**Delivery date prediction**

**Problem statement:**

The logistics team at Olist uses heuristics to provide an estimated delivery date for the orders placed. It is very conservative about the delivery dates. As a result, it is able to deliver the products much in advance. Although this is beneficial for the logistics team’s 'on time delivery' KPI, it is not favourable for the CMO. He found that on average, the estimated time to deliver products that are given to customers is twice that of the actual delivery time. Such a high expected delivery time is driving away Olist's customers. So, the CMO is looking to use ML to get a far more accurate expected delivery date.

**Proposed Solutions:**

ML solutions -

Use one regression model using features like destination location and source location, and historic delivery date, item size, proximity to distribution hubs, seller, etc to predict the estimated time of delivery. Add the delivery time to the order date and calculate the estimated delivery.

Divide the entire delivery process into individual stages of transport.

Estimate time for getting the goods from the vendor.

Estimate the time that the item will stay in the warehouse.

Estimate the time required to send the item from the warehouse to the delivery location.  
  
Build individual models for estimating these times, add the time to get the overall delivery time and add it to the order time to get the estimated delivery date.   
The advantage of building separate models is more fine control over the processes and better prediction.

Non - ML solutions -

A rule-based approach to predicting delivery times. There are set times for each step of the delivery process such as set time for getting goods from the warehouse to the transport hub, from the transport hub to the next transport hub etc.   
  
Each of these times is mapped to the nearest shipping cutoffs wherever appropriate. Fixed additional processing times are added for weekends and holidays. Thus, the rule-based model doesn’t adapt based on recent performance changes and is designed based on heuristics.

Selecting First ML solution - because the data points for individual steps are not present.

**Benefits of Proposed Solution**

**Process improvements -**

Predicting delivery dates more accurately will not impact any process in the e-commerce setup.

Reduction in Inventory costs.

**Monetary benefits -**

The monetary benefits of accurate delivery can be divided into two streams

Reducing customer churn because of accurate delivery date prediction. Hence, increasing the revenue of the company.

Reduction in the number of people who drop out - 10% (Assumed)

Number of customers daily - 272.44   
(total number of orders in 1 year = 99442 / number of days in a year 365)

Average order value - 100 Brazilian real (Assumed)

Revenue increased by = 0.1 \* 272.44 \* 100 = 2724.4 per day.

Inventory management   
  
Predicting the delivery date correctly will reduce the pressure on the supply chain. The through out rate of deliveries of the entire system increases but predicting the delivery date correctly. The warehouses will have to hold lesser inventory.

Reduction of time items stay in warehouse = 40% (assumed)

In the same time frame warehouse will be able to hold more items = 40% (assumed)

Warehouse cost per item delivered will reduce = 25% (assumed)

Initial warehouse cost per item delivered = 5 BR (assumed)

Total savings per day = 272.44 \* 0.75 \* 5 BR = 1021.65 BR

**Summarise the DS approach**

The estimation of accurate delivery dates is a regression problem to be solved. You use various data to estimate the time needed for delivery, then add the time to order date to get the right delivery date.

Delivery date prediction is also a kind of balancing act between competitiveness and accuracy.   
You can always have long delivery dates and always be accurate, but might lose on sale to some competitor who can deliver quickly.

Or you can have extremely short delivery time promises and disturb the customer sentiment. Hence add a buffer.  
[There is a constant trade-off between being accurate and being competitive and, of course, we would aim to optimize both.](https://towardsdatascience.com/delivery-date-estimation-5aff1a0ff8dc)

**Limitation**

RMSE might be high

High complexity model - non linearity , thus non linear model might be needed - is the team capable or not.

Think of edge cases - data for special cases is not available

Define Appropriate Success Metrics

Early delivery rates

Late delivery rates

Churn rate after estimated dates is shown

**Prioritising Use Case**

Refer to the use case prioritisation framework linked [here](https://docs.google.com/spreadsheets/d/1EGoVubzdetsv8YhrKeenSq_uaBSubtsgZm2WydL1U1I/edit#gid=704154787).

