Approve or deny?

Classifying credit applicants

unine
Université de Neuchâtel

Aritz Lizoain

31st May 2022

Goal

Accurately and efficiently determine if new applicants present good or bad credit risk

Current procedure

- Time consuming
- Requires the lender to check the applicant's background and information
- Fixed criteria based on knowledge

Classification model

- Automatic, instant
- Provides insights into most important factors that determine the final decision
- Criteria based on knowledge + experience

Data

- 1000 credit applicants
- 30 variables (credit history, purpose of credit, credit amount, etc.)

Process

- A machine learning model is trained with the data
- The model classifies the credit rating as good/bad
- The model is mainly focused on correctly classifying applicants with bad credit
- The model prediction is compared to the true credit rating

Initial insights

Can we guess which variables determine the decision?

Decisive variables are likely to show different tendencies for applicants with good and bad credit risk

Checking account status

- Clear distinction

- Good credit risk applicants tend to have over 200DM*, or no checking account.

- Bad credit risk applicants tend to have less than 200DM*

- However, it is not possible to approve/deny an application based uniquely on one variable


Checking account status
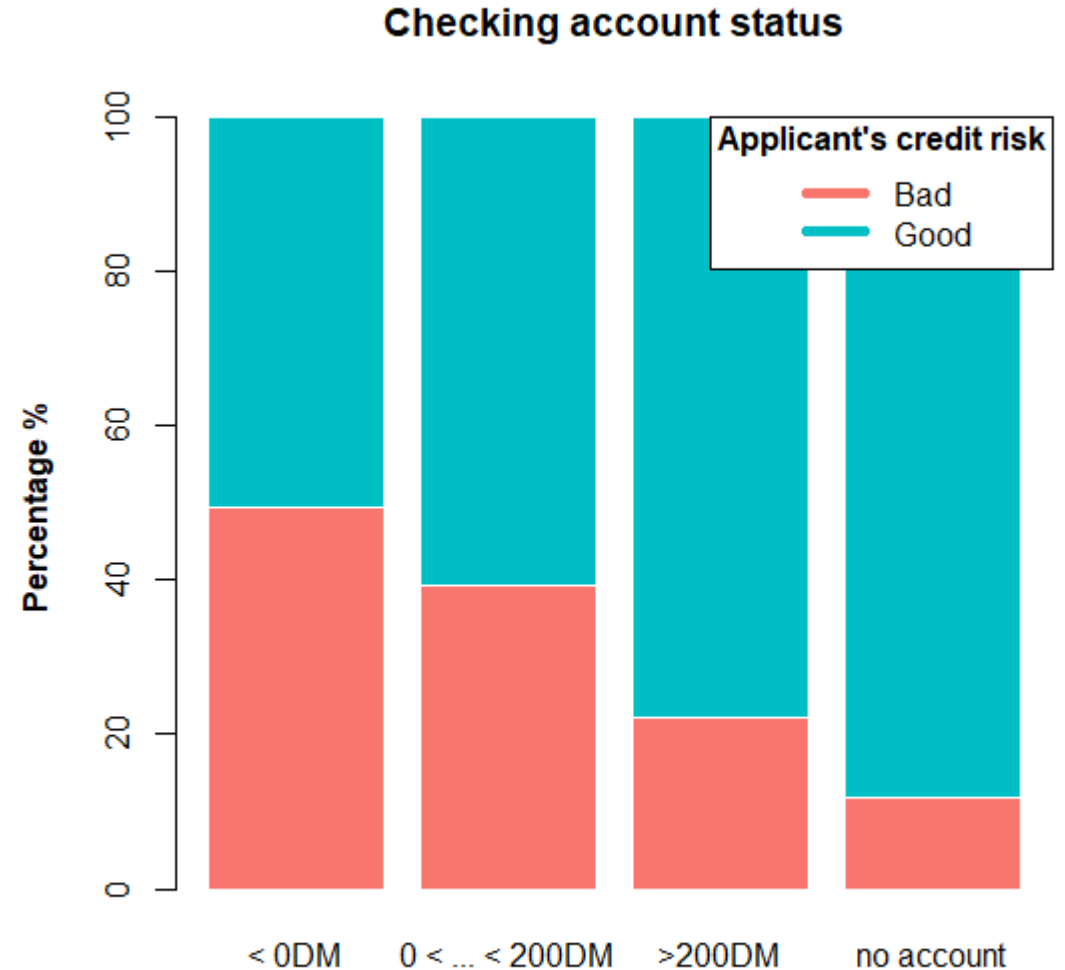
* official currency of West Germany from 1948 until 1990

