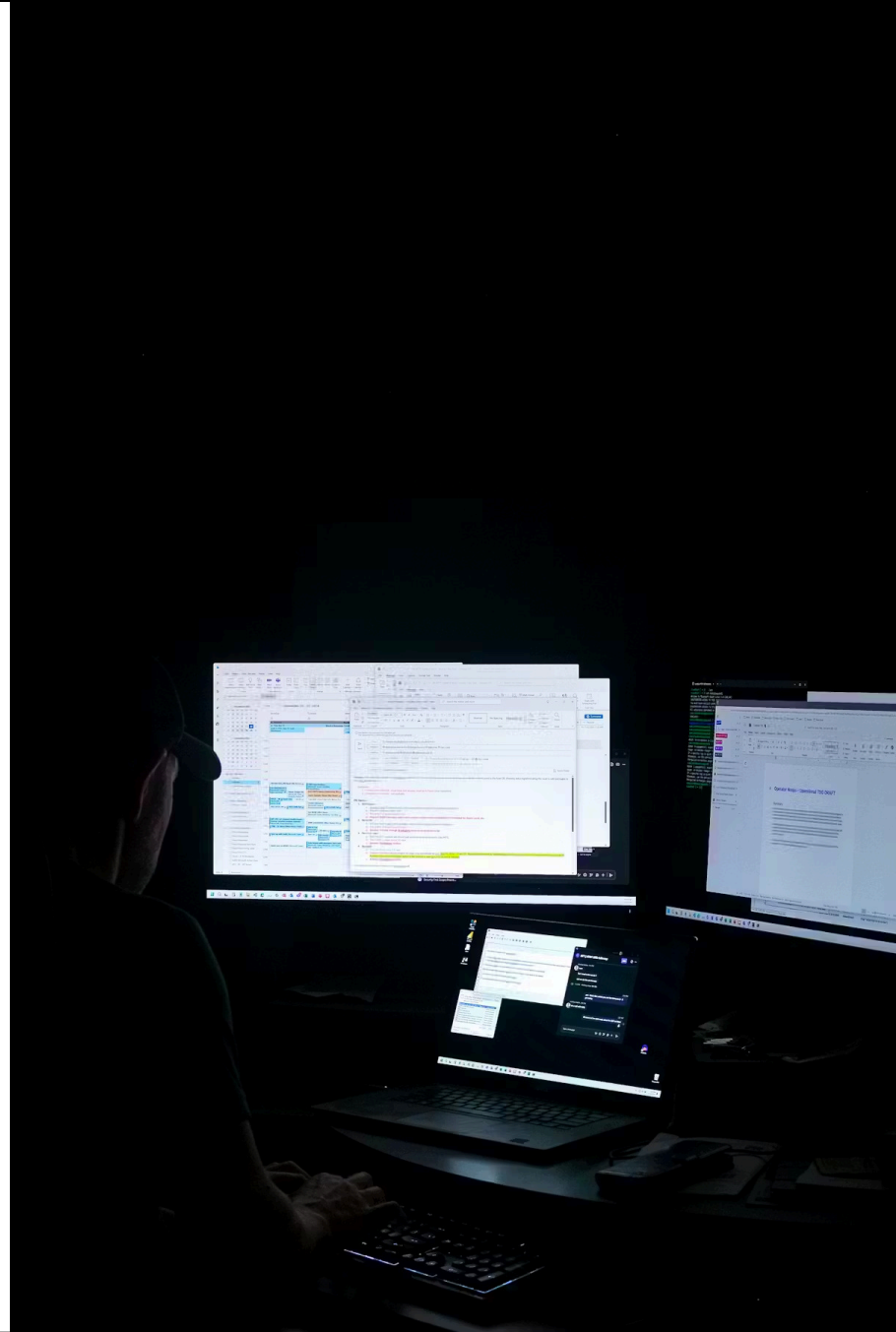
NeuronAI

Loan Default Prediction System

Dataset & Merging

Data Preprocessing Pipeline

- Merged application, credit history, demographics, financial rations, geographic and load details data by unique identifiers.
- Removed unnecessary signs and noisy random columns
- Normalized values: formats
- Removed duplicates and handled missing values
- Filtered the ID columns



