Business Case

Assuming you are working in a bank and a customer is requesting for loan.

You need to create descriptive (statistic), predictive (machine learning), prescriptive (qualitative) models on the loan default data.

Notes:

Please use all the attributes except the index.

Remember to split train test using default rates.

Use dummy variables for Employed.

Measurement is accuracy.

Data visualization is needed. At least use one of pandas, seaborn and matplotlib.

Use Python and Orange.

Non-Programming:

1. Please answer the question below:
2. Explain the pros and cons about your model, including limitation (can be both quantitative and qualitative).
3. How to overcome the weakness of your model (future study).
4. Any descriptive analysis you could think of for this case. Example, confidence interval.
5. What historical data variables are considered most influential in predicting loan defaults, and how are they weighted in your analysis?
6. How do economic indicators and market trends impact the accuracy of your loan default predictions, and what strategies are in place to adapt to changing conditions?
7. How do you balance the need to mitigate default risk with the goal of providing access to credit for underserved or high-risk borrowers? Are there any ethical considerations in this decision-making process?
8. The prescriptive analysis on loan default aims to enhance decision-making, reduce default risk, and optimize lending practices while maintaining a balance between profitability and risk mitigation. How do you think accurately predicting loan default can help in any decision making? (The importance of your model to the bank.)