Results

- The model classifies correctly 70% of applicants
- The model correctly detects 77% of applicants with bad credit risk
- Low (23%) false negative rate

Most important variables

1) Checking account status
2) Duration of credit
3) Credit amount
4) Credit history
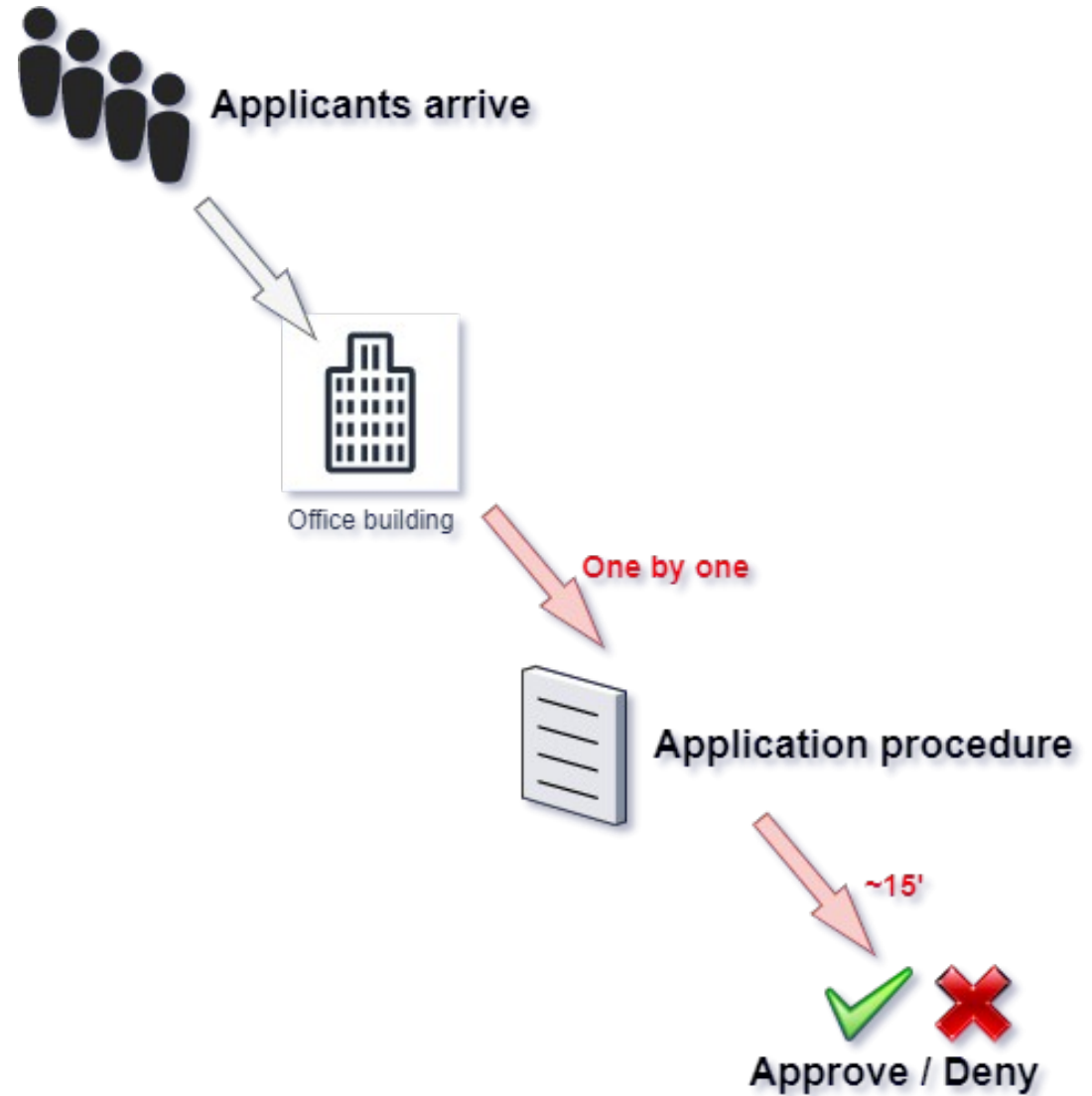5) Average balance in savings account
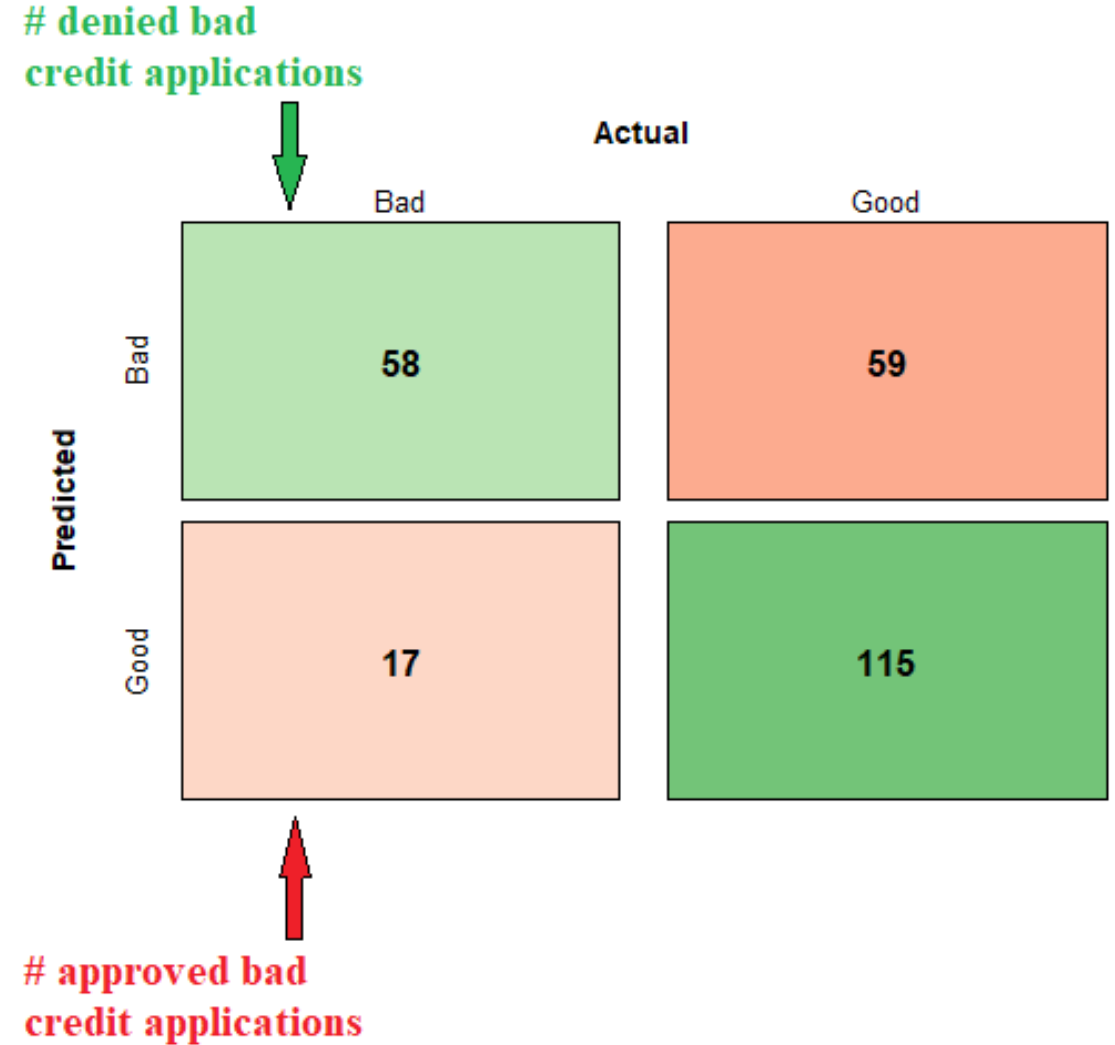
Example