**Cite references**

* Similar delivery date estimation application form example on amazon blogs. <https://aws.amazon.com/blogs/industries/how-to-predict-shipments-time-of-delivery-with-cloud-based-machine-learning-models/>
* Paper suggesting various solutions for delivery date production <https://arxiv.org/pdf/2105.00315.pdf>

**Sentiment Analysis**

**Problem statement:**

The Chief Marketing Officer at Olist wanted to understand the experience of the customers based on the reviews received after the delivery of the orders. He also wanted to identify the areas of improvement based on these reviews. He had heard that NLP can be used for sentiment analysis and topic modeling, which will be useful in finding topics in customer reviews. However, he was also cognizant of the fact the customer reviews are in Portuguese, whereas the NLP algorithms are not so sophisticated in Portuguese.

**Proposed Solutions:**

ML Solution -

To analyze the sentiment of customer reviews in Portuguese, we can use machine learning (ML) techniques. This involves creating a dataset of reviews labeled as positive, negative, or neutral. We can then train a machine learning model on this dataset to learn to classify the sentiment of new reviews.

There are different machine learning algorithms that we can use for this task, such as Naive Bayes, Support Vector Machines (SVM), and Recurrent Neural Networks (RNN). We can also use pre-trained models such as BERT, which can be fine-tuned on our dataset to improve accuracy.

The best choice of model will depend on the specific requirements of our task, such as the desired accuracy, speed, and resource constraints.

Non ML solution –

To analyze customer reviews without machine learning, you can follow these steps:

1. Collect and organize the reviews. This involves gathering all of your customer reviews in one place and organizing them by date and order number.
2. Read through the reviews and identify common themes and sentiments. This means looking for patterns in the reviews and identifying the overall feelings that customers are expressing. You can develop a list of topics that are relevant to your business and track the sentiment expressed by customers towards each topic.
3. Categorize the topics and sentiments. Once you have identified the relevant topics and sentiments, you can group them into different areas, such as product quality, delivery, customer service, etc.
4. Provide recommendations for improvement. Based on the topics and sentiments you have identified, you can recommend ways to improve your business to the relevant departments. For example, if you notice that a common complaint is about delayed delivery, you can recommend that your logistics team work on improving their delivery timelines.
5. Monitor progress. Once you have implemented your recommendations, you can monitor their progress and track improvements in customer sentiment over time.

This approach is less computationally intensive than using machine learning, but it may not be as accurate. However, it can be a good option for businesses with small or limited resources.

**Benefits of Proposed Solutions:**

Benefits of using machine learning:

* Efficiency: Machine learning can process large volumes of customer reviews quickly and efficiently, saving time and resources compared to manual analysis.
* Consistency: Machine learning can provide a consistent analysis of customer reviews, avoiding biases that may arise from different team members' interpretations.
* Scalability: Machine learning solutions can be easily scaled to handle an increasing volume of customer reviews as a business grows.
* Continuous improvement: Machine learning models can be fine-tuned over time to improve accuracy and relevance to a business's specific needs.

Benefits of using non-ML solution:

* Language expertise: The non-ML solution requires a team that is proficient in Portuguese and has a good understanding of Olist's business. This can lead to a more accurate and nuanced analysis of customer reviews than an NLP algorithm that is not specifically trained for Portuguese.
* Human intuition: The team can use their intuition and knowledge of Olist's business to identify patterns and trends in the customer reviews that an NLP algorithm might miss.
* Flexibility: The non-ML solution can be tailored to Olist's specific needs and requirements. The team can adjust their analysis and recommendations based on Olist's unique business context.

Monetary benefits of the ML solution:

* Improved accuracy: The ML solution can provide a consistent and accurate analysis of customer reviews, reducing the likelihood of missed insights or errors that could lead to negative customer experiences, which can cost businesses money.
* Increased revenue: By identifying areas of improvement based on customer reviews, businesses can make targeted improvements to their products and services, which can lead to increased customer satisfaction and loyalty, and ultimately, increased revenue.
* Improved efficiency: By automating the analysis of customer reviews, the ML solution can save businesses time and resources compared to a manual analysis. This can lead to cost savings.

Monetary benefits of the non-ML solution:

* Reduced customer churn: Addressing customer concerns and improving satisfaction can also reduce customer churn, which can be a significant cost savings for businesses.
* Increased customer satisfaction: By analyzing customer reviews and providing recommendations for improvement, businesses can address customer concerns and improve their satisfaction with the business's products and services. This can lead to increased customer loyalty and repeat business, which can boost revenue.
* Increased revenue: By addressing customer concerns and improving satisfaction, businesses may be able to increase revenue by retaining customers and generating positive word-of-mouth referrals.

