Project Summary

This project presents an end-to-end credit risk assessment system, utilizing a dataset of 500 synthetic loan applicants to closely simulate industry-standard lending and underwriting workflows. The system integrates financial metric analysis, rule-based credit policies, and predictive modeling to assess borrower creditworthiness and estimate default probability, reflecting the best practices in modern banking risk assessment (TrustDecision, 2025; OCC, 2017; Community Banking Connections, 2024).

Dataset Construction & Feature Engineering

A synthetic applicant portfolio was created with attributes typical of banking data—annual income, existing debt, loan amount and term, credit score, collateral value, and employment status. Core risk indicators were engineered, including Debt-to-Income (DTI) and Loan-to-Value (LTV) ratios, monthly debt obligations, and categorical risk tiers. These indicators mirror the quantifiable inputs used in contemporary risk models and serve as foundational variables for both rule-based and statistical decision-making (TrustDecision, 2025; Community Banking Connections, 2024).

Rule-Based Underwriting Policy

A structured underwriting framework was developed, operationalizing realistic lending rules:

- Approval for strong applicants (Credit Score ≥ 680, DTI $\leq 40\%$, LTV $\leq 80\%$)
- Review for intermediate risk (Credit Score 640–679, moderate DTI or LTV)
- Rejection for high-risk (Credit Score < 640, DTI $> 45\%$, LTV $> 100\%$)

This approach is consistent with industry practices, distinguishing risk levels through transparent criteria that align with regulatory requirements and internal policy (OCC, 2017; Community Banking Connections, 2024).

Default Labeling

To facilitate predictive modeling, applicants were labeled as default (1) if they exhibited at least one high-risk indicator—Credit Score < 620, DTI $> 45\%$, or LTV $> 100\%$. This method provides a realistic target variable for model training and mirrors key default triggers acknowledged by lenders (TrustDecision, 2025).

Predictive Modeling

A logistic regression classifier was trained using the engineered risk features, following an 80/20 train-test split and proper data preprocessing. The model outputs both binary and probabilistic

default predictions, supporting risk-aligned lending decisions. Results were exported for further portfolio analysis, demonstrating best practices in the integration of predictive analytics within credit workflows (TrustDecision, 2025; Nected, 2025).

Portfolio Insights and Impact

Dual analysis of rule-based and model-based decisions produced actionable insights:

- DTI and LTV were the strongest predictors of risk.
- Model recommendations closely mirrored banking-sector risk stratification.
- The portfolio exhibited a realistic risk distribution, facilitating effective loan monitoring and resource allocation.

Practical Relevance

The project demonstrates how structured analytics and data-driven models can increase the consistency, fairness, and transparency of credit decisions while reducing institutional risk exposure. This mirrors contemporary expectations of risk management in financial institutions and connects with broader themes like risk-based pricing, regulatory compliance, and operational efficiency (OCC, 2017; NCino, 2024).

Conclusion

The complete process—data generation, feature engineering, risk rule development, and statistical modeling—exemplifies modern credit evaluation lifecycle management. These skills are directly applicable to domains such as cybersecurity risk assessment and anomaly detection, highlighting the transferability of risk analytics expertise.

Link to all files: Please click here

References

Bhati, M. (2025, October 13). Credit risk assessment guide: 7 proven techniques & monitoring tips. *Nected.Ai*. https://www.nected.ai/us/blog-us/credit-risk-assessment

Comptroller's Handbook. (n.d.). *Rating Credit Risk*. Occ.gov. Retrieved November 21, 2025, from https://www.occ.gov/publications-and-resources/publications/comptrollers-handbook/files/rating-credit-risk/pub-ch-rating-credit-risk.pdf

The importance of credit risk management for FIs. (n.d.). nCino. Retrieved November 21, 2025, from https://www.ncino.com/blog/importance-credit-risk-management-financial-institutions

The importance of loan risk rating systems. (n.d.). Communitybankingconnections.org. Retrieved November 21, 2025, from https://www.communitybankingconnections.org/Articles/2025/R2/the-importance-of-loan-risk-rating-systems

What is credit risk assessment? A beginner's guide. (n.d.). Trustdecision.com. Retrieved November 21, 2025, from https://trustdecision.com/articles/what-is-credit-risk-assessment-a-beginners-guide