

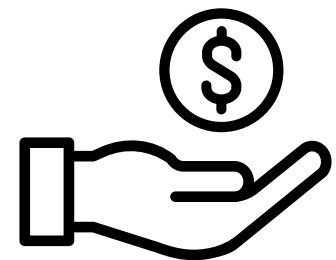
UPDATE 2

LendingClub Case

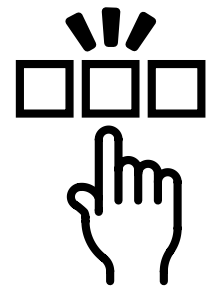
Yang Yue, Ted Wang, Goyo Lozano Palacio



A **data-driven investment strategy** is all about running the right race and running it well

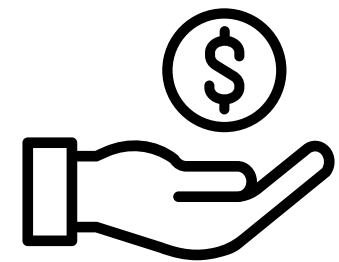


Which loans should you select and invest in?



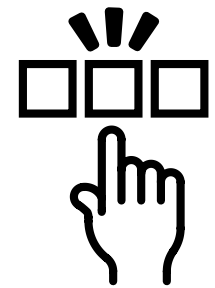
How do we select those loans?

A **data-driven investment strategy** is all about running the right race and running it well



Which loans should you select and invest in?

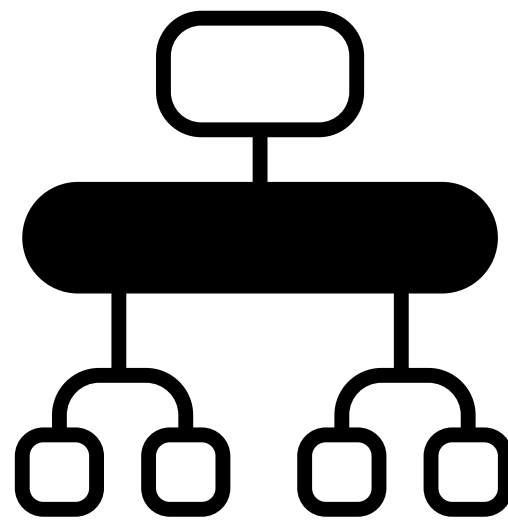
Loans with the most return on investment and lowest risk



How do we select those loans?

Employ different strategies that select loans with highest return, lowest risk, etc.

The **decision tree classifier** predicts whether a loan defaults...



It outputs a **label**:
0 - the loan does not default
1 - the loan defaults

... using a **set of rules** that are easy to interpret



loan term > 36 months



renting home ownership
status > 0.5



debt-to-income ratio <= 0.043

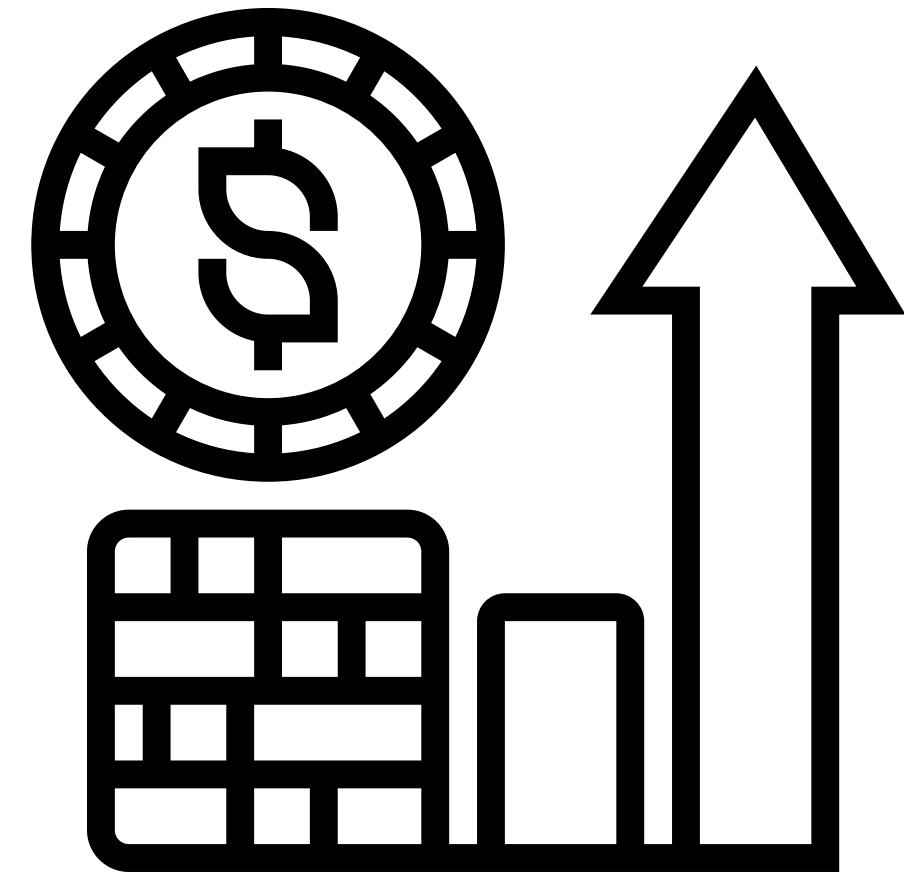
The **logistic regressor** estimates
a number...