**Limitations of the ML solution:**

* NLP algorithms are not as sophisticated in Portuguese as they are in English. This could lead to lower accuracy when analyzing Portuguese customer reviews.
* The model is only as good as the data it is trained on. If the dataset used to train the model is biased or limited in scope, the model may not accurately reflect the sentiment of all customers.

**Limitations of the non-ML solution:**

* It is time-consuming and expensive to manually analyze customer reviews, especially for companies with a large customer base.
* Human analysts may have their own biases, which could influence the results.

**Define appropriate success metrics:**

* **Sentiment analysis accuracy:** This metric measures how well the sentiment analysis model can correctly classify the sentiment of customer reviews as positive, negative, or neutral.
* **Topic modeling accuracy:** This metric measures how well the topic modeling model can identify the main topics discussed in customer reviews.
* **Net Promoter Score (NPS):** This metric measures the willingness of customers to recommend a business to others. It is calculated by subtracting the percentage of detractors (customers who rate the business 0-6 on a scale of 0-10) from the percentage of promoters (customers who rate the business 9-10 on the same scale).
* **Improvement rate:** This metric measures the change in customer sentiment over time. It can be calculated by subtracting the average sentiment score from the previous period from the average sentiment score from the current period.

**Summarise the Solution**

* The non-ML solution is more accurate and nuanced, but it is also more time-consuming and expensive.
* The ML solution is less accurate and nuanced, but it is more efficient, scalable, and affordable.
* The best solution for Olist will depend on their specific needs and resources.

**Prioritising Use Case**

* We could refer the attached excel having the detailed framework..
* Plus. refer to the use case prioritisation framework linked [here](https://docs.google.com/spreadsheets/d/1EGoVubzdetsv8YhrKeenSq_uaBSubtsgZm2WydL1U1I/edit#gid=704154787).

**References**

* <https://towardsdatascience.com/delivery-date-estimation-5aff1a0ff8dc>
* Paper suggesting various solutions for delivery date production <https://arxiv.org/pdf/2105.00315.pdf>

**Customer Churn**

**Problem statement:**

Customer churn is when customers stop using a company's products or services. It is a critical metric for e-commerce companies because it can have a significant impact on their profits.

Olist wants to develop customer churn models to identify customers who are at risk of churning so that they can develop targeted retention strategies to keep them.

Customer churn models work by identifying patterns in the behaviour of customers who have churned in the past. These patterns can then be used to identify customers who are at risk of churning in the future.

Retention strategies can vary depending on the factors that are driving customer churn. For example, if a customer is churning because they are unhappy with the quality of a product, the company may offer them a discount or a refund.

**Proposed Solutions:**

**ML Solution for Customer Churn Prediction-**

Machine learning (ML) can be used to develop customer churn prediction models that can identify customers who are at risk of churning. This is done by using historical customer data to train a predictive model. The model can then be used to score new customers and identify those who are most likely to churn.

One popular ML technique for customer churn prediction is logistic regression. Logistic regression is a binary classification algorithm that predicts whether a customer is likely to churn or not, based on a set of features such as purchase history, browsing behavior, and customer demographics.

Other ML algorithms that can be used for customer churn prediction include decision trees, random forests, and support vector machines. These algorithms can also provide insights into the most important factors driving customer churn.

Non-ML Solution for Customer Churn Reduction-

There are a number of non-ML solutions that businesses can use to reduce customer churn. These include:

* Conducting customer surveys to gather feedback and identify the reasons for customer dissatisfaction and churn.
* Dividing customers into different segments based on their demographic and behavioral data to identify customers who are at risk of churning and develop targeted retention strategies for each segment.
* Calculating the CLV of each customer to understand their value to the e-commerce platform and prioritize retention efforts for high-value customers.
* Improving customer service by providing personalized assistance, timely responses to queries and complaints, and easy access to customer support channels.
* Developing loyalty programs that reward customers for their continued engagement with the e-commerce platform.

