## 1.4 Research Objectives

### 1.4.1 General Objective

1. To analyze unclear loan repayment trends among youth borrowers in Jipange SACCO and explore borrower reliability analytics approaches to support risk-informed lending decisions.

### 1.4.2 Specific Objectives

1. To assess the repayment trends of youth loans issued by Jipange SACCO from 2021 to 2024.
2. To explore borrower attributes that influence repayment behavior.
3. To investigate the main reasons behind repayment delays and defaults among youth.
4. To review the current borrower evaluation methods used at Jipange SACCO.
5. To compare traditional borrower screening techniques with modern data-driven models.

(V) To evaluate how machine learning can be applied to predict borrower risk

1. To identify the technical requirements (software, hardware, skills) for implementing such systems.
2. To examine how borrower reliability analytics can support financial stability at Jipange SACCO.
3. To recommend a viable and feasible approach to borrower risk assessment for system development in the next semester.

## 1.5 Scope of the Study

This study focuses on exploring how data analytics can be used to understand borrower reliability at Jipange Youth Sacco in Makueni County. The primary objective is to examine the issue of unclear or inconsistent loan repayment patterns among young borrowers. The study will consider individual Sacco members *aged between 18 and 35 years*, using data from *2021 to 2024.*

It does not cover corporate or group loans. Instead, it focuses on analyzing personal borrowing and repayment records to find patterns that could help improve how the Sacco assesses and manages loan risks among youth.

## 1.6 Research Questions

1. What borrower attributes contribute most to their repayment behavior in Jipange Youth Sacco?
2. How can data analytics be used to identify reliable borrowers in Jipange Youth Sacco?
3. What algorithm or scoring method can aid in evaluating borrower reliability?
4. How will borrower reliability scores enable Jipange Youth Sacco to reduce financial risk?
5. What are the main factors affecting the financial stability of Jipange Youth Sacco due to unclear repayment trends?
6. What financial data is most relevant in assessing borrower credibility?
7. How can machine learning methods outperform traditional methods in evaluating borrower reliability?
8. What policy measures can be implemented alongside borrower scoring to aid financial stability?
9. How do borrower repayment trends affect the operational sustainability of Jipange Youth Sacco?
10. What is the significance of developing a borrower reliability algorithm for financial institutions in Kenya?

# CHAPTER TWO: LITERATURE REVIEW

## 2.1 Introduction to Literature

This chapter reviews recent studies conducted between 2020 and 2024 that relate to borrower behavior, repayment patterns, and the application of data analytics in loan risk assessment. It draws upon credible sources, including peer-reviewed journals *such as MDPI and SpringerLink, institutional reports from the Central Bank of Kenya, and practical case studies from financial technology (FinTech)* platforms. These sources are credible because they are grounded in real-world evidence, reviewed by experts, and focused on practical financial outcomes. The literature examines the factors that influence repayment behavior, the functioning of borrower scoring models, and the role of digital solutions in supporting youth-based financial institutions.

Additionally, the review identifies where different authors agree or disagree, while highlighting the strengths and limitations of each approach. Comparisons between traditional and data-driven methods are discussed to provide a balanced view. The researcher draws on these findings to build an informed foundation for the proposed system and make a judgment about which approach is most suitable for implementation at Jipange SACCO.

## 2.2 The Review.

### 2.2.1 Borrower Attributes That Influence Repayment Behavior

Borrower repayment behavior is often influenced by specific characteristics such as age, income level, employment status, business type, and prior loan history. Studies such as *Gitau et al. (2021)* found that self-employed youth with seasonal incomes were more likely to delay repayments compared to those with steady salaried employment. Similarly, a 2023 CBK report showed that repayment consistency improved when borrowers received financial literacy training. The pros of understanding borrower traits include better risk segmentation and targeted support. However, the drawback is that over-reliance on static borrower attributes can lead to generalization and overlook behavioral factors. This analysis helps SACCOs like Jipange personalize their credit decisions.

