



VIT[®]
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

Vellore – 632014, Tamil Nadu, India
SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

Fall Semester 2025-2026

Case Study 1

SLOT : F2
Programme Name : B.Tech.
Course Name & code : Financial Data Analytics & BCSE336L
Class Number : VL2025260101796
Faculty Name : Dr. Mehfooza M

Due date: 30/09/2025

AI-Powered Credit Risk Assessment Using Alternative Data in R

Case Study Question:

Collect **10 research papers** that apply **machine learning techniques** (e.g., Logistic Regression, Random Forest, Neural Networks) for credit scoring using alternative data (e.g., mobile transactions, social media, or utility payments). Implement a credit risk model in **R** using non-traditional data sources, compare performance against conventional models, and analyze how your approach addresses financial inclusion challenges. Your submission must include:

1. **Abstract** (Summary of objectives, methods, and key findings)
2. **Keywords** (e.g., "Alternative Credit Scoring," "Machine Learning," "Financial Inclusion," "R Programming")
3. **Introduction** (Background on credit risk assessment, problem statement, and literature review from the 10 papers)
4. **Methodology** (Data collection, preprocessing, feature engineering, and model training in R)
5. **Comparison and Interpretation** (Performance metrics: AUC-ROC, Precision-Recall, F1-Score vs. traditional bureau scores)
6. **Future Work** (Challenges in deployment, ethical AI considerations)
7. **References** (Proper citations of the 10 papers)

Dataset Options:

- **Mobile Money Transaction Logs** (e.g., simulated M-PESA data from GSMA)
- **Utility Payment Histories** (e.g., water/electricity bills from World Bank datasets)
- **Social Media Activity Patterns** (e.g., anonymized metadata from research partnerships)

Expected Deliverables:

- A structured report (convertible to a conference paper)
- R code for data processing, model training, and fairness auditing
- Performance comparison table (your model vs. FICO-based approaches)
- **Ethical Impact Assessment** (Using frameworks like O'Neil Risk Checklist)