Dataset Overview

Key Features

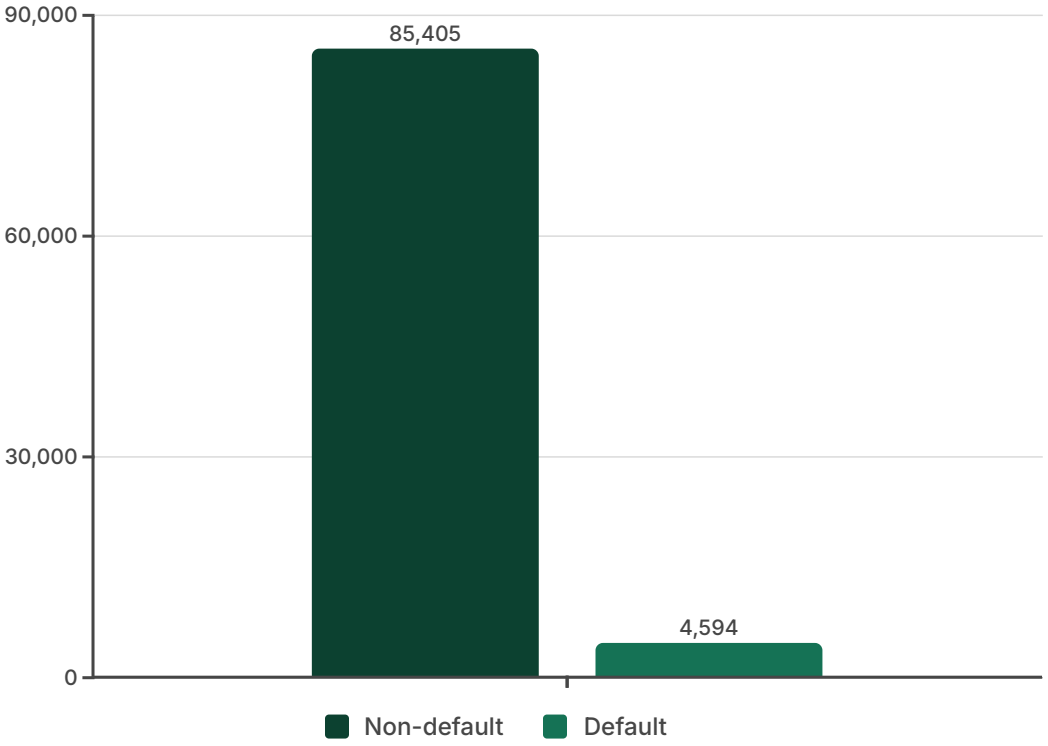
Credit score, income levels, age demographics, account metrics, login behaviour, and customer service interactions.

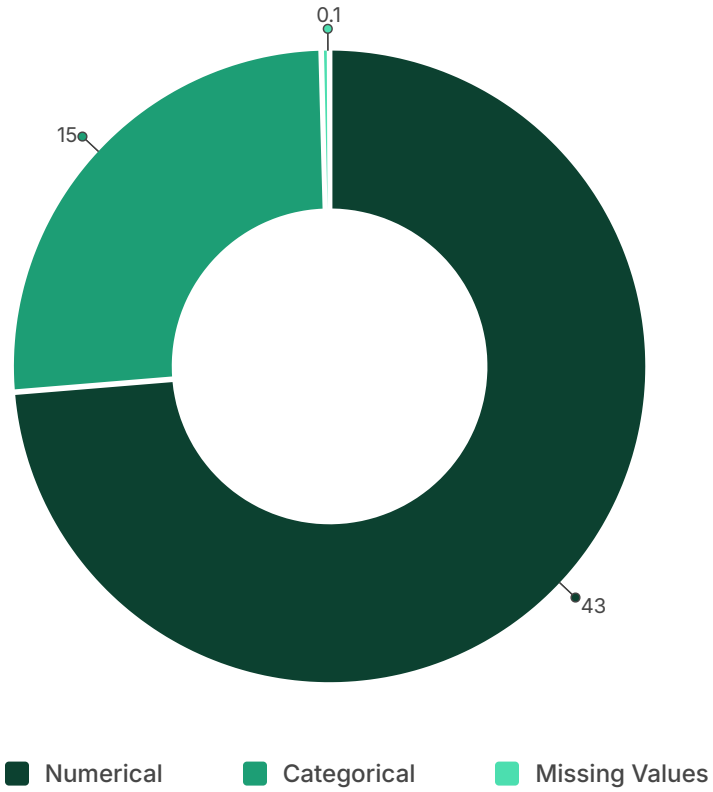
Data Structure

Comprehensive dataset with categorical and numerical variables capturing customer financial behaviour and demographics.

