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 Chief Marketing Officer (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 long 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:

Since sufficient data is available on actual delivery dates, this is a classic case of regression to estimate the delivery duration. However, it is important to consider the impact of three factors:

1. Weekends

Contrary to popular belief, the weekend is the quietest time for e-commerce and sales begin to dip from Friday. Also, weekends can cause delays in delivery due to reduction in manpower as people are on leave. Thus, it is essential to understand the impact of weekends on delivery dates of orders.

2. Holiday season

Holiday season increases sales, creating increased traffic for a lot of delivery systems. Thus, one must consider this factor before designing a delivery date prediction system. A lot of companies have a separate ML model for prediction during holiday season and one for non holiday season. Apart from that there is a reduction in manpower during the holiday season.

E-commerce sales made up 20.9 percent of total retail sales in the holiday season of 2021, which is slightly higher from 20.6 percent and 14.6 percent in the holiday seasons of 2020 and 2019, respectively.

ML solutions

1. Divide the entire delivery process into individual stages of transport.
    i. Identify the order date and day and find the duration of the year.
    ii. Estimate the time to get the goods from the vendor.
        1. Split the vendors into three categories based on their consistency of delivering to the platform.
        2. Identify the warehouse where it will be delivered.
    iii. Estimate the time required to send the item from the warehouse to the final delivery location.
    iv. Estimate the time required to deliver to the customer.
    v. Merge the three processes.

3. Non - ML solutions
    i. Based on EDA alone, it is possible to manually estimate the delivery dates. You can compute average delay in delivery dates and add that to the actual delivery dates to improve your delivery dates system.

Benefits of Proposed Solution:

The solution can improve warehouse utilisation, decrease customer churn, improve profits and brand image. Assuming a reduction in 10% of transit time, a simple formula to gauge the scope for transport cost reduction is as follows:

10% Reduction in Transit Time = (100% - 10%) * (Cost per trip * Number of Trips)

Similar computations can be achieved for all other parameters.

Summarise the DS Approach:

Delivery time promising is critical to managing customer expectations and improving customer satisfaction. Simply over-promising or under-promising is undesirable due to its negative impacts on short-term/long-term sales. A regression technique is designed to identify the actual delivery date based on historical data. The system needs to be optimised on a regular basis as the company will tend to streamline its delivery process. Thus, one must take this aspect into account and constantly update the machine learning (ML) model for delivery date prediction.

Prioritising Use Case:

Amongst the six use cases given to the Data Science team, this use case was ranked first based on the associated challenges and commercial value.

Success Metrics and Key Performance Indicators

Starting with a success metric of improving delivery date prediction by 30%, the success metrics and key performance Indicators can be measured as follows

1. Percentage of orders which are delivered before Actual Prediction Date using existing technique (A)
2. Percentage of orders which are delivered before Actual prediction date using AI prediction (B)
3. Percentage Improvement (PI) in delivery prediction is
$$\left[\frac{(A-B)}{Average(A,B)}\right]x100$$
4. If PI>30%, the project is justified.

Cite References

http://www.parcelperform.com/insights/estimated-delivery-date

https://www.aftership.com/features/delivery-date-prediction

https://aws.amazon.com/blogs/industries/how-to-predict-shipments-time-of-delivery-with-cloud-based-machine-learning-models/

https://docs.microsoft.com/en-us/dynamics365/business-central/sales-how-to-calculate-order-promising-dates

Yanchuk, V. M., Tkachuk, A. G., Antoniuk, D. S., Vakaliuk, T. A., & Humeniuk, A. A. (2020). Mathematical simulation of package delivery optimization using a combination of carriers	The authors use the Monte Carlo Simulation by considering urban zone, rural zone and distant zone orders with different carriers, including vendor and supplier delivery vehicles. Investigations show that delivery and cost can be optimised using different routes and multiple delivery partners.
Cardenas, I. D., Dewulf, W., Beckers, J., Smet, C., & Vanelslander, T. (2017). The e-commerce parcel delivery market and the implications of home B2C deliveries vs pick-up points. The e-commerce parcel delivery market and the implications of home B2C deliveries vs pick-up points, 235-256.	Delivery date prediction extensively relies on the last mile, which is expensive and highly complicated in the e-commerce delivery space. This article describes the challenges and techniques to overcome the same.

