What are Credit Risk Analysis Models?

Financial institutions use credit risk analysis models to determine the probability of default of a potential borrower. The models provide information on the level of a borrower's credit risk at any particular time. If the lender fails to detect the credit risk in advance, it exposes them to the risk of default and loss of funds. Lenders rely on the validation provided by credit risk analysis models to make key lending decisions on whether or not to extend credit to the borrower and the credit to be charged.

With the continuous evolution of technology, banks are continually researching and developing effective ways of modeling credit risk. A growing number of financial institutions are investing in new technologies and human resources to make it possible to create credit risk models using machine learning languages, such as Python and other analytics-friendly languages. It ensures that the models created produce data that are both accurate and scientific. What is Credit Risk?

Credit risk arises when a corporate or individual borrower fails to meet their debt obligations. It is the probability that the lender will not receive the principal and interest payments of a debt required to service the debt extended to a borrower.

On the side of the lender, credit risk will disrupt its cash flows and also increase collection costs, since the lender may be forced to hire a debt collection agency to enforce the collection. The loss may be partial or complete, where the lender incurs a loss of part of the loan or the entire loan extended to the borrower.

The interest rate charged on a loan serves as the lender's reward for accepting to bear credit risk. In an efficient market system, banks charge a high interest rate for high-risk loans as a way of compensating for the high risk of default. For example, a corporate borrower with a steady income and a good credit history can get credit at a lower interest rate than what high-risk borrowers would be charged.

Conversely, when transacting with a corporate borrower with a poor credit history, the lender can decide to charge a high interest rate for the loan or reject the loan application altogether. Lenders can use different methods to assess the level of credit risk of a potential borrower in order to mitigate losses and avoid delayed payments.

AMEX Project Background

A bank can offer two types of cards:

    a.   Charge card: The balance is required to be paid in full each month

    b.   Lending card: Lending cards allow the customer to pay the balance over a period of time subject to interest being charged

An individual can apply for any one of the two types of card on offer. In order to extend the card to the individuals, banks must first underwrite the applicant. Underwriting is the process by which the lender decides whether an applicant is creditworthy and should receive a credit line. Along with the data present in application forms, banks also have access to the consumer bureau. Bureau is an agency that aggregates consumer borrowing and payment information for the purpose of assessing credit-worthiness of an individual and setting a limit on the cumulative credit that can be extended to an individual by lenders.

Problem Statement

We at Amex want you (the students) to develop a credit risk model that predicts the likelihood of a customer to default a payment. The Data present for a customer is at any given point of time. We need you to develop a model to predict the likelihood of the customer to default after 12 months.

Dataset Details:

Training Data_2021.csv – Training data (83K Observations)
Test Data_2021.csv – Dataset that needs to be scored by the students.

The dataset has the customer application and bureau data with the default tagging i.e., if a customer has missed a cumulative of 3 payments across all open trades, his default indicator is 1 else 0. Data consists of independent variables at the time T0 and the actual performance of the individual (Default/ Non Default) after 12 months i.e., at time T12.

Sample Variables:

1. Application variables:
    a. Annual Income
    b. Type of Card applied for – Charge / Lending
    c. Duration of stay at current address
2. Bureau variables:
    a. Performance on external credit card trades
    b. Performance on other trades (Auto / Personal / Education / Mortgage Loan)

Benchmark Results:
Confusion Matrix:

		TRUE Value	
		1	0
Predicted Value	1	a	c
	0	b	d

Accuracy	(a+d)/(a+b+c+d)
Recall	a/(a+b)
Precision	a/(a+c)
F1	2/((1/Recall)+(1/Precision))
Balanced Accuracy	((a/(a+b))+(d/(c+d)))/2