**Project Objectives and Scope**

1. **What is the primary goal of your fraud detection model?**

The objective of this fraud detection model would be to ensure that as many fraudulent transactions are detected and identified correctly and quickly while at the same time ensuring that legitimate transactions are not impacted. Once the model has been trained, it should be scalable and adaptable to new frauds evolving over time.

1. **Why are sensitivity and precision important for this project?**

**Sensitivity**- Sensitivity refers to the proportion of actual fraudulent transactions correctly identified by the model.

**Significance of Sensitivity-**

* A high sensitive model ensures that the model will be in a position to detect large number of fraud transactions.
* Sensitivity = True Positive/(True Positive + False Negative)

**Precision- Precision refers to the proportion of transactions which are flagged as fraud by the model which are actually fraudulent.**

**Significance of Precision-**

* A higher precision means that it will lower the legitimate transactions incorrectly marked as fraudulent.
* With lower legitimate transactions marked as fraud, it will reduce the overhead costs for any financials institution.
* Help customer experience. State Bank of India has a low precision meaning any deduction from a account after a long time to the tune of Rupees 20,000/- is by default flagged as fault. If the callback is not received by the customer for any reason whatsoever, then the transaction is termed as fraud, despite being legitimate, thus leading to poor customer experience.
* Precision = True Positive/(True Positive + False Positive)

Given the interdependence on each other, there needs to be an optimal balance to ensure model functions smoothly.

1. **Does the 'Time' feature help in predicting fraud? How?**

* Yes. Time feature gives us a reference point as to how many transactions are carried out from reference point. For instance, a sudden increase in the number of the transactions from a referral temporal point would strongly indicate a fraud has taken place. Scamsters try to withdraw small amount multiple times to limit the visibility of the fraud.
* Most of the frauds are observed to take place at night where there is less supervision. Eg. At ATM kiosks, the lack of security personal at night encourages scamsters to fraudulently withdraw money rather than during day time.
* A dormant account seeing only deposit, when it suddenly sees a withdrawal where the amount is on the higher side could trigger that the fraud has taken place.

1. **Why should the 'Amount' feature be standardized?**

* The amount feature is varied from small values to large values. This can have an impact on the models which use distance for calculation such as clustering algorithms as they can be heavily influenced by large values thus discarding the other smaller values if not standardized.
* It helps in improving performance of certain models as standardization helps in improving optimization of certain models more efficiently.
* When a feature is standardized, then the sensitivity of the model can be along the predicted or known lines. A small change in the input will not have sweeping changes in the output.
* Having standardized values helps in detection of anomalies by comparing normalized deviations from standardized deviations.

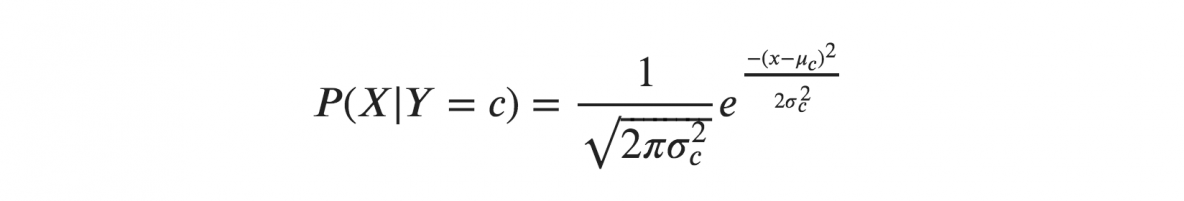
1. **Which features are dropped during preprocessing and why?**

* From the dataset, the feature ‘Class’ is dropped as it is the independent/target variable which determines if a fraud has taken place or not.

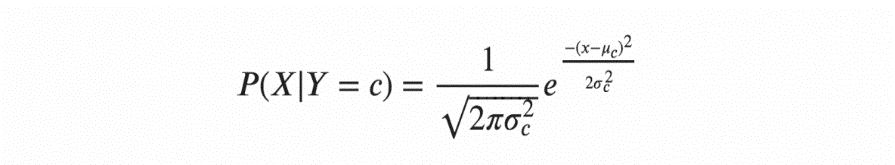
1. **How does Gaussian Naive Bayes handle continuous features?**

* Gaussian naïve Bayes considers features which are on a continuous scale (instead of categories). It assumes that all features exhibit a normal (or Gaussian) distribution and so calculates a normal distribution curve with a mean and standard deviation for each feature in each category (where the x axis represents the measurement of the features, and the y axis represents the likelihood of the category given an x axis measurement).

To illustrate, suppose the training data contain a continuous attribute, x. We first segment the data by the class, and then compute the mean and variance of x in each class. Let μc be the mean of the values in x associated with class c, and let σ2c be the variance of the values in x associated with class c. Then, the probability distribution of some value given a class, p(x=v|c), can be computed by plugging v into the equation for a Normal distribution parameterized by μc and σ2c. That is:



1. **What are the steps in training the Naive Bayes model?**
   * + - 1. Prepare the data- Organizing the dataset into features and corresponding class labels.
         2. Calculate prior probability and evidence- Calculate the probability using the formula- P(X)= Number of instances in class X / Total number of instances​.
         3. Find Likelihood probability with each attribute for each class-
   * For categorical features- P(a|X) = Number of instances where x =a/Total number of instances in class X.
   * For continuous features- Compute mean, variance for each feature in each class, assuming Gaussian distribution.