Yu, Y., Wang, X., Zhong, R. Y., & Huang, G. Q. (2016). E-commerce logistics in supply chain management: Practice perspective. *Procedia Cirp*, *52*, 179-185.	The authors have extensively studied the challenges in supply chain logistics for e-commerce companies in different geographical locations. The practical outcomes are discussed in detail and can help understand transfer metrics when designing the AI solution.
Zhang, Y., Liu, Y., Li, G., Ding, Y., Chen, N., Zhang, H., ... & Zhang, D. (2019). Route prediction for instant delivery. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *3*(3), 1-25.	Literature has consistently shown that identifying the optimal flow path of goods is important to achieve shorter delivery dates. This work focusses on creating multiple features to model the decision-making.

Sentiment Analysis

Problem Statement:

The Chief Marketing Officer (CMO) 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 natural language processing (NLP) can be used for sentiment analysis and topic modelling, 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:

1. ML Solution

Sentiment Analysis can be carried out using either the supervised or the unsupervised approach. In either case, it is critical to understand and decompose the text into morphemes and associate each morpheme with a Lemma word, Parts of Speech and affix. Though solutions for the morphological techniques are mature and used in English, a couple of tools are currently available for the Portuguese language. NLTK and Polyglot provide various tools and functions to support the Portuguese language. The process to be followed to create machine learning (ML) models is as follows:

a. Acquire the sentiment data and assign Positive Sentiment or Negative Sentiment based on the star rating. All ratings greater than 3 stars are stated to be positive and less than 3 stars as negative.
b. Remove all URLs and special symbols, including emojis.
c. Understand slangs and abbreviations. For example,

        i. kkk in Portuguese stands for laughing and
        ii. blz stands for cool, all-good.
    d. Create a data dictionary for slang that can strengthen the prediction.
    e. Create sentence level tokens and apply Part of Speech (POS) Tagging.
    f. Select attributes based on adjectives and adverbs.
    g. Vectorise the word level tokens.
    h. Train with a supervised algorithm like SVM or ANN if classification is preferred and map to the sentiment value. Use a clustering technique to find clusters that can be tagged to understand the sentiment better.
    i. Test the training set and achieve benchmarks of at least 0.85 in the F1 Score

2. Non-ML solution

Since star ratings for each of the sentiment is already available, the sales/marketing team can measure the sentiment and manually check actions to be taken for one and two star rated reviews, which can create customer churn. They can also use lower ratings to identify customer pain-points and improve existing processes.

Benefits of Proposed Solution:

The proposed solution can help the organisation in

1. Building brand value
Tracking your audience's sentiments lets you comprehend how and what their feelings are behind your social media content and posts. Knowing the reason for your posts' reaction can be a source of valuable context for your brand on how to proceed and respond with your next business plan.

2. Reducing customer churn
You can calculate customer churn based on revenue. Businesses that take this approach typically use monthly recurring revenue (MRR) as a baseline figure. MRR is the total predicted revenue for an organization in a particular month. . Calculation of customer churn is slightly more complicated when using MRR.

For example, suppose a company had $400,000 of MRR at the beginning of the month and $350,000 at the end. Now, let's say that the company brought in $55,000 from existing customers who purchased additional material that same month. The churn can be calculated as follows:

[(400,000 - 350,000) - 55000] / 400,000 = -0.0125 x 100 = -1.25%

The churn rate is negative, which implies that the company actually ended up making money despite the $55,000 loss in MRR. This is known as negative churn.

