In credit risk world, statistics and machine learning play an important role in solving problems related to credit risk. Hence role of predictive modelers and data scientists have become so important.

What is Credit Risk Modelling?

Credit risk modelling refers to data driven risk models which calculates the chances of a borrower defaults on loan (or credit card). If a borrower fails to repay loan, how much amount he/she owes at the time of default and how much lender would lose from the outstanding amount. In other words, we need to build probability of default, loss given default and exposure at default models as per advanced IRB approach under Basel norms.

What is Credit Risk?

In simple words, it is the risk of borrower not repaying loan, credit card or any other type of loan. Sometimes customers pay some instalments of loan but don't repay the full amount which includes principal amount plus interest. For example, you took a personal loan of USD 100,000 for 10 years at 9% interest rate. You paid a few initial instalments of loan to the bank but stopped paying afterwards. Remaining unpaid instalments are worth USD 30,000. It's a loss to the bank.

It's not restricted to retail customers but includes small, medium and big corporate houses. In news, you might have heard of Kingfisher Company became non-performing asset (NPA) which means the company had not been able to pay dues. High NPAs lead to huge financial losses to the bank which turns to reduction of interest rate on the deposit into banks. Serious honest borrowers with good credit history (credit score) would have to suffer. Hence it is essential that banks have sufficient capital to protect depositors from risks

Why Credit Risk is important?

Do you remember or aware of 2008 recession? In US, mortgage home loan were given to low creditworthy customers (individuals with poor credit score). Poor credit score indicates that one is highly likely to default on loan which means they are risky customers for bank. To compensate risk, banks used to charge higher interest rate than the normal standard rate. Banks funded these loans by selling them to investors on the secondary market. The process of selling them to investors is a legal financial method which is called Collateralized debt obligations (CDO). In 2004-2007, these CDOs were considered as low-risky financial instrument (highly rated).

As these home loan borrowers had high chance to default, many of the them started defaulting on their loans and banks started seizing (foreclose) their property. The real estate bubble burst and a sharp decline in home prices. Many financial institutions globally invested in these funds resulted to a

recession. Banks, investors and re-insurers faced huge financial losses and bankruptcy of many financial and non-financial firms. Even non-financial firms were impacted badly because of either their investment in these funds or impacted because of a very low demand and purchasing activities in the economy. In simple words, people had a very little or no money to spend which leads to many organizations halted their production. It further leads to huge job losses. US Government bailed out many big corporate houses during recession. You may have understood now why credit risk is so important. The whole economy can be in danger if current and future credit losses are not identified or estimated properly.

Basel Regulations

A committee was set up in year 1974 by central bank governors of G10 countries. It is to ensure that banks have minimum enough capital to give back depositors' funds. They meet regularly to discuss banking supervisory matters at the Bank for International Settlements (BIS) in Basel, Switzerland. The committee was expanded in 2009 to 27 jurisdictions, including Brazil, Canada, Germany, Australia, Argentina, China, France, India, Saudi Arabia, the Netherlands, Russia, Hong Kong, Japan, Italy, Korea, Mexico, Singapore, Spain, Luxembourg, Turkey, Switzerland, Sweden, South Africa, the United Kingdom, the United States, Indonesia and Belgium.

Basel I

Basel I accord is the first official pact introduced in year 1988. It focused on credit risk and introduced the idea of the capital adequacy ratio which is also known as Capital to Risk Assets Ratio. It is the ratio of a bank's capital to its risk. Banks needed to maintain ratio of at least 8%. It means capital should be more than 8 percent of the risk-weighted assets.

In Basel I, fixed risk weights were set based on the level of exposure. It was 50% for mortgages and 100% for non-mortgage exposures (like credit card, overdraft, auto loans, personal finance etc). See the example shown below -

Mortgage $5,000

Risk Weight 50%

Risk Weighted Assets $2500 (Mortgage * Risk Weight)

Minimum Capital Required $200 (8% * Risk Weighted Assets)

Basel II

Basel II accord was introduced in June 2004 to eliminate the limitations of Basel I. For example, Basel I focused only on credit risk whereas Basel II focused not only credit risk but also includes operational and market risk. Operational Risk includes fraud and system failures. Market risk includes equity, currency and commodity risk.

In Basel II, there are following three ways to estimate credit risk.