### 2.2.2 Use of Data Analytics to Identify Reliable Borrowers

Data analytics enable SACCOs to move beyond intuition and guesswork. By processing historical loan data, institutions can generate predictive insights. For instance*, Mwacharo (2022)* reported that integrating data dashboards in community banks reduced default rates by 18%. Fintech platforms like Tala and Branch have also demonstrated that mobile phone usage patterns, payment frequency, and airtime purchases can be indicators of creditworthiness. While analytics tools are powerful, they require clean data and skilled analysts. The advantages are efficiency and scalability, while the limitation is the risk of algorithm bias if the algorithm is poorly trained. Overall, analytics provides a structured, evidence-based path for assessing borrower trustworthiness.

### 2.2.3 Machine Learning Models in Borrower Risk Evaluation

Machine learning has become a preferred tool in borrower scoring due to its adaptability. Unlike traditional models, machine learning algorithms learn from trends and patterns without requiring explicit programming. According to a 2020 study published in MDPI, random forest models outperformed logistic regression by 15% in predicting defaults among microfinance borrowers. While these tools offer accuracy and adaptability, they also require significant computational resources and technical skills. SACCOs like Jipange can benefit from these tools if they invest in capacity-building or partner with technology providers. Strength lies in real-time risk scoring, while the limitation is complexity.

### 2.2.4 Role of Borrower Scoring in Reducing Financial Risk

Borrower scoring translates borrower behavior into quantifiable scores that help lenders estimate default probability. A World Bank *(2021)* report highlighted how SACCOs that adopted score-based lending models had improved loan performance by 23%. These scores allow institutions to segment risk, price loans accordingly, and improve portfolio quality. The main pro is objective decision-making; however, scores may not capture real-time shocks such as illness or emergencies. For Jipange, scoring provides a foundation for fairness and accountability, but it must be supplemented with flexible policies.

### 2.2.5 Importance of Relevant Financial Data in Credibility Assessment

Accurate credibility assessment depends on reliable financial data, including repayment history, cash flow statements, and savings behavior. A study by *Kirui (2022)* found that data completeness was the biggest bottleneck in SACCO credit analysis. SACCOs using digital record keeping had better borrower performance tracking than those using manual ledgers. The benefit of using quality financial data is improved analysis and prediction. However, poor data collection practices or inconsistent updates reduce the usefulness of analytics tools. For Jipange, investing in better data collection systems is a critical step toward predictive reliability modeling.

### 2.2.6 A Viable Borrower Reliability System for SACCOs

A reliable system for assessing borrower reliability should combine historical data analysis, scoring techniques, and machine learning models tailored to SACCO dynamics. Comparing solutions such as credit scoring platforms, mobile-based data trackers, and custom in-house tools reveals differences in cost, setup time, and technical complexity. For example, third-party tools like Credit Info are ready-made but may not suit SACCO-specific needs, while in-house tools offer control but require more resources. A viable system for Jipange should be cost-effective, use available member data, require moderate training, and support gradual improvement. From a judgment perspective, a hybrid system that blends traditional knowledge with data-driven scoring seemsmost feasible.

### 2.2.7 Comparing Traditional Methods and Data-Driven Approaches in Borrower Evaluation

Traditional borrower assessments in SACCOs often rely on face-to-face interviews, community reputation, and limited paper-based records. While these methods are familiar and accessible, they lack precision and consistency. In contrast, data-driven tools such as Excel-based scoring, Power BI dashboards, and Python models provide structured, evidence-based decision-making. The advantage of modern tools lies in their ability to scale, reduce bias, and provide timely insights. However, they require training and proper data management. For Jipange SACCO, shifting from intuition to analytics presents a learning curve, but it could strengthen loan performance and accountability.

### 2.2.8 Addressing Policy Gaps to Support Borrower Scoring Systems

Introducing a borrower scoring system requires more than just technology; it also needs proper policies. SACCOs must establish guidelines on data use, borrower privacy, and how scores influence loan decisions. Without such frameworks, there is a risk of misuse or confusion. A study by Karanja (2022) emphasized that SACCOs with clear digital lending policies saw smoother adoption of scoring tools. Policy development helps ensure fairness, transparency, and member trust. For Jipange, aligning the scoring system with internal rules and national data protection standards will be essential for smooth integration.