Status Codes

Binary target variable indicating default vs non-default outcomes for credit risk assessment.





Analytics & Insights

Feature Importance Analysis

Behavioural Metrics Lead

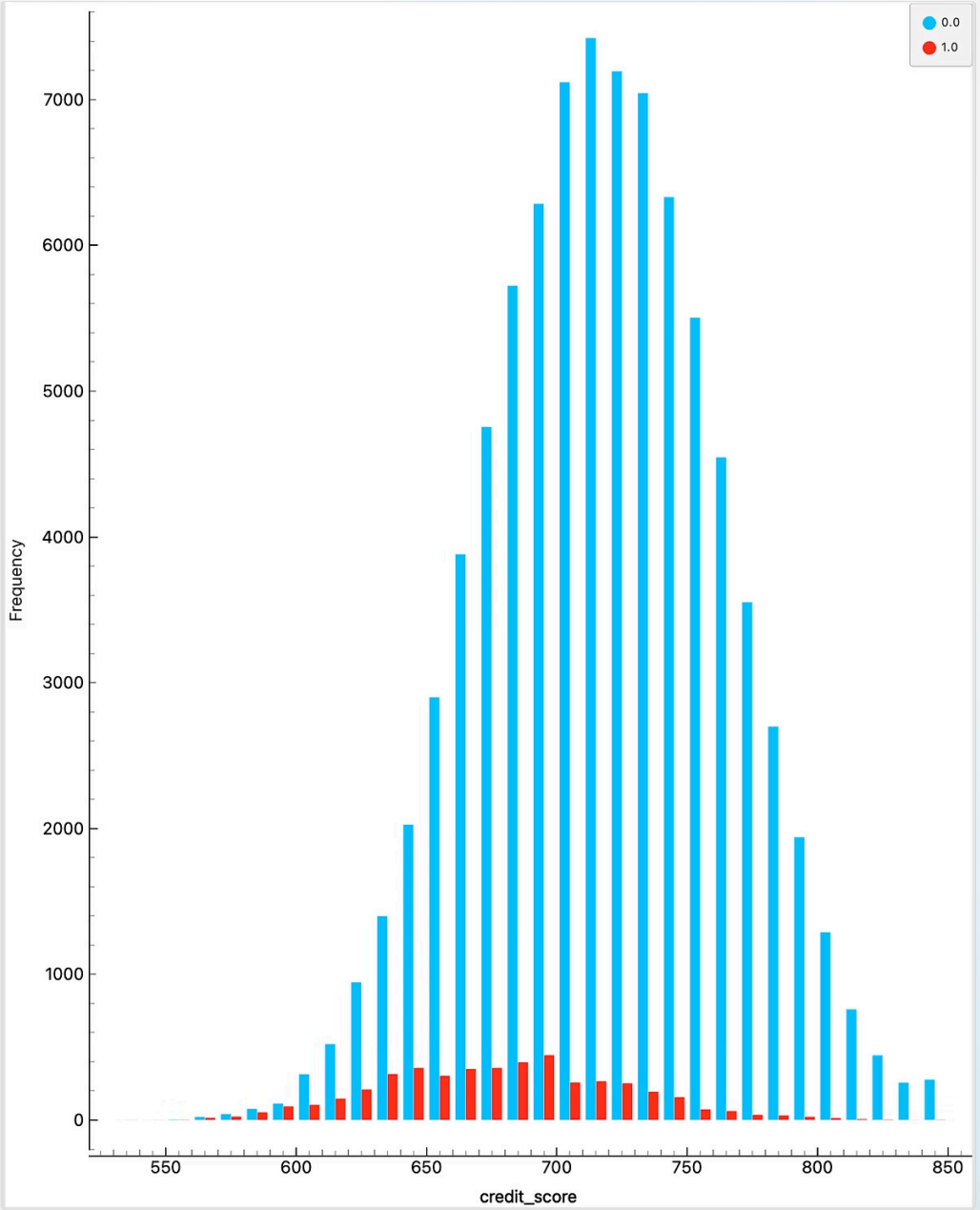
Customer service calls and login patterns show strongest correlation with default risk, outperforming traditional demographic factors.

