Proposal of a Lost Parcel mitigation plan focused in the US market

José Francisco Carvalho Mendes

Master's Dissertation

FEUP supervisor: Prof. Pedro Amorim

U. PORTO

FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Mestrado em Engenharia e Gestão Industrial

2022-07-11

Abstract

E-commerce increased tenfold in the first three months of the epidemic, and with the inclusion of footwear and jewelry, the global fashion industry brought in 1.64 trillion dollars in revenue in 2019. The fashion business has expanded by 4 to 5 percent since 2013, with the premium segment posting the largest gains.

Data will be the key to unlocking the insights needed to adapt to change and reengage clients in the upcoming months and years, through predictive models and data mining, with executives anticipating supply chain difficulties, an increase in domestic luxury spending amid sluggish overseas travel, and the further evolution of digital channels.

The main objective of this project is to develop a mitigation strategy for the lost parcels issue that Farfetch, the largest luxury fashion marketplace in the world, is currently experiencing in the US market. This was done through the Business Operations Team, a company talent incubator and internal consulting team that is focused on innovation.

A claim for lost parcels is made by the consumer and can have numerous different meanings, such as whether there are missing products in the order, the ordered box is empty, damaged goods, or even the non-delivery of orders. And the big picture problem of this is that in addition to the decreasing of the customer satisfaction score, It also results in financial losses.

For that reason, this project faced a number of difficulties, including developing the model that determined which orders were most likely to receive lost parcel refunds, figuring out the plan's triggers and mitigation initiatives, and ensuring these initiatives' technical and financial sustainability.

This way, assessing the LP root causes and beginning the business case preparation were the initial steps.

The following exploratory analysis measured the LPs that fell under the primary root causes, which helped to clarify which aspects might be the most crucial to consider when doing the future analysis on the net benefit.

After that, understanding the outcomes on the Lost Parcels project from the Fraud and the Data Science teams to position Business Operations in the right place to communicate and interact with both teams.

This way, this thesis provides an overview of what are the main characteristics of the orders that make them be more prompt to become refunded lost parcels, besides clarifying the whole process from the beginning to the end of the mitigation plan for lost parcels in the US market construction.

Resumo

O comércio eletrônico aumentou dez vezes mais nos primeiros três meses da epidemia e, com a inclusão de calçados e joias, a indústria global da moda gerou 1,64 trilhão de dólares de receita, em 2019. O negócio da moda expandiu de 4 a 5% desde 2013, com o segmento de luxo a apresentar os maiores ganhos.

Os dados serão a chave para desbloquear os insights necessários para a indústria da moda e do e-commerce se adaptarem às mudanças e fixarem os clientes nos próximos meses e anos, por meio de modelos preditivos e mineração de dados, já que os executivos antecipam dificuldades na supply chain, um aumento nos gastos domésticos de luxo e a evolução dos canais digitais.

O principal objetivo deste projeto é desenvolver uma estratégia de mitigação para o problema da perda de encomendas que a Farfetch, o maior marketplace de moda de luxo do mundo, vive atualmente no mercado norte-americano. Isso será feito por meio da equipa de Operações de Negócios, uma incubadora de talentos da empresa e equipa de consultoria interna com foco em projetos de inovação.

Uma reclamação de encomendas perdidas é feita pelo consumidor e pode ter inúmeras conotações diferentes, como produtos em falta na encomenda, uma caixa vazia, mercadorias danificadas ou mesmo a não entrega das encomendas. Além de diminuir o índice de satisfação do cliente, também resulta em perdas financeiras.

Por isso, este projeto enfrentou uma série de desafios, incluindo o desenvolvimento do modelo que determina quais pedidos são mais propensos a receber reembolsos de encomendas perdidas, descobrindo os gatilhos e as medidas de mitigação do plano e garantindo a sustentabilidade técnica e financeira das iniciativas.

Avaliar as causas-raiz do LP e iniciar a preparação do caso de negócios foram os passos iniciais.

A análise exploratória, a seguir, mediu os casos de encomendas perdidas que se enquadraram nas causas primárias, ajudando a esclarecer quais aspectos podem ser os mais cruciais a serem considerados, ao fazer a análise futura do benefício líquido.

Depois disso, entender os resultados em encomendas perdidas das equipas de Fraude e Ciência analítica para posicionar a equipa no lugar, de forma a comunicar e interagir com ambas as equipas.