3. Improving services and processes
Creating a database of negative reviews will help identify customer pain points better, which can then be used to improve the existing processes.

Summarise the Solution:

The benefits of the proposed ML solutions are as follows:

ii. Avoid English translation, which can create errors in context
iii. Use existing open-source tools, thus reducing costs
iv. Reduce the number of steps and have a simple pipeline

v. Create a database of negative sentiments can help identify various pain points faced by the customers and help improve the existing services

Prioritising Use Case:

Amongst the six use cases given to the Data Science team, this use case was ranked 4th based on the challenges and commercial value. The sales team can use the non-ML model to address very poor sentiments.

Success Metrics and Key Performance Indicators

Addressing negative reviews can help in reducing customer churn. The churn rate formula is [(Lost Customers ÷ Total Customers at the Start of Time Period) x 100] is not an accurate metric for e-commerce websites as it is difficult to identify customers' buying patterns since no subscription fees are applicable on a monthly or quarterly basis. Monthly Recurring Revenue (after adjusting against revenue growth along with percentage increase in new customer acquisition can be a good indicator to identify Sentiment. The third metric is the reduction in negative sentiments after action taken by the customer success team. The KPI can be measured as

$$\frac{[w1(MRR)+w2*(New\ customers\ acquired\ by\ referral)]}{Ratio\ of\ negative\ sentiment\ over\ positive\ Sentiment} Where\ w1\ +\ w2\ =\ 1$$

Cite References:

https://monkeylearn.com/application/review-analysis/

https://www.lexalytics.com/semantria

https://www.brandwatch.com/

https://www.meaningcloud.com/products/sentiment-analysis

https://www.rosette.com/capability/sentiment-analyzer/

Souza, M., & Vieira, R. (2012, April). Sentiment analysis on twitter data for Portuguese language. In *International Conference on Computational Processing of the Portuguese Language* (pp. 241-247). Springer, Berlin, Heidelberg.	The authors investigate the OpLexicon library for the Portuguese language sentiment analysis. Investigations were carried out with good precision.
Balage Filho, P., Pardo, T. A. S., & Aluísio, S. (2013). An evaluation of the Brazilian Portuguese LIWC dictionary for sentiment analysis. In *Proceedings of the 9th Brazilian Symposium in Information and Human Language Technology*.	The authors investigate Linguistic Inquiry and Word Count (LIWC), which has been made available for the Portuguese language. As a result, the authors were able to achieve 88% accuracy on positive sentiments.
Yang, L., Li, Y., Wang, J., & Sherratt, R. S. (2020). Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. *IEEE access*, *8*, 23522-23530.	The authors extensively investigate product reviews on e-commerce for sentiment reviews using deep learning.

Customer Churn

Problem Statement:

Customer churn is a critical metric for the CMO of any e-commerce company. OLIST wants to develop customer churn models to identify 'at-risk' customers so that an appropriate retention strategy can be built. This will provide insights into the factors that drive customer churn, thus reinforcing the retention efforts of the company. Maintaining a large customer base is an important way of increasing revenue. However, as it happens in many businesses, customers tend to move between e-commerce companies. To prevent customers from constantly migrating, the company has built a churn model. The model is used to identify the customers who are likely to migrate. Now, the company wants to come up with a strategy to prevent churn.

Proposed Solutions:

1. ML solution

Sufficient data is not available for customer churn as specific data is not available on account closure, which indicates a churn. A model can be built on the hypothesis that dormant accounts are churn but the data may not be sufficient for an individual. Another way to hypothesise a churn is by looking at the star ratings, which is not sufficient as they have to be mapped to an individual and the number of negative ratings is less. A combination of star rating and identifying mean time between purchases can be used to build an ML model and the data can be used to take feedback from the customers to build a more rugged model.