**Monetary benefits:**

* Increased revenue: By identifying customers who are at risk of churning and implementing strategies to keep them, e-commerce businesses can reduce the number of customers they lose. This results in more sales and more money.
* Reduced customer acquisition cost: It is generally more expensive to acquire new customers than to retain existing ones. By retaining existing customers, e-commerce businesses can reduce the amount of money they spend on customer acquisition, which saves them money.
* Improved customer satisfaction: By understanding and addressing the reasons why customers churn, e-commerce businesses can make customers happier and more loyal. This leads to more sales and more money over time.

**Limitation:**

ML solutions:

* Limited interpretability: ML models can be complex and difficult to understand, which can make it difficult for e-commerce companies to understand how the model is making predictions and recommendations. This can make it difficult to trust the model and to use its insights to make informed decisions.
* Data bias: ML models can be biased, which can lead to inaccurate predictions and recommendations. For example, if the model is trained on a dataset that is biased towards a certain type of customer, it may be more likely to predict that customers who do not fit that profile are at risk of churning.
* Limited domain knowledge: ML models may not have access to the domain knowledge and expertise that is needed to provide accurate and actionable insights. For example, the model may not understand the specific factors that drive customer churn in the e-commerce industry.

Non-ML solutions:

* Inability to handle real-time data: Non-ML solutions may not be able to process data in real time, which can make it difficult for e-commerce companies to respond quickly to changes in customer behavior and market conditions.
* Limited scalability: Non-ML solutions may not be able to analyze large and complex datasets, which can make it difficult to identify the factors that drive customer churn accurately.
* Lack of predictive capabilities: Non-ML solutions may not be able to predict which customers are at risk of churning, making it difficult for e-commerce companies to take proactive steps to retain them.

**Appropriate success metrics for customer churn prediction:**

* Churn rate: The percentage of customers who stop using a product or service during a given period of time.
* Customer retention rate: The percentage of customers who continue to use a product or service during a given period of time.
* Customer lifetime value (CLV): The total revenue that a customer is expected to generate over the course of their relationship with a company.
* Customer satisfaction score (CSAT): A measure of how satisfied customers are with a product or service.

**Prioritising Use Case:**

* Refer to the use case prioritisation framework linked [here](https://docs.google.com/spreadsheets/d/1EGoVubzdetsv8YhrKeenSq_uaBSubtsgZm2WydL1U1I/edit#gid=704154787).

**Summary of the proposed solutions:**

E-commerce companies can use the proposed solutions to improve customer retention, increase customer satisfaction, reduce costs, gain data-driven insights, and implement scalable solutions to address customer churn.

Both non-ML and ML-based solutions can be used to identify customers who are at risk of churning. Non-ML solutions involve analysing historical data and conducting surveys to identify trends and patterns that can indicate which customers are likely to churn. ML-based solutions use predictive models to identify at-risk customers based on a variety of features, such as purchase history, customer demographics, and customer behaviour.

**References:**

* <https://towardsdatascience.com/delivery-date-estimation-5aff1a0ff8dc>
* Paper suggesting various solutions for delivery date production <https://arxiv.org/pdf/2105.00315.pdf>

**Customer Acquisition Cost**

**Problem statement:**

Olist's marketing team is spending a lot of money to attract new customers, and the CFO wants to make sure that this money is being spent wisely. The CFO wants to start a new process to measure how effective the marketing campaigns are by comparing them to how much money Olist makes from each customer over time.

**This problem statement can be broken down into two key parts:**

1. Olist's marketing team is spending too much money on customer acquisition.
2. The CFO wants to measure the effectiveness of the acquisition campaigns by comparing them to the lifetime value of customers.

**Solutions to these two problems can help Olist reduce its customer acquisition cost and improve its profitability.**

**Solution Summary:**

The proposed solution is to optimize the customer acquisition cost by measuring the effectiveness of acquisition campaigns against the lifetime value of customers. Two possible solutions are suggested: a Machine Learning (ML) solution and a Non-ML solution.

ML Solution:  
ML can be used to optimize customer acquisition cost in the e-commerce industry. One ML solution is to use predictive modelling to identify the most profitable acquisition channels and campaigns. This approach involves analyzing historical data to identify patterns and insights into customer behaviour and preferences. Machine learning algorithms such as decision trees, random forests, and gradient boosting can be used to develop a predictive model that can identify the most effective acquisition channels and campaigns.