### 2.2.9 Impact of Repayment Patterns on SACCO Sustainability

Repayment behavior directly affects SACCO’s ability to grow and serve its members. Frequent delays or defaults reduce cash flow, hinder lending capacity, and can damage SACCO’s credit reputation. According to a 2023 report by the Kenya Union of Savings and Credit Cooperatives (KUSCCO), irregular repayments are the leading cause of liquidity strain in youth-focused SACCOs. Reliable borrowers, on the other hand, support long-term sustainability. For Jipange, mapping these trends through analytics can help predict risks and take initiative-taking measures to safeguard operations.

## 2.3 The Conclusion of Literature

The literature review provides a comprehensive foundation for understanding borrower reliability analytics in SACCO settings, particularly in the context of youth borrowers. Key findings show that borrower attributes such as income type, financial discipline, and prior history significantly influence repayment behavior. The integration of data analytics emerges as a powerful approach to improve borrower assessment, allowing SACCOs to move beyond subjective methods.

Machine learning models offer advanced capabilities in predicting risk, though they require technical expertise and quality data to function effectively. Borrower scoring plays a critical role in reducing financial risk by standardizing evaluations, while relevant financial data remains a vital resource for any successful assessment model. Importantly, the comparison of available digital tools and system requirements highlighted the need for customized, cost-effective, and practical solutions for SACCOs, such as Jipange.

From the review, combining data-driven tools with traditional knowledge enhances decision-making and supports long-term sustainability. These insights will directly inform the development of a tailored borrower reliability assessment system, guiding the selection of methods and technologies in the subsequent phases of the study.

# CHAPTER THREE: WORKPLAN

## 3.1 Introduction

This section outlines the preparation for the project that will soon begin. After taking a deep look into the issue of unclear loan repayment trends at Jipange Youth Sacco and understanding how this affects Sacco’s ability to identify reliable borrowers, the study explored various digital approaches that could help solve this challenge. These included real-life systems that use data analytics and scoring models, along with an examination of their features, how they work, and what it would take, both in terms of tools, skills, and time, to build them. With all that background in place, this chapter now focuses on deciding which solution is most practical and achievable. It presents a detailed plan covering the system’s functional requirements, the objectives of the project, the software and hardware needed, and the specific skills required to implement the solution. It also includes a timeline that breaks down the entire process step by step, from the start of development to the final stage, planned for next semester.

## 3.2 Feasibility Assessment

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Project Alternative** | **Technical Feasibility** | **Budget Feasibility** | **Schedule Feasibility** | **Conclusion** |
| Excel-based scoring tool | Laptop-MS Excel  Basic Excel skills | No extra cost | 2-3 Weeks | Feasible |
| Python Machine Learning Model | Laptop with  Internet  Python libraries  Coding skills | Free tools | 4-5 weeks | Feasible |
| Power BI dashboard | Basic data cleaning and visualization skills | Power BI Desktop is free | 3-4 weeks | Feasible |
| Web app with a dashboard | Front-end + backend coding, hosting setup, and higher development skills | Hosting domain host is involved; time and technical skills are demanding | 8-10 weeks | Not feasible |

***Table 3.2 Feasibility Assessment Table***

## 3.3 Project Justification

After comparing different system alternatives based on cost, required skills, and the time it would take to develop each one, the Excel-Based Scoring Tool emerged as the most suitable choice for this project. All the other options, including Python machine learning, Power BI dashboards, and web applications, were useful in their own ways, but they either needed extra resources, advanced technical skills, or a longer period to implement.

The Excel tool, on the other hand, is simple, cost-effective, and easy to understand. Most SACCO staff already use Excel for basic operations, so this option will not require more training. It also fits well with the time allocated for this project and can be completed within 2 to 3 weeks. Even though Excel is not as advanced as some other tools, it will still allow Sacco to enter borrower details, apply basic scoring formulas, and generate meaningful reports for loan evaluation.