- Standardized Approach
- Foundation Internal Rating Based (IRB) approach
- Advanced Internal Rating Based (IRB) Approach

Standardized Approach

For corporate, the banks relies on ratings from certified credit rating agencies (CRAs) like S&P, Moody etc. to quantify required capital for credit risk. Risk weight is 20% for high rated exposures and goes up to 150 percent for low rated exposures. For retail, risk weight is 35% for mortgage exposures and 75% for non-mortgage exposures (no rating by credit rating agencies required for retail).

Corporate Exposure $5,00,000

Credit Assessment AAA

Risk Weights 20%

Risk Weighted Assets $1,00,000

Minimum Capital Required $8,000

Internal Ratings Based (IRB) Approach

It has four credit risk components:

- Probability of Default (PD)
- Exposure at Default (EAD)
- Loss given Default (LGD)
- Effective Maturity (M)

Probability of Default (PD)

Probability of default means the likelihood that a borrower will default on debt (credit card, mortgage or non-mortgage loan) over a one-year period. In simple words, it returns the expected probability of customers fail to repay the loan. Probability is expressed in the form of percentage, lies between 0% and 100%. Higher the probability, higher the chance of default.

Exposure at Default (EAD)

It means how much we should expect the amount outstanding to be in the case of default. It is the amount that the borrower has to pay the bank at the time of default.

Loss given Default (LGD)

It means how much of the amount outstanding we expect to lose. It is a proportion of the total exposure when borrower defaults. It is calculated by (1 - Recovery Rate).

LGD = (EAD – PV(recovery) – PV(cost)) / EAD

PV (recovery)= Present value of recovery discounted till time of default.

PV (cost) = Present value of cost discounted till time of default.

Expected Loss

The amount a lender might loose by lending to borrower.

Expected Loss is calculated by (PD * LGD * EAD).

Example

Someone takes $100,000 home loan from bank for purchase of flat. At the time of default, loan has an outstanding balance of $70,000. Bank foreclosed flat and sold it for $60,000. EAD is $70,000. LGD is calculated by dividing ($70,000 - $60,000)/$70,000 i.e. 14.3%.

Probability of Default 2%

Exposure at Default $20,000

Loss Given Default 20%

Expected Loss $80

Foundation and Advanced IRB Approach

There are two types of Internal Rating Based (IRB) approaches which are Foundation IRB and Advanced IRB.

Foundation IRB

PD is estimated internally by the bank while LGD and EAD are prescribed by regulator.

Advanced IRB

PD, LGD, and EAD can be estimated internally by the bank itself.

Completed task-

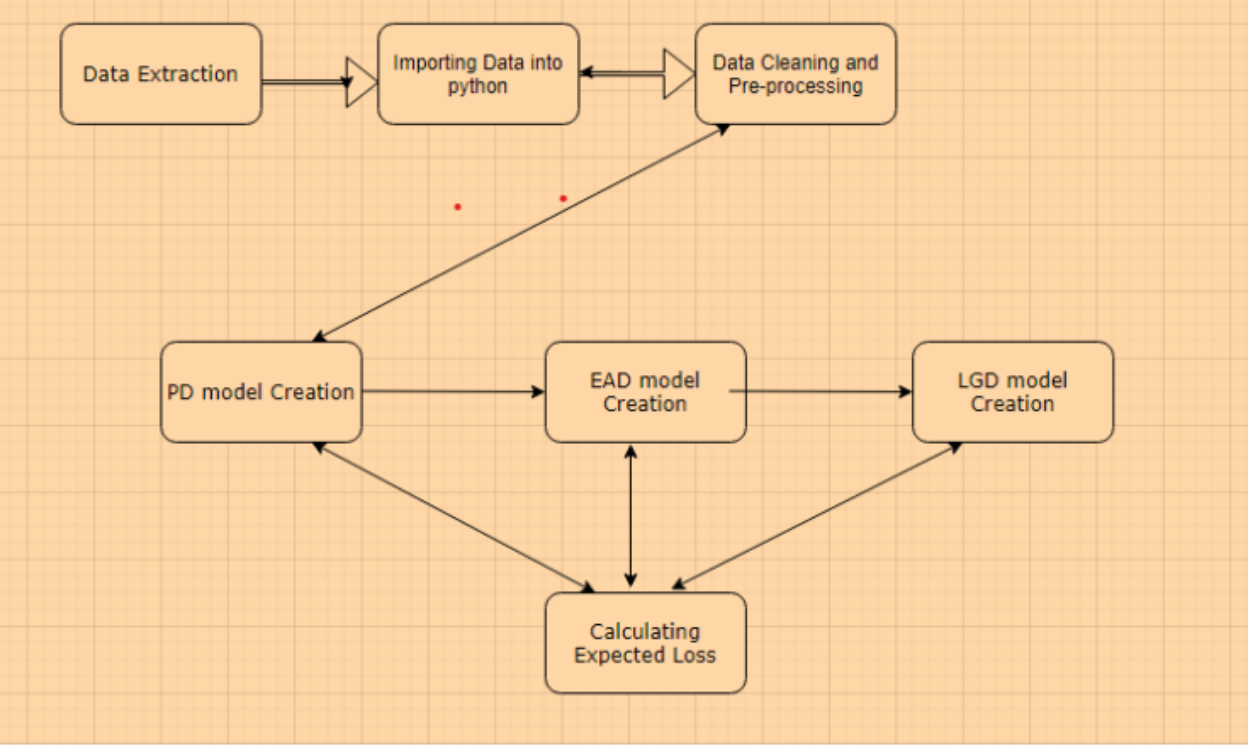
1. Understand dependent and independent variables.
2. Importing data into python.
3. Preprocessing few continuous variables.
4. Preprocessing few discrete variables.
5. Checking for missing values and clean.