Income Significance

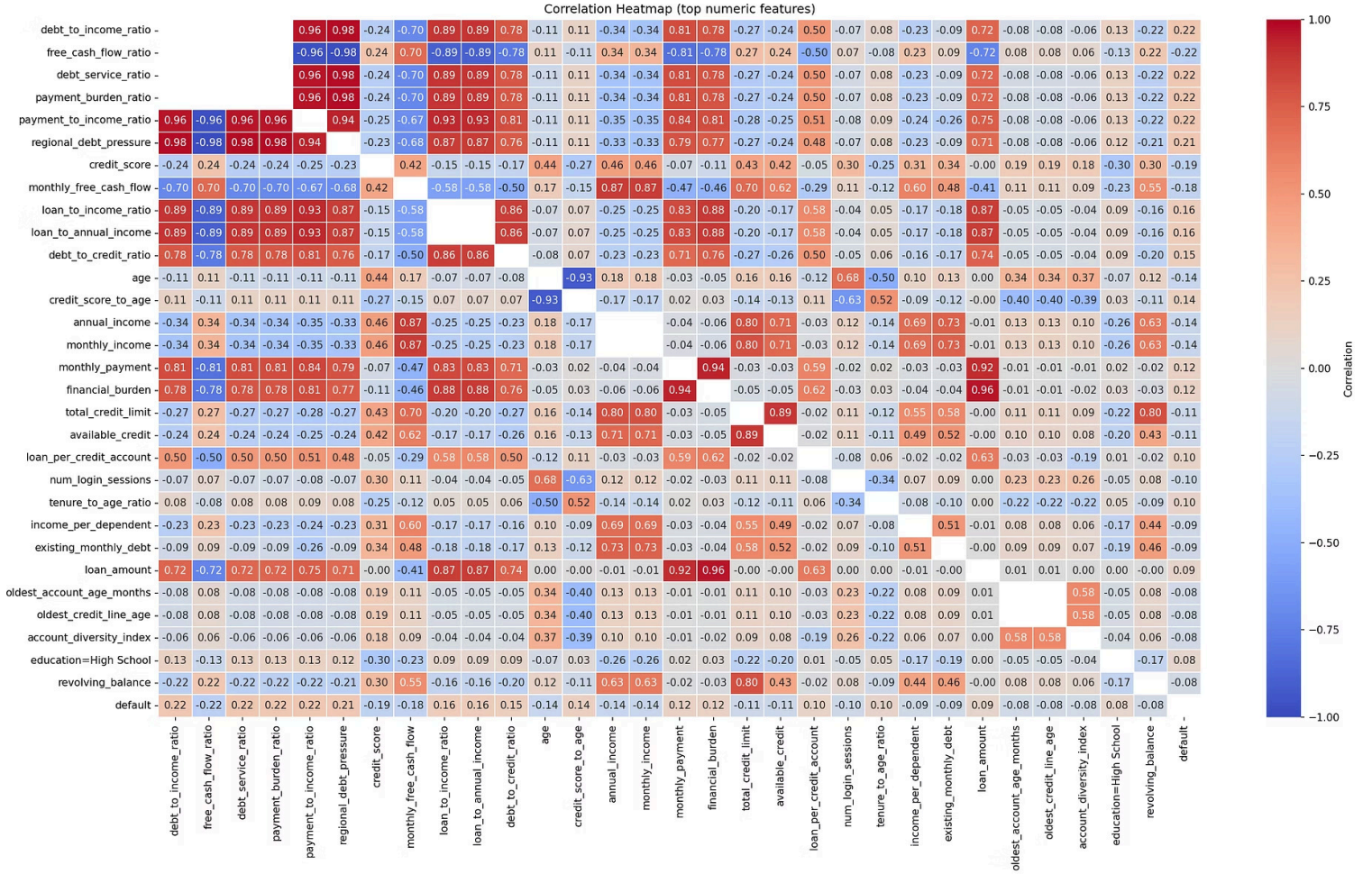
Annual income demonstrates moderate predictive power when combined with account age and transaction behaviour.