a. Bucketise the data according to frequency of purchase.
    i. Singleton
    ii. Less than one month
    iii. One month to three months
    iv. Three months to six months
    v. Greater than 6 Months
b. Bucketise the average purchase value.
    i. Very high value
    ii. High value
    iii. Average value
    iv. Low value
    v. Very low value
    vi. Combination of all
c. Identify the number of negative sentiments.
    i. Only positive
    ii. Only negative
    iii. Mostly positive
    iv. Mostly negative
d. Cluster the data by starting off with four clusters and increasing to find out if any unique patterns are available
e. Formulate a hypothesis, connect with customers for a specific cluster and check if the hypothesis holds good.
f. The customers in the cluster that has maximum churn can then be used by the OLIST team for corrective actions. Analyse the behaviour patterns of the customers in this cluster to gain insight into why the customers churn.

Though Recency Frequency Monetary (RFM) and customer lifetime value looks attractive, it may not be able to provide accurate prediction for the given scenario due to insufficient data.

2. Non-ML solution

Since star ratings for each of the sentiments are already available, the sales/marketing team can measure the feedback manually, and based on feedback, it may take action.

Benefits of Proposed Solution:

The benefits of the proposed ML solutions are as follows:

i. Reduce potential customer churn
ii. Understand customers better, which can help in customer acquisition
iii. Understand the Lifetime Value (LTV) better

The LTV is calculated by multiplying the value of the customer to the business by their average lifespan. It helps a company identify how much revenue they can expect to earn from a customer over the life of their relationship with the company.

The average sales in an electronic spare store are $200 and, on average, a customer shops four times every two years. The lifetime value is calculated as: LTV = $200 x 4 x 2 = $1600.

Furthermore, the profit margin in the electronic spare store is 20%, hence the LTV is computed as follows: LTV = $200 x 4 x 2 x 20% = $320.

LTV can help a business estimate future cash flows and the number of customers they need to obtain to achieve profitability.

Summarise the Solution:

A clustering model is proposed to bucketise potential customer churns and, in the process, understand more about customer acquisition and LTV. The clusters can be further used to identify whether the cluster belongs to churn, potential churn or loyal category.

Prioritising Use Case:

Amongst the six use cases given to the Data Science team, this use case was ranked 5th based on the challenges and potential underfitting of the model.

Success Metrics and Key Performance Indicators

The churn rate formula is [(Lost Customers ÷ Total Customers at the Start of Time Period) x 100] is not an accurate metric for e-commerce websites as it is difficult to identify customers' buying patterns since no subscription fees are applicable on a monthly or quarterly basis. Monthly Recurring Revenue (after adjusting against revenue growth along with percentage increase in new customer acquisition can be a good indicator to identify Sentiment. The KPI can be measured as

$$w1(MRR) + w2 * (New\ customers\ acquired\ by\ referral) Where\ w1 + w2 = 1$$

Cite references

https://hevodata.com/learn/understanding-customer-churn-analysis/

https://www.netsuite.com/portal/resource/articles/human-resources/customer-churn-analysis.shtml

https://www.churnly.ai/compare.html

https://www.trifacta.com/churn-analytics/

Yu, X., Guo, S., Guo, J., & Huang, X. (2011). An extended support vector machine forecasting framework for customer churn in e-commerce. *Expert Systems with Applications*, *38*(3), 1425-1430.	The authors create a data warehouse to analyse customer behaviour data using extract transform load (ETL). Samples are obtained to train the classifier and prove that a modified Support Vector Machine Classifier performs better than ANN and standard SVM.