While future versions of the system could explore more powerful options like machine learning, for now, Excel offers a practical starting point that meets Sacco’s needs and the limits of this academic project. Choosing it means the project can deliver real value within a short time, without stretching beyond Sacco’s current capabilities.

## 3.4 System Objectives

### 3.4.1 General Objectives

Developing an Excel-Based Tool to Assess Borrower Reliability for Improved Loan Decisions at Jipange Youth SACCO

### 3.4.2 Specific Objectives

1. To provide a clean borrower data input interface in Excel for entering details like age, loan history, and savings behavior.
2. To automatically assign and calculate weighted scores based on borrower characteristics that influence repayment ability.
3. To generate a final borrower reliability score using embedded Excel formulas and logic.
4. To categorize borrowers into defined risk levels, such as high-risk, medium-risk, or low-risk, based on their score range.
5. To offer dynamic visual reports such as charts or dashboards summarizing overall borrower trends and risk distribution.
6. To allow for easy updates or edits to borrower records so the tool remains reusable and current over time.

## 3.5 Functional Requirements

|  |  |
| --- | --- |
| **Participants/case user** | **Functional Requirements** |
| Loan Officer | -Input borrower details such as name, age, loan history, income, and savings into a structured Excel sheet.  -View automatically calculated borrower scores based on weighted attributes.  -Update or edit borrower records without damaging the structure of the tool. |
| Credit Committee member | -Review of the borrower risk category (High, Medium, Low) generated from final scores. |
| System Administrator | -Maintain the scoring formulas and ensure correct logic is applied in the calculations.  -Backup data regularly and ensure the tool is functioning correctly across all devices using it. |
| SACCO Manager | -Generate and print summary reports of borrower scores and risk categories for use in meetings. |

***Table 3.5 Functional Requirements Table***

## 3.6 Project Schedule Breakdown (Gantt Chart)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **WEEKS** | **PROJECT MILESTONES** | | | |
| **1** | **2** | **3** | **4** |
| **Project Design & Modeling** | **Project Development & Testing** | **Project Deployment** | **Project Presentation** |
| 1. 1-5 Sep |  |  |  |  |
| 1. 8-12 Sep |  |  |  |  |
| 1. 15-19Sep |  |  |  |  |
| 1. 22-26 Sep |  |  |  |  |
| 1. 29 Sep – 3 Oct |  |  |  |  |
| 1. 6 – 10 Oct |  |  |  |  |
| 1. 13 – 17 Oct |  |  |  |  |
| 1. 20-24 Oct |  |  |  |  |
| 1. 27-31 Oct |  |  |  |  |
| 1. 3-7 Nov |  |  |  |  |
| 1. 10-14 Nov |  |  |  |  |
| 1. 17-21 Nov |  |  |  |  |
| 1. 24-28 Nov |  |  |  |  |

***Table 3.6 Project Schedule Gantt Chart***

# REFERENCES

Central Bank of Kenya. (2023).“*Bank supervision annual report 2023”.*https://www.centralbank.go.ke/uploads/banking \_sector\_annual\_reports/2023.pdf

Gitau, M. ,Njoroge, L, & Chege, p. (2021).”*Socioeconomic Factors Affecting Loan Repayment Among Youth in SACCOs in Kenya. International Journal of Finance and Development, 5(2),44-58.*

Kirui, T. K. (2022).”D*ata management challenges in SACCO credit scoring systems.” A*frica journal of information Systems, 10(1), 31-45.https://ajis.org/article/data-management -Sacco

Mwacharo, R. (2022).*” Using data dashboards to manage microfinance defaults:”* A case of Kenyan community banks. Journal of Financial Technology and Innovation, 4(3),65-78. https://jfti.org/article/data-dashboardkenya.

World Bank.(2021).”*Digital financial service:* “How borrower scoring improves credit access in East Africa. Https://www.worldbank.org/en/publication/dfscore-ea-2021