

Credit Risk

Business Applications of Operations Research and Econometrics

Petra Tomanová

Summer Semester 2023/2024

Department of Econometrics
Prague University of Economics and Business

Roles

Roles

YOU

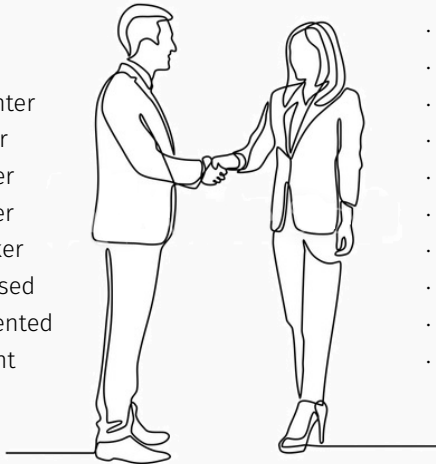
Consultant

- Smart
- Responsible
- Team Player
- Perfect Presenter
- Communicator
- Problem Solver
- Decision Maker
- Creative Thinker
- Solution Focused
- Customer Oriented
- Stress Resilient

ME

Head of Risk

- Busy
- Busy
- Busy
- Busy
- Busy
- Busy
- Busy
- Busy
- Busy
- Busy
- Busy



Project

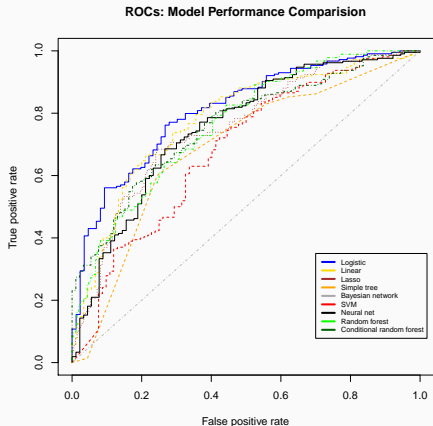
YOU are assigned to a project related to the credit risk modeling. Your goal is to find a suitable model/algorithm that answers the question: "Should the **WESTEALALLYOURMONEY BANK** provide a loan to a new customer or not?". You have approximately 1 man-day to deliver your solution (you might work overtime as consultants usually do ☺).

ME, I work in **WESTEALALLYOURMONEY BANK** as the Head of Risk. I provided you a data sample of 307,511 applications, so you can learn who our customers are. I will also give you a brief introduction to the credit risk. You are allowed to ask me questions and I will try to answer them. But remember, **I am BUSY**, so do not waste my time!

Introduction to Credit Risk

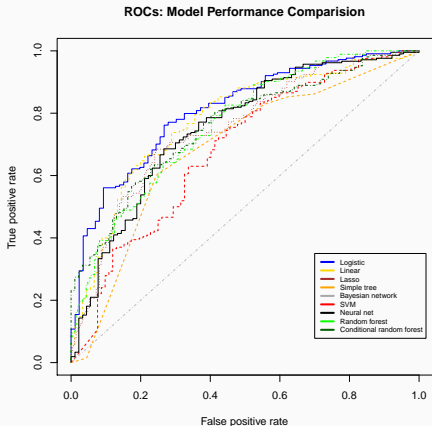
Key concepts

- Default definition, Probability of Default (PD)



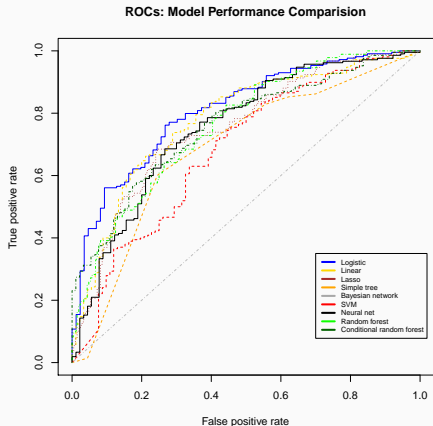
Key concepts

- Default definition, Probability of Default (PD)
- Odds



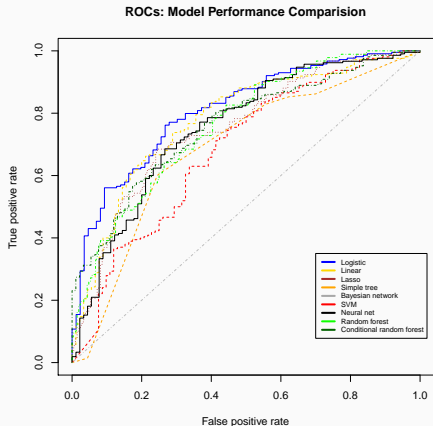
Key concepts

- Default definition, Probability of Default (PD)
- Odds
- Weight of Evidence (WOE)



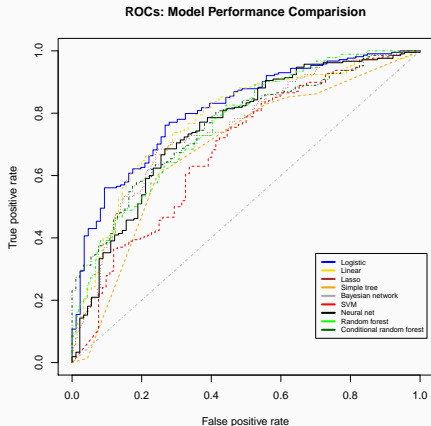
Key concepts

- Default definition, Probability of Default (PD)
- Odds
- Weight of Evidence (WOE)
- Information Value (IV)