Gordini, N., & Veglio, V. (2017). Customers churn prediction and marketing retention strategies. An application of support vector machines based on the AUC parameter-selection technique in B2B e-commerce industry. *Industrial Marketing Management*, *62*, 100-107.	The authors extensively investigate machine learning techniques for class imbalanced and noisy data for customer churn prediction. The modified SVM classifier addresses the challenges of generalisation in class imbalanced data.
Berger, P., & Kompan, M. (2019). User modeling for churn prediction in E-commerce. *IEEE Intelligent Systems*, *34*(2), 44-52.	A successful prediction of churn of a specific customer provides an opportunity to change their decision to leave. The authors propose a novel complex user model focussed on user churn intent prediction based on composing of multiple sets of features representing the user's interaction with the web application.
Cao, J., Yu, X., & Zhang, Z. (2015). Integrating OWA and data mining for analyzing customers churn in E-commerce. *Journal of Systems Science and Complexity*, *28*(2), 381-392.	Customers are of great importance to e-commerce in intense competition and have been shown to follow the Pareto Principle, with 20% of customers generating 80% of the revenue. Thus, finding these customers is very critical. Customer lifetime value (CLV) is presented to assess the customers with respect to recency, frequency and monetary (RFM) variables. A novel model is proposed to analyse customers' purchase data and RFM variables based on ordered weighted averaging (OWA) and the K-Means cluster algorithm. OWA is employed to determine the weights of RFM variables in evaluating customer lifetime value or loyalty.

Customer Acquisition Cost Optimisation

Problem Statement:

The marketing team at OLIST runs multiple promotional campaigns to acquire new customers. However, the CFO believes that the marketing team is burning significant cash by offering huge discounts on products and other benefits, which is inflating the customer acquisition cost. The CFO wants to initiate a new process to measure the effectiveness of the acquisition campaigns by comparing them against the lifetime value of customers. Another way of increasing revenue is to gain more customers. The money that a company spends on getting one customer is called the acquisition cost. For instance, suppose OLIST needs to spend 30 Brazilian Real (BR) to acquire one customer. In this case, 30 BR is the acquisition cost of the customer. Obviously, it would be worth spending the 30 BR only if the customer generates more than 30 BR of lifetime revenue. So, the company wants to solve this optimisation problem..

Proposed Solutions:

1. ML solution

Customer acquisition cost (CAC) is the cost of acquiring every new customer. It gives an idea of how effective the marketing efforts have been. The ML team can try to predict the optimal CAC by understanding the CAC/LTV ratio from the business team. The steps involved in this process are as follows:

1 Compute the current LTV for each group of products based on the timeline over which the CAC is justified (The dataset shows that customers' frequency can be very low in many cases).

2 Obtain the ideal CAC/LTV ratio for each product group.

3 Obtain expected churn rate.

4 Build a model to predict the expected CAC for a given LTV and churn data based on historical data.

2. Non-ML solution

Benchmark CAC /LTV value to about 4.

Benefits of Proposed Solution:

The benefits of the proposed ML solutions are as follows:

i. Dynamically identify the optimal value based on product groups
ii. Understand customers better, which can improve the customer satisfaction

Summarise the Solution:

An ML solution to map CAC to LTV and customer churn is proposed. CAC is calculated by dividing all the sales and marketing costs involved to acquire a new customer within a certain timeframe. To get your CAC, divide all sales and marketing costs by the number of customers acquired over a given time period. CAC is an important metric for growing companies to determine profitability and efficiency.

For example, if a company spent $1500 to acquire 500 new customers in 1 month, their CAC is $3.00.

Prioritising Use Case:

Amongst the six use cases given to the Data Science team, this use case was ranked 3rd as an organisation needs to balance LTV and CAC ratios.

Success Metrics and Key Performance Indicators

Though the CAC/LTV ratio is used as a good success metric in subscription-based business models, the same does not hold good in e-commerce websites. To compute for the e-commerce value we can introduce $LTV_{[1]}$ which is the revenue generated by the customer in a given financial year. A good KPI can be

