Credit_Risk_Modelling

Credit Risk Modelling Using Machine Learning

Problem Statement

Based on customer data we are trying to predict whether to give loan or not.

Asset: All banking product. Things which give profit to the bank are asset.
Eg: Housing Loan, Car Loan, Education Loan, Credit Card Loan

Liability: Things which gives doesn't profit to the bank are asset.
Eg: Current Account, Savings Account, Fixed Deposit, Recurring Deposit

- Current Account, Savings Account are called CASA (Current Account Savings Account).
- Fixed Deposit & Recurring Deposit are called Term Deposit.

NPA

- It stands for Non Performing Asset.
- Loan that is deafaulted is known as NPA.

1. Disbursed Amount:
- Loan amount given to a customer is called Disbursed amount.
2. OSP:
- It stands for Out Standing Principle.
- 1 Lakh Loan & 8000 EMI & After paying 40000 through EMI & Left with 60000 is OSP. After Loan OSP should be 0.
3. DPD:
- It stands for Days Past Due. How many days after due date EMI amount has beeen paid.
- DPD should be ideally 0. If it is not 0 that means it is defaulted.
4. PAR:
- It stands for Portfolio At Risk.
- It means OSP when DPD > 0 days.
5. NPA:
- Loan Account When DPD > 90 days.

Credit Risk Types in Banking

1. DPD (Zero): NDA (Non Delinquint Account). No default account i.e., timely payment.
2. DPD (0 to 30): SMA1 (Standard Monitoring Account)
3. DPD (31 to 60): SMA2 (Standard Monitoring Account)
4. DPD (61 to 90): SMA3 (Standard Monitoring Account)
5. DPD (90 to 180): NPA
6. DPD (> 180): Written-Off (Loan which is not present). Bank does this to improve NPA figure. NPA improve: Loan portfolio quality will be better. So, market sentiments will be good.

Two types of NPA:

1. GNPA:

- It stands for Gross Non Performing Asset.
- If it is in range of (3% -5 %) i.e., OSP default.

2. NNPA:

- It stands for Net Non Performing Asset.
- If it is in range of (0.01 - 0.06 %) i.e., Provisioing Amount Subtracted
- When assessing Bank Quality, go for GNPA. Because this is the more accuracted parameter to check Bank Quality.
- Trade line means loan account.

Command to create exe file:

- python -m PyInstaller --onefile exe.py

Business Interpretation

- P1: Best
- P2: Second Best
- P3: Third Best
- P4: Last Explanation to business end user:
- Risk appetite: Low > target already achieve, P1
- Risk appetite: High > target are far away, P1, P2 and P3
- Risk appetite: Severely High > target are very far away, P1, P2, P3 and P4

Feedback loop:

- Based on the feedback, you will relabel the rows.
- Model retraining: Model should keep on evolving with time.

Types of Hyperparameter tuning:

- Gridsearch CV: Check all combinations so more accurate and slower
- Randomsearch CV: Choose combinations randomly so less accurate and faster
- Bayesian: Take baye's theorem into account to choose the combinations

Correlation vs Causation:

- Income vs savings - +ve correlation and +ve causation
- Ice Cream Sales vs Shark attacks - +ve correlation and no causation because of unerlying factor of warm temperature
- Master Degrees vs Box Office Revenue - +ve correlation and no causation because of unerlying factor of population increase
- Exercise vs Body Weight - -ve correlation and yes causation
- Smoking vs Lifespan - -ve correlation and yes causation
- TV time vs Better health - -ve correlation and no causation because of unerlying factor of more sleep time