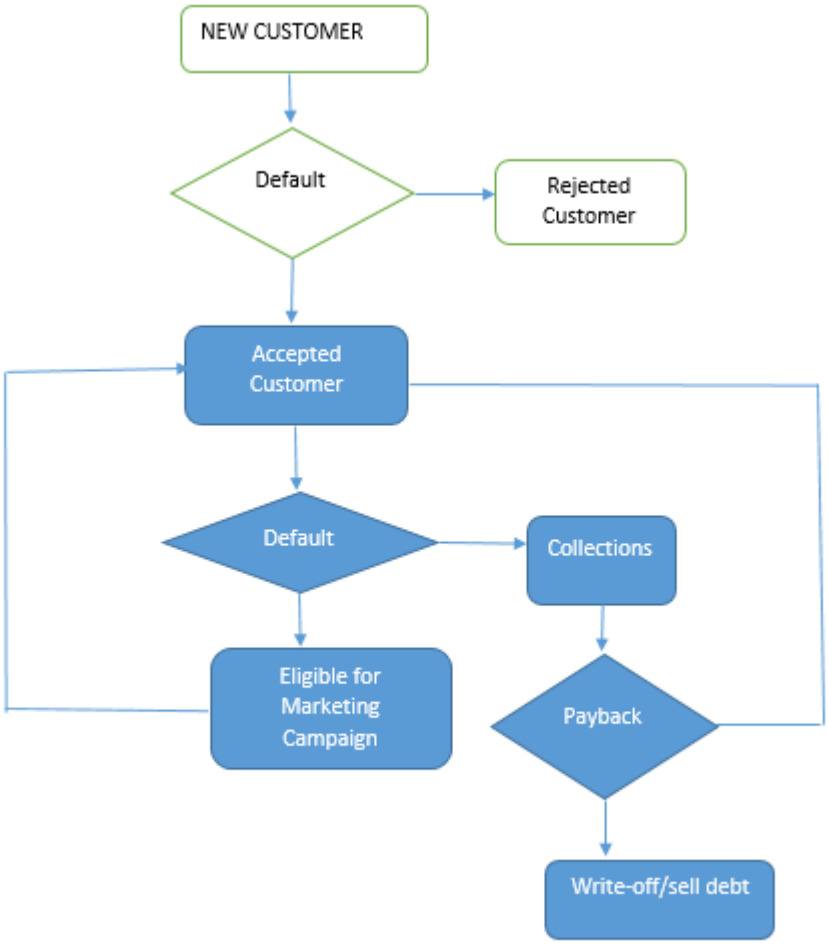
Next steps-

PD model - Logistic regression

6. PD model creation
7. PD model estimation
8. PD model validation
9. Applying PD model for decision making.
10. PD model monitoring.
11. LGD and EAD models.
12. Calculating Expected Loss.

Project flow Design




NEW CUSTOMER
Default
Rejected Customer
Accepted Customer
Default
Collections
Eligible for Marketing Campaign
Payback
Write-off/sell debt