* + - * 1. See which class has a higher probability, given the input belongs to higher probability class.

1. **How are sensitivity and precision calculated?**

Sensitivity = True Positive/(True Positive + False Negative).

Precision = True Positive/(True Positive + False Positive).

1. **What metrics do you use to evaluate model performance?**

**Following are the metrics used to evaluate model performance-**

* Sensitivity or Recall: Sensitivity = True Positive/(True Positive + False Negative)
* Specificity = True Negative/(True Negative + False Positive)
* ROC(Receiver Operating Characteristic) Curve: It is a graph showing the performance of the model at all threshold levels. It plots True Positive Rate versus False Positive Rate at all levels.

TPR = TP/(TP+FN)

FPR = FP/(FP+TN)

To compute points in the ROC curve, we would need to evaluate logistic regression with different classification thresholds, which is an inefficient process. Hence, we use AUC, i.e. Area under ROC curve.

* AUC Curve- AUC provides an aggregate measure of performance across all possible classification thresholds.
* F1 Score = It is the measure of the harmonic mean of precision and recall.

F1 Score = 2x(Precision x Recall)/(Precision + Recall)

1. **What are the key findings from your model's predictions?**

The key findings are as follows-

* Accuracy- 97.73%. A higher accuracy indicates that the model’s predictions align more closely with the actual labels or ground truth values.
* Recall- 99.97%. High recall means the model has a low rate of false negatives.
* Precision- 97.75. High Precision means model makes fewer false positive predictions.
* Area Under Curve-92.08%- Higher the AUC, better the model’s performance at distinguishing between positive and negative classes.
* F1 Score-98.85%- A higher f1 score generally indicates a well-balanced performance.

1. **How do different threshold values affect model performance?**

**PENDING**

1. **What are the limitations of Naive Bayes for fraud detection?**

* May assign zero probability to unseen events, leading to poor generalization.

Impact on fraud detection- When new patters of fraud emerge, it may lead to unseen feature values.

* Can be influenced by irrelevant attributes.

Impact on fraud detection- It may degrade a model’s performance, as it may not effectively assign importance to different features.

* Assumes that features are independent, which may not always hold in real-world data.

Impact- In real world, features intricately interact with each other. If ignores, it might lead to suboptimal performance.

* Naïve Bayes does not handle co-related features that well.

Impact- In fraud detection, many features are co-related such as transaction amount and frequency. Assuming them independent may lead to suboptimal performance.

1. **What other algorithms could improve performance?**

Combine Naive Bayes with other algorithms like decision trees to leverage the strengths of both models.

1. **What are the ethical implications of deploying your fraud detection model?**

* **Sampling Bias:** Occurs when the data used to train an AI model doesn’t represent the whole target population, leading to skewed or incomplete outcomes.
* **Selection Bias:** This happens when the data for training an AI model is chosen selectively or non-randomly, making the model less effective for the entire target group.
* **Labelling Bias:** Arises during the data labelling process if the labels reflect subjective judgments or biases, influencing the AI model’s learning and decision-making.
* **Cultural Bias:** Emerge if the training data predominantly represents one culture or language, causing the model to perform poorly with data from different cultural backgrounds.
* **Data Collection Bias:** Develops from flawed data collection methods, such as biased survey questions or relying heavily on self-reported data, distorting the AI model’s performance.
* **Algorithmic Bias:** Occurs when the AI model’s algorithm inherently favours or discriminates against certain groups, leading to unfair outcomes.
* **Temporal Bias:** Refers to biases that come into play due to changes over time in societal norms, data distributions, or the relationships between variables in the training data, making the AI model outdated or misaligned with the current context.

1. **What are the steps to implement Naive Bayes in Python?**

* **Data collection**. Aggregating vast amounts of transactional and behavioural data from various sources.
* **Feature engineering**. Identifying and selecting relevant attributes or features of the data that could indicate fraudulent behaviour.
* **Model training**. Using historical data to train the machine learning models to recognize fraud patterns.
* **Anomaly detection**. Applying statistical techniques to identify outliers that diverge from standard patterns.
* **Continuous learning**. Updating the model with new data, ensuring the system evolves with changing fraud tactics.
* **Alerting and reporting**. Flagging suspicious activities and providing detailed reports for further investigation.

1. **How can cross-validation improve your model?**

K-fold cross-validation is a technique to evaluate the performance of a machine learning model which splits the dataset into k-folds or equal-sized subsets.

The dataset is randomly divided into k unique set’s, each referred to as a fold. This model is trained over k-1 folds and tested on remaining fold. The results are averaged to obtain the final performance of the model, it’s generally used over the unseen data to avoid overfitting.

The k value is chosen based on the size of the dataset, for example, if the dataset has 100 instances, then randomly choose the k value as 10 or 20. K-Fold cross validation is majorly used when the dataset is small or imbalanced or when performance of the model needs to be estimated accurately.