Account Activity

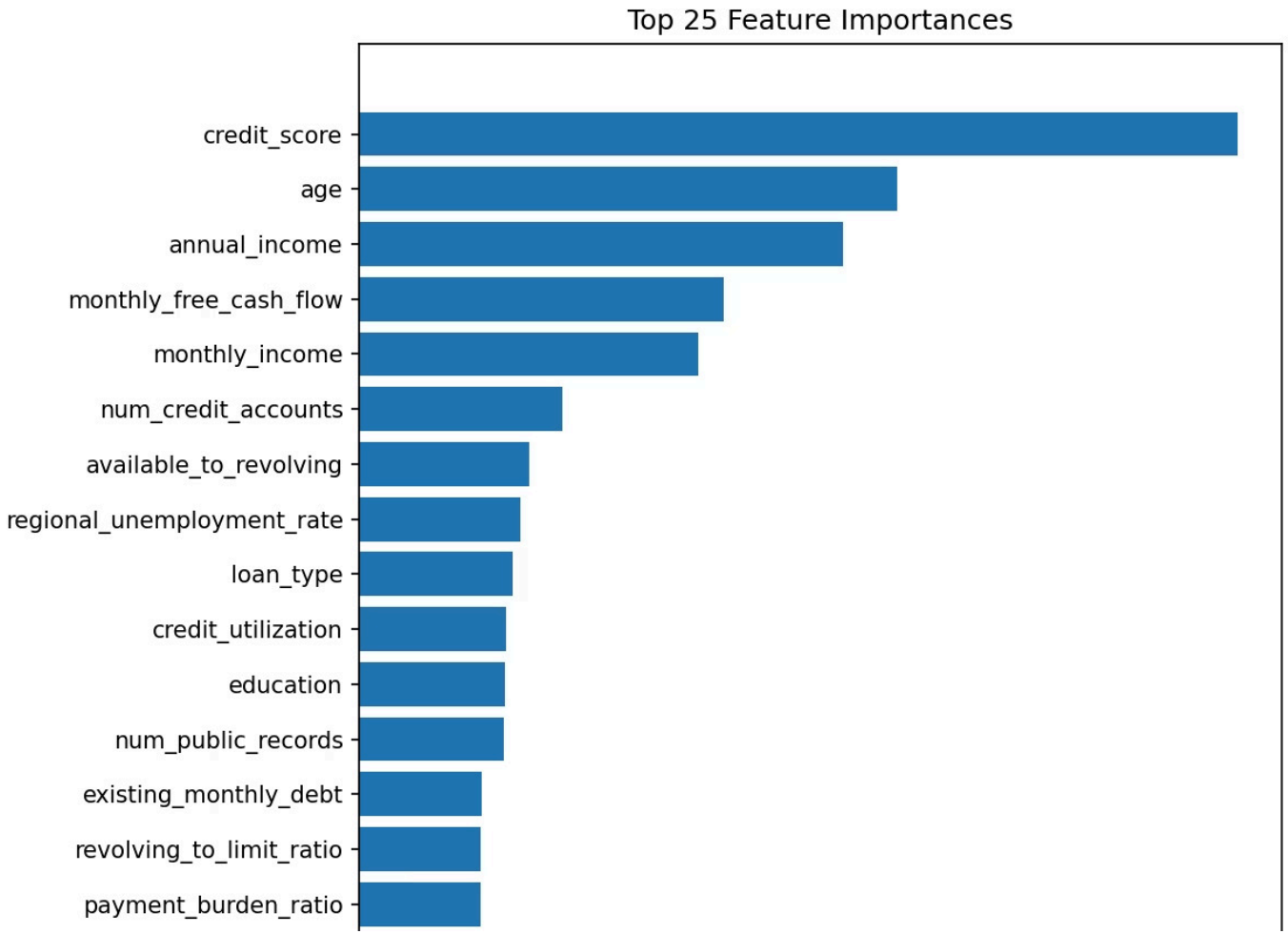
Account age and transaction frequency provide crucial context for assessing creditworthiness and default probability.



Correlation diagram



Important features



AI Model Architecture

1

Feature Engineering

Created interaction features and temporal patterns from raw data and removed outliers.

2

Encoding

Applied target encoding for categorical variables.