Another ML solution is to use customer segmentation to target specific customer groups with tailored acquisition campaigns. By segmenting customers based on demographics, purchase history, and other relevant factors, e-commerce companies can develop acquisition campaigns that are more likely to resonate with each group.

Non-ML Solution:

One non-ML solution is to conduct a customer acquisition cost analysis. This analysis involves tracking the cost of each acquisition campaign and the total revenue generated by the acquired customers over their lifetime. The analysis can help identify the most cost-effective acquisition channels and campaigns.

Another non-ML solution is to implement a customer referral program. This program incentivizes existing customers to refer new customers, which can be a cost-effective way to acquire new customers. By providing incentives to existing customers, e-commerce companies can reduce their customer acquisition cost.

**Benefits of proposed solutions:**

**Increased profits:** Machine learning (ML) and non-ML solutions can help e-commerce companies reduce the cost of acquiring new customers, which leads to higher profits.

**Improved customer experience:** By tailoring acquisition campaigns to specific customer groups, companies can improve the customer experience and increase their chances of acquiring new customers.

**Scalability:** Both ML and non-ML solutions can be applied at scale, making them ideal for e-commerce companies where we have to acquire a large number of customers.

**Cost-effectiveness:** Non-ML solutions such as referral programs can be a cost-effective way to acquire new customers, while ML-based solutions can help companies identify the most cost-effective acquisition channels and campaigns.

**Monetary benefits:**

* **Increase in revenue:** E-commerce companies can make more money by using ML and non-ML solutions to acquire new customers more efficiently and effectively.
* **Reduce customer acquisition cost:** ML and non-ML solutions can help e-commerce companies save money on customer acquisition by identifying the most effective channels and campaigns, and by tailoring campaigns to specific customer groups.
* **Improve ROI:** ML and non-ML solutions can help e-commerce companies improve their return on investment by reducing the cost of customer acquisition and increasing the revenue generated by acquired customers.

**Limitation:**

ML Solution –

* A ML model trained on data from a predominantly white and affluent population may be biased against customers from other demographic groups. This could lead to the model recommending different products or services to customers based on their race or ethnicity, even if they have similar needs and preferences.
* A ML model used to predict customer churn may be difficult to interpret, making it hard to understand why the model is predicting that certain customers are likely to churn. This can make it difficult to develop effective strategies to retain these customers.
* A ML model used to identify the most effective customer acquisition channels may not have access to data on all of the relevant channels. This could lead the model to recommend channels that are not as effective as other channels that the model is not aware of.

Non-ML Solution –

* A non-ML solution that relies on human analysis of customer data may not be able to identify subtle patterns and trends that could be indicative of effective customer acquisition channels and campaigns.
* A non-ML solution may not be able to process large and complex data sets in a timely manner, which can limit its ability to provide actionable insights.
* A non-ML solution may not be able to provide real-time insights into the effectiveness of customer acquisition campaigns, which can make it difficult for e-commerce companies to quickly adjust their campaigns in response to changes in customer behaviour and market conditions.

**Define Appropriate Success Metrics:**

* **Customer acquisition rate (CAR)**: The percentage of new customers acquired in a given period of time.
* **Customer acquisition cost (CAC)**: The average cost of acquiring a new customer.
* **Customer lifetime value (CLV)**: The total revenue that a customer is expected to generate over the course of their relationship with the company.
* **Return on advertising spend (ROAS)**: The ratio of revenue generated from advertising to the amount of money spent on advertising.
* **Conversion rate**: The percentage of visitors to a website or landing page who take a desired action, such as signing up for a newsletter or making a purchase.

**Prioritising Use Case:**

* Refer to the use case prioritisation framework linked [here](https://docs.google.com/spreadsheets/d/1EGoVubzdetsv8YhrKeenSq_uaBSubtsgZm2WydL1U1I/edit#gid=704154787).