$$\frac{CAC}{LTV[1]} < 0.33$$

Cite references

https://www.bigcommerce.com/blog/customer-acquisition-engagement/#cac-formula

https://www.drift.com/platform/prospector/

https://unbounce.com/product/features/

https://www.referralcandy.com/

https://outgrow.co/increase-conversions/

Tillmanns, S., Ter Hofstede, F., Krafft, M., & Goetz, O. (2017). How to separate the wheat from the chaff: Improved variable selection for new customer acquisition. *Journal of Marketing*, *81*(2), 99-113.	Steady customer losses create pressure for firms to acquire new accounts, a task that is both costly and risky. Moreover, firms often use a large array of predictors obtained from vendors without knowledge about their prospects, which rapidly creates massive high-dimensional data problems. Therefore, selecting the appropriate variables and their functional relationships with acquisition probabilities is a substantial challenge. The authors use probability-based classification techniques and found them effective.
D'Haen, J., & Van den Poel, D. (2013). Model-supported business-to-business prospect prediction based on an iterative customer acquisition framework. *Industrial Marketing Management*, *42*(4), 544-551.	This article discusses a model designed to help sales representatives acquire customers in a business-to-business environment. Sales representatives are often overwhelmed by the available information, so they use arbitrary rules to select leads to pursue. The goal of the proposed model is to generate a high-quality list of prospects that are easier to convert into leads and ultimately customers in three phases. The model uses

	logistic regression, decision trees and neural networks.
Schwartz, E. M., Bradlow, E. T., & Fader, P. S. (2017). Customer acquisition via display advertising using multi-armed bandit experiments. *Marketing Science*, *36*(4), 500-522.	The authors show that customer acquisition would decrease by about 10% if the firm were to optimise click-through rates instead of conversion directly, a finding that has implications for understanding the marketing funnel.

Hansotia, B. J., & Wang, P. (1997). Analytical challenges in customer acquisition. Journal of Interactive Marketing, 11(2), 7-19.

Fraud Detection

Problem Statement:

Fraud is one of the most challenging areas to deal with in the e-commerce industry, as it can result in huge financial losses. There can be fraud in the areas of merchant identity, advanced fee, wire transfer scams, chargeback transactions, etc. The CFO wants to use the power of analytics to identify fraudulent transactions so as to help guard the organisation against such actions. An e-commerce marketplace is a platform that brings together sellers and buyers. Any fraud that happens between independent sellers and buyers will harm the company's image. The harm caused will have a direct impact on the revenue of the company

Proposed Solutions:

E-commerce fraud detection helps e-commerce businesses detect high-risk transactions using fraud detection and prevention practices. It generally uses algorithm-based analysis to assess each transaction's potential risk.

1. ML solution

Currently, no data related to fraud is available in the given datasets and hence it is not feasible. However, being a platform, the company can devise ways to capture such cases so that it may be possible to identify the frauds in the future.

For an ML approach to detect frauds, additional data, including IP geolocation, device fingerprinting and the number of transactions on new address deliveries, must be captured.

2. Non-ML solution

For effective fraud detection, frequent audit of site security is critical. The business can ensure PCI compliance and have a dedicated team to monitor the website for suspicious activity. In addition, the company should incorporate Address Verification Services (AVS) to detect suspicious credit card transactions in real-time. AVS ensures that the billing address and the credit card address match for every purchase.

Benefits of Proposed Solution:

Currently, no solutions are possible using the ML approach due to insufficient data. Many e-commerce businesses simply aren't prepared for the influx in fraud attacks that have occurred in the last year. Merchants new to e-commerce may not have the proper tools to effectively prevent fraud.

Summarise the Solution:

The organisation should start collecting more data to identify fraudulent transactions. Abnormal transactions should be monitored to understand the type of frauds. Traditional identity verification methods that use physical attributes (IP address, date of birth, social security number, etc.) are becoming less effective at accurately identifying fraudulent orders.

Prioritising Use Case:

Among the six use cases given to the Data Science team, this use case was ranked 5th as data is currently unavailable.