3

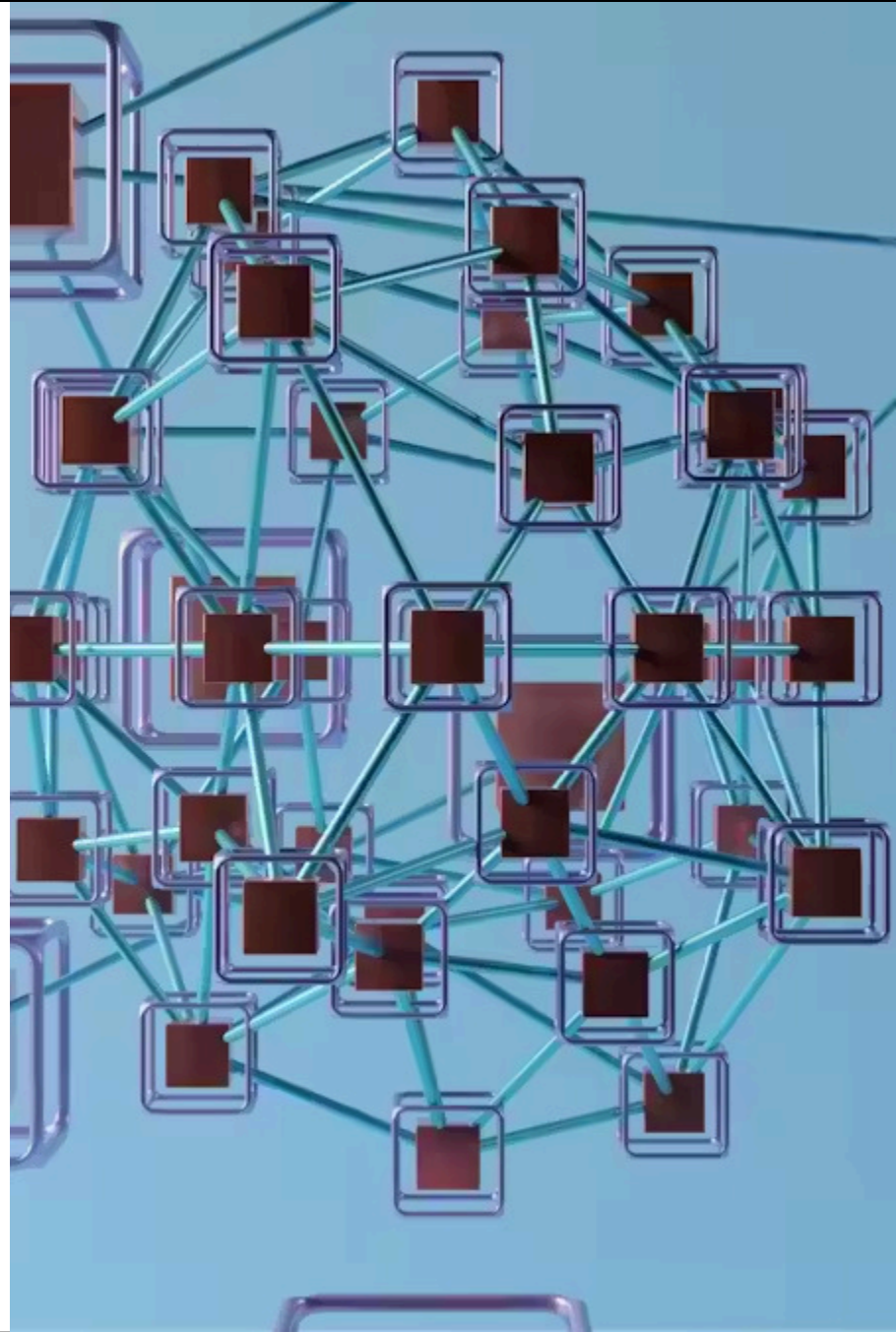
Balancing

Used Random Undersampling to address class imbalance.