Desta forma, esta tese fornece uma visão geral de quais são as principais características dos pedidos que os tornam mais prováveis de serem encomendas perdidas reembolsadas, para além de esclarecer todo o processo desde o início até o fim da construção do plano de mitigação de encomendas perdidas no mercado dos Estados Unidos da América.

Acknowledgements

Gratitude and best of luck:

To all those who helped me to complete this dissertation and who accompanied me throughout my academic journey: to FEUP and to the incredible team of professors and staff who, in the meantime, enriched my experience, in particular to Professor Pedro Amorim for his availability, good advices and support whenever I needed; to Farfetch, for trusting me and my abilities to carry out such an interesting and complex project, especially to Raluca and Pedro for their critical spirit and interest, even in the smallest details, which contributed so much to the realization of this thesis and to my growth.

To the people who accompanied me personally, for the fellowship and for sustaining the meaning of the cliche sentence "friendship for life": João, Miguel, Guga, Neves, Luís Paulo, and Bárbara.

Finally, to my family, especially to the generous pillars from whenever I bornt until now, and hopefully forever, who never expected anything in return: my parents, brother and grandparents.

For all those whose ways crossed with mine, thank you. We are other than the sum of us.

"The whole is other than the sum of the parts."

Kurt Koffka

Contents

Acronyms and Symbols

LP	Lost Parcel
LP CCC	Lost Parcel Customer Compensation Cost
DS	Data Science
EDD	Estimated Delivery Date
DSS	Direct Signature Service
Q4	Quarter 4
US	United States
RoW	Rest of the world
CS	Customer Service
RTO	Return to Origin
CB	Chargeback
AOV	Average Order Value

List of Figures

List of Tables

Chapter 1

Introduction

This chapter focuses on providing an introductory context of the project.

It presents the e-commerce marketplace business setting used as a case study to investigate the operations behind a major project inside the Business Operations Team.

Moreover, the motivation for this thesis is understood so as the goal of optimizing the savings for the problem.

Finally, the main project stakeholders are identified and the methodology followed throughout the document is displayed.

1.1 Company Description

FARFETCH Limited is the world's leading luxury fashion marketplace, founded in 2007 by José Neves for the love of fashion and launched in 2008. The FARFETCH Marketplace now connects customers in over 190 countries and territories with items from more than 50 countries and over 1,400 of the world's best brands, boutiques, and department stores, providing a truly unique shopping experience and access to the most extensive selection of luxury on a single platform.

Browns and Stadium Goods, which provide luxury products to consumers, and New Guards Group, a platform for the development of global fashion brands, are among FARFETCH's other businesses as well. FARFETCH's Luxury New Retail initiative provides the luxury industry with a wide range of consumer-facing channels and enterprise-level solutions. The Luxury New Retail initiative also includes FARFETCH Platform Solutions, which provides e-commerce and technology capabilities to enterprise clients, as well as in-store innovations such as the FARFETCH Connected Retail suite of technology products.

The problem of this thesis is mainly about the higher incidence of lost parcels claims on orders where the store's residence of the items is in the United States, so as the customer's. This type of orders will be called US Domestic. The rest, where the country of residence of the store is different from the United States and the customer's residence is in the US, are called US international.

The main factor for this difference is proven to be the carrier. Domestic orders are mostly carried by the carrier UPS, while the rest are carried DHL and FEDEX.

However, the crux of the problem is not just the quality of the carrier but the type of features the contracts Farfetch have with some carriers and not with others. In the case of DHL and FEDEX, the carriers have DSS (Direct Signature Service), which is the signature proof of receipt of the package by the customer. In the case of UPS, this feature is not available, thus making it impossible for the consumer to confirm receipt of the order, and for Farfetch to track the handover to the final customer. This increases the propensity for fraud on the part of the consumer, but also the propensity for sloppy delivery on the part of the carrier. So much so that the incidence of Lost Parcels definitely increases when looking at DHL's figures and UPS's.

The economic efficiency and practical feasibility of applying the DSS feature to the UPS carrier was then studied, so that the volume of lost parcels decreased significantly. Besides, other proposed actions will be discussed to prove the efficiency of the mitigation plan.

1.2 Team and Stakeholders Description

The Business Operations Team works like an internal consulting team focused on innovation, as well as a talent incubator for the company. The mindset is always of servicing and consulting within other teams in the operations department but also accountable for others. Business Operations will only be successful if making other business areas successful.

Based on a collaborative problem-solving strategy, Business Operation facilitates cross-functional engagement to identify business opportunities, design tailored solutions, and support during execution. The strategy always or almost always tries to have an analytical basis, which is then complemented with business sense.