Cite references
https://shield.com/

https://castle.io/

https://thegood.com/insights/ecommerce-fraud/

https://www.maxmind.com/en/solutions/minfraud-services

https://spd.group/machine-learning/e-commerce-fraud-detection/

Dornadula, V. N., & Geetha, S. (2019). Credit card fraud detection using machine learning algorithms. *Procedia computer science*, *165*, 631-641.	The authors use clustering techniques to divide the cardholders into different clusters/groups based on their transaction amount, i.e., high, medium and low, using range partitioning. Using Sliding-Window, the transactions were aggregated into respective groups. Features including maximum amount, minimum amount of transaction, followed by the average amount in the window and time lapse was generated. The authors use the Synthetic Minority Oversampling Techniques to address the class imbalance data and obtain an accuracy to the tune of 99.9% in laboratory conditions.
Nanduri, J., Jia, Y., Oka, A., Beaver, J., & Liu, Y. W. (2020). Microsoft uses machine learning and optimization to reduce e-commerce fraud. *INFORMS Journal on Applied Analytics*, *50*(1), 64-79.	One major challenge in tackling e-commerce fraud results from dynamic fraud patterns, which can degrade the detection power of risk models and can lead to them failing to detect fraud that has emerging unrecognised patterns. The problem is further exacerbated by the conventional decision frameworks that ignore the follow-up decisions made by other associated parties, including

	payment-instrument-issuing banks and manual review agents. Microsoft fraud-management system (FMS) effectively tackles these two challenges using ML. After implementing these innovations over two years (2016–2018), Microsoft reduced its fraud loss by 0.52%, thus generating $75 million in additional savings. They reduced the incorrect fraud rejection rate by 1.38% and improved their bank authorisation rate by 7.7 percentage points.
Wang, S., Liu, C., Gao, X., Qu, H., & Xu, W. (2017, September). Session-based fraud detection in online e-commerce transactions using recurrent neural networks. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (pp. 241-252). Springer, Cham.	The authors propose a data-driven model using ML algorithms on big data to predict the probability of a transaction being fraudulent or legitimate. The model was trained on historical e-commerce credit card transaction data to predict the probability of any future transaction by the customer being fraudulent. Supervised ML algorithms, like Random Forest, Support Vector Machine, Gradient Boost or combinations of these, are implemented and their performances are compared. At the same time, the problem of class imbalance is taken into consideration and The techniques of oversampling and data pre-processing are performed before the model is trained on a classifier. Using a pipelined classifier, the authors achieve an accuracy of 97%.

Price Optimisation

Problem Statement:

The success of any e-commerce company depends on multiple factors in terms of variety, pricing, delivery and customer experience, to name a few. Price is a very sensitive component and needs to be very competitive to all the stakeholders, including the retailer registered in the platform. Price optimisation refers to the pricing approach where the best sales ratios are obtained. The sales ratios change dynamically due to demand and supply, competition and the stage of the product life cycle the item is currently in. Price optimisation cannot be static but needs to change dynamically based on the market trends. Though heuristic based approaches work, ML solutions can prove ideal to understand and learn the pricing dynamics.
However, before optimising the price you must make sure that you do not optimise the price of everyday items as customers do not wish to pay varying prices for such items. These items can include FMCG items, standard products, etc. On the other hand, customers are willing to pay variable prices for less frequently bought items, such as electronics, luxury items, etc.

Proposed Solutions:

1. ML solution

When done well, optimising pricing can positively impact the company's success. By setting the right price, various business elements will improve, such as sales, marketing, growth and profitability. The price optimisation strategy ensures changing the prices based on sales data and expected demand.

1. Categorise the sales data for each product as one of the following:
    1. Intermittent
        a. Whether there is a pattern to the intermittent data
    2. Lumpy
        a. If the intermittent pattern also generates variable volumes
    3. Regular and consistent volumes
    4. Erratic

2. Select the Intermittent and Regular categories for price optimisation as the data science team is not mature.
3. Squared coefficient of variation and Mean Inter demand interval can be used to classify the four different sales patterns.
4. Aggregate the data every 10 days.
5. Compute stock holding cost.
6. Compute the revenue loss due to out-of-stock items.
7. Estimate demand.
8. Build either a regression model (difficult) or classification model (about six categories).