A value:

A number estimating an intermediate
return value at a fixed interest rate

... so that we can predict the
return on investment



Our initial hypotheses line up with our findings

- *Loans with higher grades = lower default rates*
Supported by EDA in Update 1
- *A higher debt-to-income = increases the likelihood of default.*
Supported - by decision tree and logistic regression
- *Higher annual income is inversely related to default rates.*
Supported - logistic regression showed higher annual income decreased odds of default.
- *A higher number of delinquencies = higher default rate.*
Weakly supported by logistic regression but impact is smaller than other features.



An **intermediate return** assumes a realistic balance between overly pessimistic and optimistic estimations

Pessimistic

Annualizes returns over the *entire original loan term* (treating early repayments as if the money were left idle)

Optimistic

Annualizes returns based on the *actual period* the loan was active and assumes immediate reinvestment

Intermediate

Computes return for a fixed investment horizon by reinvesting cash flows at predetermined modest rates

Random Selection

5%

Tiered Grade Selection

7.6%

Conditional Combined
Expected Return

6.7%

Comparing **investment
strategies** reveals the **best**
portfolio option

Lowest Default Risk

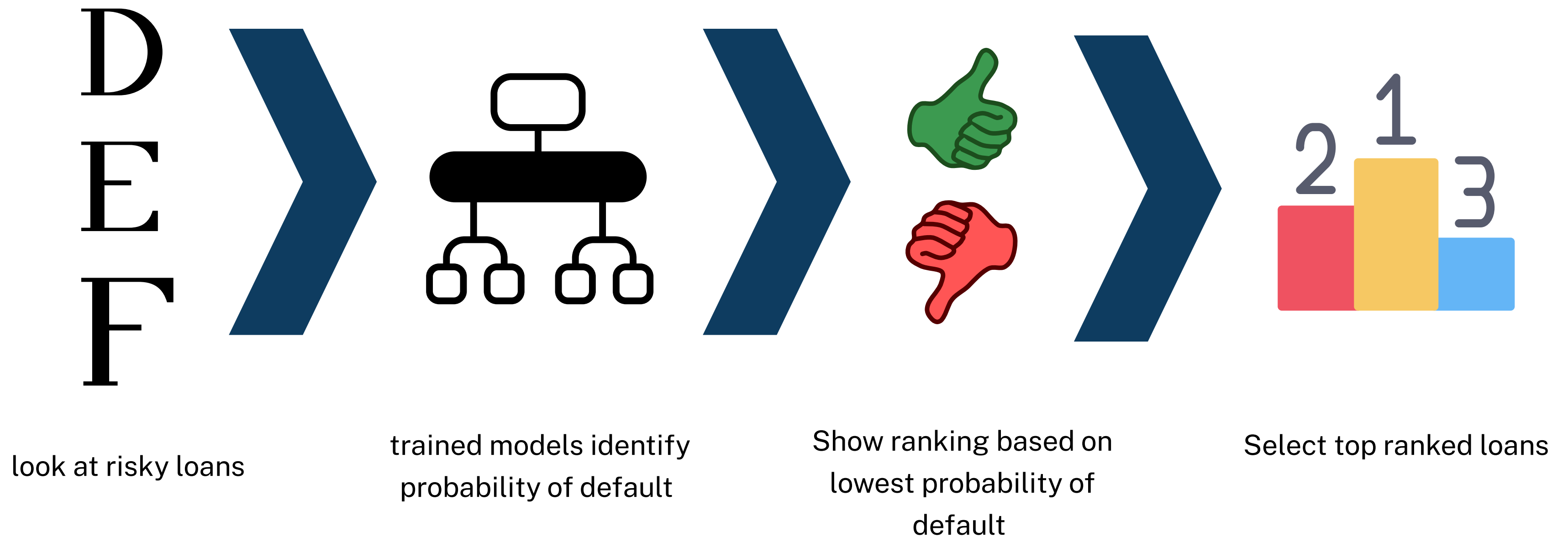
5.2%

Highest Predicted Return

5.4%

*for 100 selected loans

A **tiered grade strategy** maximizes return by considering riskier loans and optimizing for success



Our models and strategies are **robust**, **flexible**, and **scalable**

STRATEGY	20 LOANS	100 LOANS	1000 LOANS
Random	5.94%	5.05%	4.24%
Lowest Probability of Default	4.9%	5.24%	5.24%
Highest Predicted Return	1.44%	5.42%	5.94%
Highest Combined Expected Return	9.22%	6.77%	6.48%
Tiered Grade Selection	7.13%	7.64%	10.58%