**Summary :**

**ML and non-ML solutions can help e-commerce companies acquire new customers more efficiently and effectively, which can lead to increased profits, better decision making, an enhanced customer experience, and cost savings.**

**References:**

* <https://towardsdatascience.com/delivery-date-estimation-5aff1a0ff8dc>
* <https://corporatefinanceinstitute.com/resources/accounting/customer-acquisition-cost-cac/>
* Paper suggesting various solutions for delivery date production <https://arxiv.org/pdf/2105.00315.pdf>

**Fraud Detection**

**Problem statement:**

**Fraud detection is a major challenge in the e-commerce industry, as it can lead to significant financial losses. Fraud can occur in a variety of forms, including merchant identity fraud, advanced fee and wire transfer scams, and chargeback fraud.**

**The CFO wants to use analytics to identify fraudulent transactions and protect the organization from these types of fraud.**

**Solution Summary:**

ML Solution –

**ML solutions can use both supervised and unsupervised learning algorithms.**

**1. Supervised learning** involves training a model on a dataset of labeled data, where each data point is labeled as either fraudulent or non-fraudulent. The model then learns to identify fraudulent transactions in new data by looking for patterns that are similar to the patterns in the training data.

**2. Unsupervised learning** involves training a model on a dataset of unlabeled data, where the data points are not labeled as either fraudulent or non-fraudulent. The model then learns to identify fraudulent transactions by looking for anomalies in the data.

**3. A hybrid approach** can also be used, where both supervised and unsupervised learning algorithms are used together to identify fraudulent transactions. This can help to improve the accuracy of the fraud detection system.

Non ML Solution –

**Non-ML solutions for fraud detection can be used to identify fraudulent transactions by using pre-defined rules, data mining, and expert knowledge.**

**1. Rule-based approach:** A rule-based approach involves setting certain rules to identify fraudulent transactions. For example, transactions above a certain threshold value or transactions from unusual locations can be flagged for further investigation.

**2. Expert knowledge:** Expert knowledge of the industry and the specific fraud types can be used to identify potential fraudulent transactions. This may involve manual analysis of transactions and other related data.

**3.Data mining:** Data mining involves analyzing large datasets to identify patterns and anomalies in transactions. This can help to identify fraudulent transactions that may not be detected by rule-based approaches.

**Benefits of Proposed Solutions:**

ML Solution -

* **Accuracy:** ML models can identify fraudulent transactions more accurately than non-ML solutions because they can analyze large volumes of data and identify complex fraud patterns.
* **Improved ROI:** ML models can improve the ROI of the organization by reducing the losses due to fraudulent transactions and increasing efficiency in fraud detection.
* **Scalability:** ML models can scale to handle large volumes of data and can adapt to changing fraud patterns over time.

Non-ML Solution –

* **Cost-effective:** Non-ML solutions such as rule-based approaches and expert knowledge are relatively inexpensive and can be implemented quickly without the need for extensive data or computing resources.
* **Transparent:** Non-ML solutions are generally more transparent and easier to understand than ML models, which can help build trust with stakeholders.

**Monetary Benefits :**

* **Reduce fraud losses:** ML and non-ML solutions can help organizations identify and prevent fraudulent transactions, which can lead to a reduction in fraud losses.
* **Improve efficiency:** ML and non-ML solutions can automate the fraud detection process, which can free up staff time and resources for other tasks.
* **Save costs:** ML and non-ML solutions can help organizations reduce the costs associated with fraud, such as the cost of investigating and responding to fraudulent transactions.

**Limitations :**

**ML solutions:**

* **Can be complex and difficult to interpret:** ML models can be complex and difficult to interpret, which can limit their transparency and trustworthiness.
* **Accuracy may decrease over time:** As fraud patterns change, the accuracy of ML models can decrease over time, requiring frequent retraining.

Non-ML Solutions:

* **Limited in their ability to identify complex fraud patterns:** Non-ML solutions rely on pre-defined rules or manual analysis, which can be ineffective at detecting complex fraud patterns that evolve over time.
* **May not be able to handle large datasets efficiently:** Non-ML solutions may not be able to scale to handle large volumes of transaction data efficiently, which can lead to missed detections.

**Define Appropriate Success Metrics:**

* **Return on Investment (ROI)**: The financial benefit of using a fraud detection solution, calculated by dividing the financial benefits by the cost of the solution.
* **Accuracy:** The percentage of fraudulent transactions that are correctly identified by the fraud detection solution.
* **Precision:** The percentage of transactions that are flagged as fraudulent by the solution that are actually fraudulent.
* **Recall:** The percentage of all fraudulent transactions that are identified by the solution.