2. Non-ML solution

Handle the problems one by one based on past sales and domain experience. The challenges increase as the number of products increases. That being said, the degree of price differentiation, whether the unique price per product, segment or personalisation, is strategic. Other factors, such as regulation, market practice and maturity of the pricing function, are all intrinsic to the process, showing each pricing decision's complexities.

Benefits of Proposed Solution:

The benefits include higher revenues, higher profits and better utilisation of old stocks. A retailer that offers 500 products may make more than thousands of pricing decisions a year. However, a retailer with 5,000 products and quarterly promotions will have to make hundreds of thousands of pricing decisions a year, which is not feasible.

Apart from increasing sales and revenue, it will also help acquire customers during periods of low demand as this system will reduce cost, which, in turn, helps acquire customers from competitors who are using constant pricing.

Summarise the Solution:

A price optimisation technique based on the available data is proposed using either regression or classification techniques. The discounting can be balanced between revenue and net profits.

Prioritising Use Case:

Amongst the six use cases given to the Data Science team, this use case was ranked 2nd as it can be directly related to the profits.

Cite references

https://www.intelligencenode.com/products/priceintelligence/

https://7learnings.com/blog/price-optimization-with-machine-learning-what-every-retailer-should-kn

ow/

https://tryolabs.com/solutions/price-optimization

https://www.sniffie.io/software-solutions/

https://www.vendavo.com/our-products/price-management-software/

Shams-Shoaaee, S. S., & Hassini, E. (2020). Price optimization with reference price effects: A generalized Benders' decomposition method and a myopic heuristic approach. *European Journal of Operational Research*, *280*(2), 555-567.	The authors consider a multi-period revenue maximisation and pricing optimisation problem in the presence of reference prices. The problem is formulated as a mixed-integer nonlinear program and a generalised Benders' decomposition algorithm is developed to solve it. A myopic heuristic model is proposed and discusses the conditions under which it produces efficient solutions. The analytical results and numerical computations illustrate the efficiency of the proposed solution approach and provide managerial pricing insights.
Vives, A., Jacob, M., & Payeras, M. (2018). Revenue management and price optimization techniques in the hotel sector: A critical literature review. *Tourism economics*, *24*(6), 720-752.	Price optimization (PO) methods seek to maximise hotel revenue and are based on inventory scarcity, customer segmentation and pricing. Different pricing policies have a greater impact than competition measurement effects in the hotel sector, as in the airline industry. This is mainly because differentiation strategies and policies at hotels can reduce the pressure of a competitive environment. The main contributions of the article were the presentation, description and classification of the principal revenue management (RM) and PO techniques in the hotel sector literature.
Kris Johnson Ferreira, Bin Hong Alex Lee, David Simchi-Levi, Analytics for an Online Retailer: Demand Forecasting and Price Optimization	The author takes the case study of Rue La La as an example of how a retailer can use its wealth of data to optimise pricing decisions on a daily basis. Rue La La is in the online fashion sample sales industry, where they offer extremely limited-time discounts on designer apparel and accessories. One of the retailer's main challenges is pricing and predicting demand for products that it has never sold before, which accounts for most sales and revenue. The authors used ML techniques to estimate historical lost sales and predict future demand of new products to tackle this challenge. The non-parametric structure of the demand prediction model, along with the dependence of a product's demand on the price of competing products, pose new challenges on translating the demand forecasts into a pricing policy and the authors develop an algorithm to efficiently solve the subsequent multiproduct price optimisation that incorporates reference price effects. They create and implement this algorithm into a pricing decision support tool for Rue La La's daily use. Field experiments were conducted to find that sales do not decrease because the tool recommended price increases for medium and high price point products. Finally, the authors estimate an increase in revenue of the test group by approximately 9.7% with an associated 90% confidence interval of [2.3%, 17.8%].