4 credit applicants show up to the office

Only 1 worker is available

Each application is estimated to take 15'

Current procedure

- 1h to finish all 4 applications
- One of the clients will need to wait for 45'
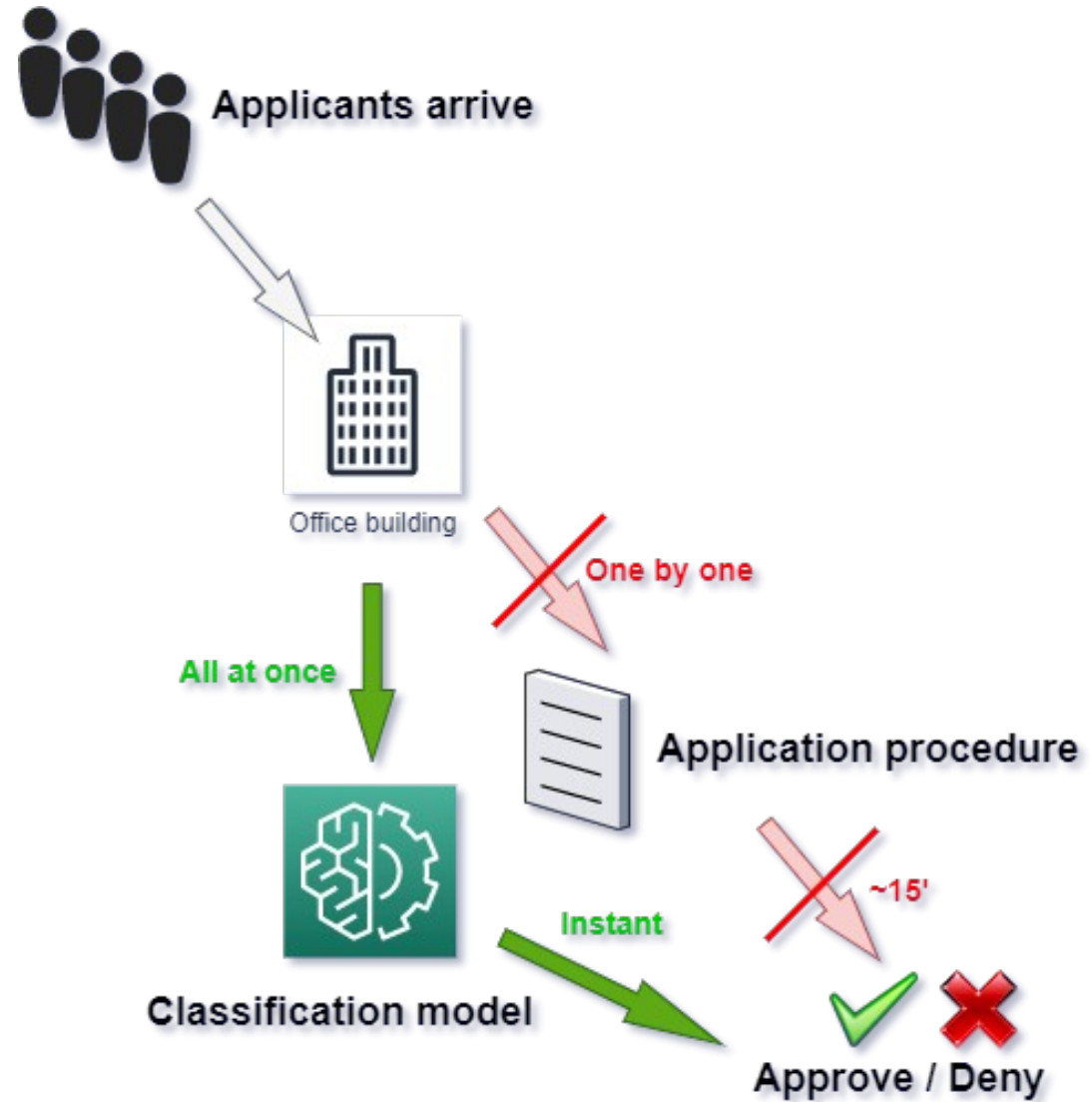- It requires for the worker to check the applicant's background and information

Applicants arrive

Office building

One by one

Application procedure

~15'

Approve / Deny

Example

4 credit applicants show up to the office

Only 1 worker is available

Each application is estimated to take 15'