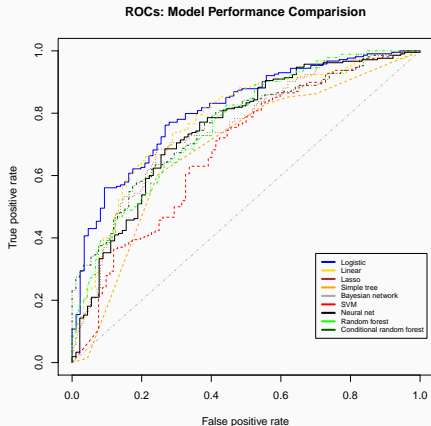
Key concepts

- Default definition, Probability of Default (PD)
- Odds
- Weight of Evidence (WOE)
- Information Value (IV)
- Gini



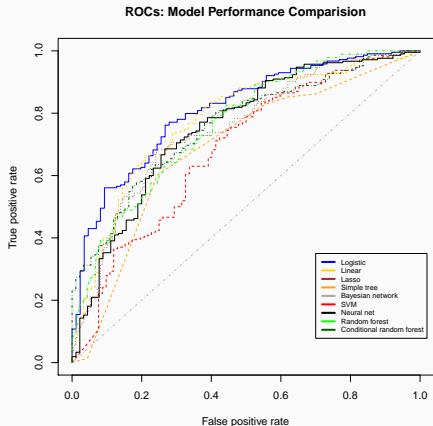
Key concepts

- Default definition, Probability of Default (PD)
- Odds
- Weight of Evidence (WOE)
- Information Value (IV)
- Gini
- Survival analysis



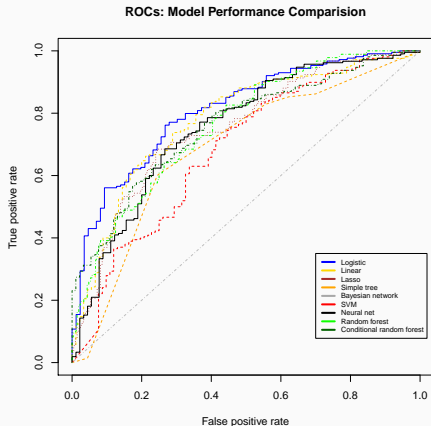
Key concepts

- Default definition, Probability of Default (PD)
- Odds
- Weight of Evidence (WOE)
- Information Value (IV)
- Gini
- Survival analysis
- Regression: logistic vs. linear



Key concepts

- Default definition, Probability of Default (PD)
- Odds
- Weight of Evidence (WOE)
- Information Value (IV)
- Gini
- Survival analysis
- Regression: logistic vs. linear
- Machine learning 😊



How to construct a model?

You might follow common steps:¹

1. Select proper **target** (default) variable.
2. Select proper **independent variables**.
3. Use different **samples** – training/testing, back test, validation.
4. **Pre-select** variables from a long list:
 - univariate Gini (e.g. $\geq 10\%$),
 - information value (e.g. $\geq 4\%$).
5. **Univariate** analysis of a short list of variables.
6. **Transformations** of variables, e.g. continuous \rightarrow categorical.
7. **Correlation** analysis I, e.g. eliminate if $\text{corr} > 0.7$.
8. **Course classification** – binning (merging of categories).

¹...taken over from my boss Pavel Charamza ☺

How to construct a model?

You might follow common steps: ~ cont.

⋮

9. **Modeling.**
10. **Correlation** analysis II, e.g. combine variables if $\text{corr} > 0.4$.
11. **Construction of 3 – 10 models:**
 - expert selection – number of variables depends on number of defaults (usually 10 – 20 for the full logistic regression).
12. **Performance measurement:** in-sample and out-sample (cross validation).

Data Description

Data Description

Data sample: *CreditData.csv* ²

- Characteristics of 307,511 loans:
 - 200,000 loans with known outcome,
 - 107,511 loans with unknown outcome.
- The **outcome** variable is *TARGET*.
 - If *TARGET* = 1 then the loan was not repaid (default).
- The data set might include **missing values**.
- The data are anonymized, only loan ID is provided (variable *SK_ID_CURR*).
- The data set includes 120 **explanatory variables**.
 - The description of all explanatory variables can be found in *columns_description.csv*.

²source: KAGGLE.COM

Assessment Criteria

Assessment Criteria

Main criteria: understanding what you are doing & having no serious issues.

Do not use complicated/fancy models you do not understand! I am the Head of Risk, I am smart, I will find out that you do not understand what you are doing.

Other criteria: number of minor issues, in-sample Gini (might not be a good measure), out-of-sample Gini (standard measure).

Keep in Mind!

Keep in mind that

1. executive summary has to be delivered until Monday, 16:00,

Keep in Mind!

Keep in mind that

1. executive summary has to be delivered until Monday, 16:00,
2. the final presentation is mandatory,

Keep in Mind!

Keep in mind that

1. executive summary has to be delivered until Monday, 16:00,
2. the final presentation is mandatory,
3. we will keep our 'Consultant & Head of Risk' roles during the presentation,

Keep in Mind!

Keep in mind that

1. executive summary has to be delivered until Monday, 16:00,
2. the final presentation is mandatory,
3. we will keep our 'Consultant & Head of Risk' roles during the presentation,
4. thus, you can be 'technical' during the meeting, however, 'business sense' of your results is important too.

? QUESTIONS ?