4

Model Selection

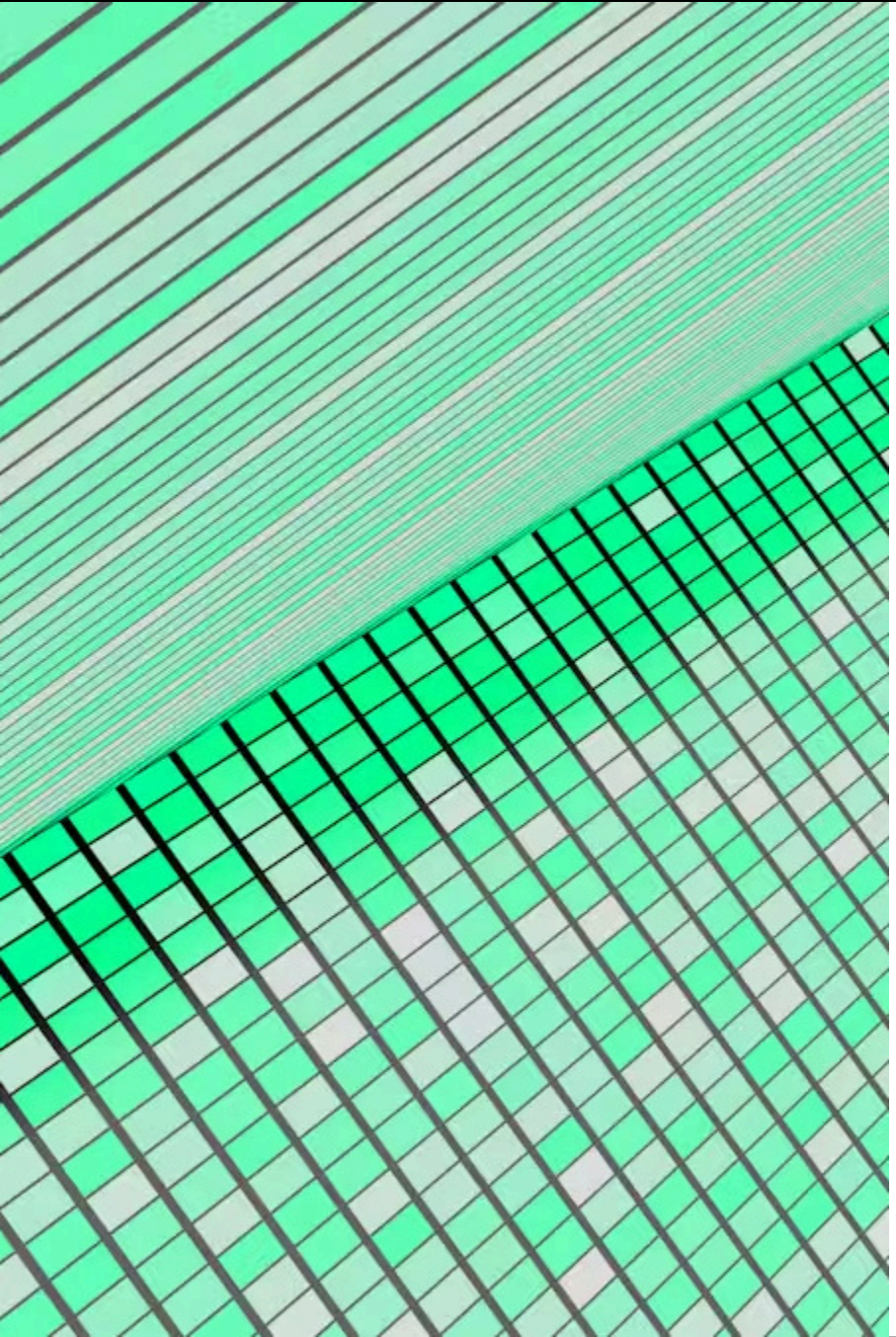
Evaluated CatBoost, XGBoost, LightGBM and Random Forest.



Model Performance Comparison

A detailed comparison of various machine learning models used in our analysis, highlighting their performance across key metrics.

Method/Model	Precision (default)	Recall (default)	F1 Score	AUC	GINI
CatBoost	**0.56**	**0.72**	**0.66**	**0.84**	**0.61**
XGBoost	0.23	0.35	0.25	0.74	0.49
LightGBM	0.19	0.26	0.24	0.80	0.60
Random Forest	0.18	0.32	0.23	0.74	0.48



Model Performance

Metrics Achieved

84%

AUC Score

Good discrimination between default and non-default cases

66%

F1 Score

Balances of precision and recall

61%

Gini

Strong discriminatory power

Our Team

Makhmudov Sunnat

Software Engineer

BSc at IUT

NeuronAI LLC

~10 years experience in IT

Ergashev Adizbek

Full-Stack Developer

PHD at TUIT

NeuronAI LLC

~10 years experience in IT

G'ayratov Muhriddin

Project Manager

MSc at Turin(TTPU)

NeuronAI LLC

~5 years experience in IT

Mamaroziqov Kamron

AI Engineer

BSc at New Uzbekistan university

NeuronAI LLC

~3 years experience in IT