Classification model

- Instant classification of all 4 applications

- No need to wait

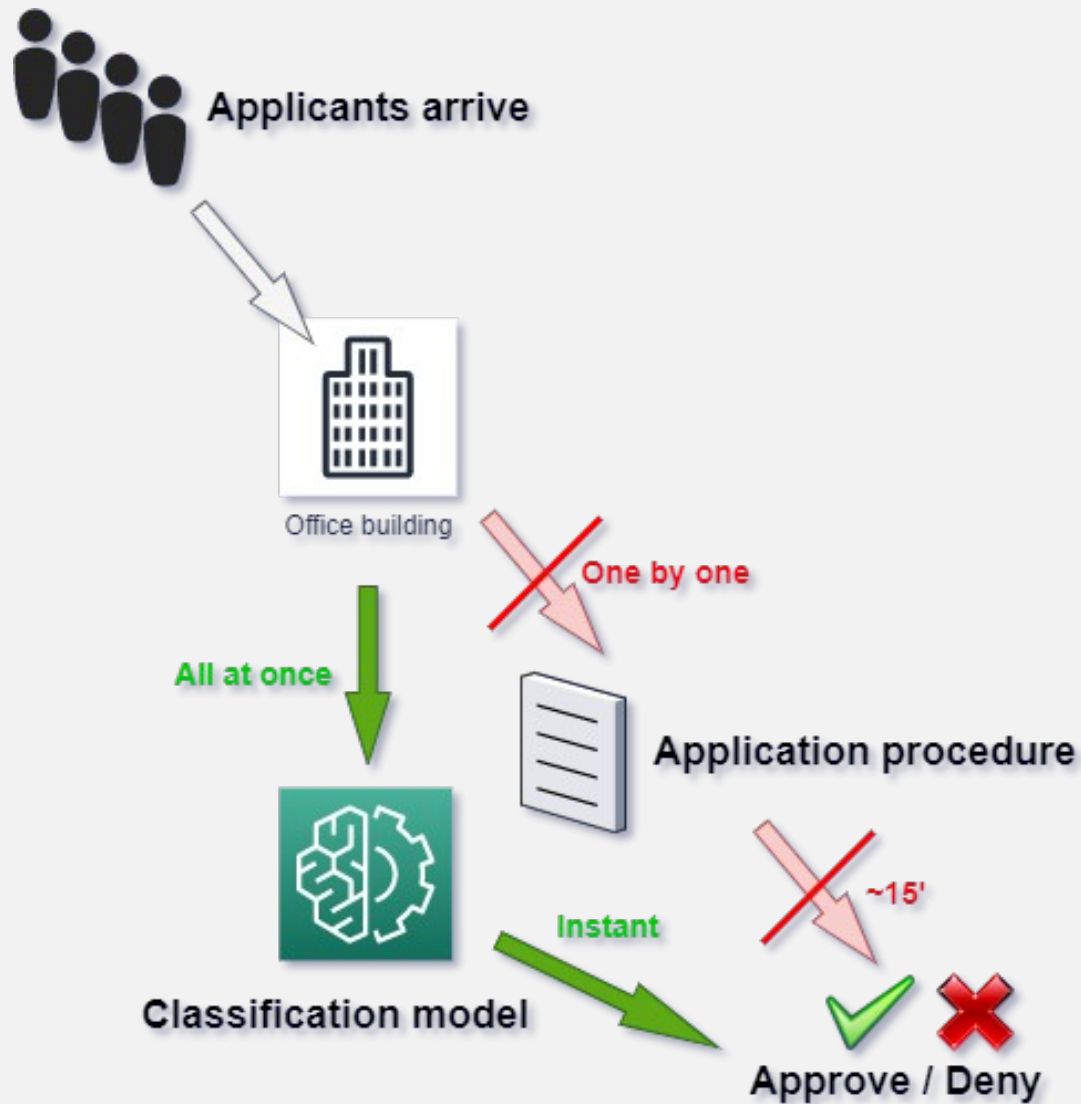- The worker can quickly verify the model's predictions, or use them to validate their decision

Applicants arrive

Office building

One by one

All at once

Application procedure

Instant

~15'

Classification model

Approve / Deny

Classification model – Random Forest

- Many "experts" vote for a decision
- Each "expert" decides which variables are most important to make a decision
- The final decision is taken based on majority voting
- Accurate and stable
- It can handle large datasets efficiently
- No new data needs to be collected – all applicant information is collected when they open a bank account

Next steps

- Currently recommended to be used as a validation tool
- Once optimized, it can fully automate the credit application process, with no supervision needed

Ready to start saving time?

Applicants arrive

Office building

One by one

All at once

Application procedure

Classification model

Instant

~15'

Approve / Deny

Classification model

- Automatic
- Instant
- Efficient
- Gain business insights

Aritz Lizoain

aritz.lizoain@unine.ch