**Prioritising Use Case:**

* Refer to the use case prioritisation framework linked [here](https://docs.google.com/spreadsheets/d/1EGoVubzdetsv8YhrKeenSq_uaBSubtsgZm2WydL1U1I/edit#gid=704154787).

**Summary:**

Machine learning (ML) solutions and non-ML solutions can both be used to detect fraud. ML solutions use machine learning algorithms to identify fraudulent transactions, while non-ML solutions use rule-based approaches and expert knowledge. Both types of solutions can provide benefits such as cost savings, increased efficiency, and improved accuracy in detecting fraud.

**References:**

* <https://towardsdatascience.com/delivery-date-estimation-5aff1a0ff8dc>
* <https://corporatefinanceinstitute.com/resources/accounting/customer-acquisition-cost-cac/>
* Paper suggesting various solutions for delivery date production <https://arxiv.org/pdf/2105.00315.pdf>

**Price Optimisation**

**Problem statement:**

**Price optimization is a critical process for e-commerce organizations, as it has a direct impact on revenue, sales, profit, and demand.**

**Price optimization algorithms use a variety of factors, such as location, customer behavior, and competitor pricing, to predict customer demand and set prices that maximize sales and revenue.**

**OLIST's sales team wants to build a price optimization algorithm to maximize sales and revenue.**

**Solution Summary:**

ML Solution –

**Machine learning (ML) algorithms can be used to analyze customer behavior and predict the optimal price for a product or service.**

**ML algorithms can be trained on a variety of factors, such as location, customer attitudes, competitor pricing, and historical sales data, to understand how customers respond to different prices.**

**ML-based price optimization solutions can set prices in real time and improve over time as new data is added.**

Non-ML Solution -

**Non-ML price optimization involves manually setting prices based on market research, competitor analysis, and historical sales data.**

**This approach can be effective if the pricing strategy is consistent and can be easily identified through manual analysis.**

**However, non-ML price optimization may not be as effective in adapting to new or changing market conditions.**

**Benefits :**

* **Increased revenue:** By setting optimal prices, organizations can maximize sales and revenue.
* **Improved customer satisfaction:** By offering fair and competitive prices, organizations can improve customer satisfaction and loyalty.
* **Competitive advantage:** Organizations can gain a competitive advantage by optimizing prices based on competitor pricing and customer preferences.

**Limitations:**

* **Data availability:** If you do not have access to all of the data that you need to implement a price optimization solution, you may want to start with a simple non-ML solution. Once you have collected more data, you can upgrade to an ML solution.
* **Implementation costs:** There are a number of ways to reduce the cost of implementing ML algorithms for price optimization. For example, you can use cloud-based ML services or open-source ML software. You can also hire a consultant to help you implement a price optimization solution.
* **Customer acceptance:** When implementing price changes, it is important to communicate with your customers and explain the reasons for the changes. You may also want to offer customers incentives to accept the new prices.

**Define Appropriate Success Metrics:**

* **Revenue:** Revenue is the most important metric for price optimization, as the goal is to maximize sales and revenue for the organization.
* **Market share:** Market share can be used as a metric to evaluate the effectiveness of the organization's pricing strategy compared to competitor
* **Customer satisfaction:** Customer satisfaction is also an important metric, as the goal is to provide optimal prices to customers and improve customer loyalty.

**Prioritising Use Case:**

* Refer to the use case prioritisation framework linked [here](https://docs.google.com/spreadsheets/d/1EGoVubzdetsv8YhrKeenSq_uaBSubtsgZm2WydL1U1I/edit#gid=704154787).

**Summary:**

Price optimization is a powerful tool that can help e-commerce businesses improve their performance in a number of ways. The best way to choose a price optimization solution is to consider the specific needs and resources of the business, as well as the specific goals that the business wants to achieve.

**References:**

* <https://towardsdatascience.com/delivery-date-estimation-5aff1a0ff8dc>
* <https://corporatefinanceinstitute.com/resources/accounting/customer-acquisition-cost-cac/>
* Paper suggesting various solutions for delivery date production <https://arxiv.org/pdf/2105.00315.pdf>