This way, the Business Operations team's mission is to support different business areas on different types of projects by finding answers to complex issues with an independent perspective, from planning to execution. The main types of projects are the transformational ones, where the team pushes the boundaries of the current knowledge to achieve a step change in the business, the strategy redesign ones, where the team rethinks the approach to the customers and/or partners through a strategic perspective, and finally, the top and bottom-line improvement one where the aim is to improve operations through continuous improvement to generate top and bottom line gains.

1.3 Project's Description

We should comprehend that every endeavor is impermanent after contextualizing where it took place. As shown in the figure 1.1, the project team should develop a specific product or service that will transform the Business As Usual (BAU) processes into a new and enhanced BAU process.

When beginning a project, it is critical to answer the following questions: when will it begin and end, who is interested and will carry it out, what problem will the project address, and what are the primary goals. Although a basic answer to clarify the primary objectives is required, it may require a brief explanation of its context and why there is an opportunity.

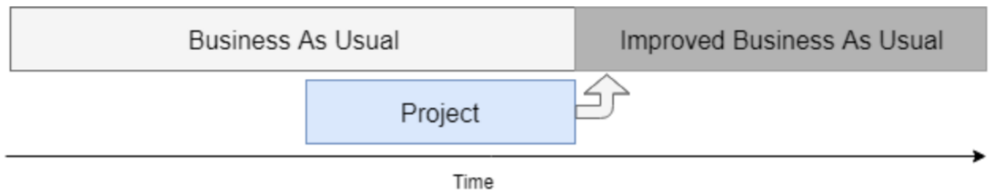


Figure 1.1: Project and Business as usual relation with time.

1.3.1 Project Motivation

The main motivation for this project was the unpredictable increment of the number of lost parcels. This fact led to financial losses, but that was not the only worry.

The reliability of the customer and their satisfaction are the most important factors, since without it, It would be impossible to do business. For that motive, the main motivations were the financial savings, to have more satisfied customers and also to keep up with Farfetch's good reputation on the delivery excellence.

1.3.2 Project Goals

The upcoming challenge is to build a framework under which orders at risk of becoming lost parcels can be identified and tackled in a targeted preventive manner, while respecting Farfetch 100% customer centric actuation, and covering both operational and fraud-born Lost Parcel claims.

This challenge comes up with the fact that the application of DSS to all the domestic orders would not be viable. This way, this service could only be addressed to orders that are more prompt to become refunded. More specifically:

- Build a model that identifies the orders that will be most prone to end up as refunded Lost Parcel claims.

- Identify the triggers and mitigating action(s) that should be activated to mitigate the risk of a Lost Parcel claim.
- Ensure the technical feasibility and the economic viability of implementing those mitigating initiatives

1.4 Methodology

The Cross-Industry Standard Process for Data Mining (CRISP-DM) was used to help guide project implementation (Schröer et al., 2021). This approach creates the framework for data mining initiatives. This strategy corresponds to the creation of the two mentioned goals: the creation of flags based on the analysis of different criteria to apply DSS to domestic orders and also the creation of a mitigation plan for Lost Parcels.

This provided a picture of the overall operational endeavors as well as refined certain key elements, opening the way for understanding the entire project, what had previously been completed, and what the specific aims for the period were. Because the information required was limited, it was important to coordinate the data collection techniques.

This genuine sales data was retrieved using SQL programming language through Google Big Query and also SQL Management Tool from many databases, the most important of which contained order, partner, and customer information. Because of the volume and complexity of the data, intensive preparation and treatment stages were required, deleting extraneous or inaccurate items.

With the data set ready for usage, it was time to determine the most important drivers - for the overall analysis as well as for each of the phases. Regarding the Catboost model, It was also necessary to prepare all the data, to remove the missing values, standardize the features, and then to create the model by dividing the dataset in Training Set, Test Set and Validation Set.

Finally, the analysis of the simple criteria and of the model's output was driven in MS Office Excel.

1.5 Thesis Outline

This document is organized in six chapters:

Chapter 1 gives the initial insights on the basis on which the project occurs, by presenting the company situation, the motivation of the project, the end objective with its development and the methodologies that were employed.

Chapter 2 evaluates the background literature, regarding the industry the company is inserted in, as well as the way raw data can be treated and visualized to extract the most reliable conclusions. Moreover, an introduction to predictive models is done.

Chapter 3 dissects the current problematic associated with Lost Parcels.

Chapter 4 introduces the the simple criteria analysis and the use of machine learning to model the probabilities of each order becoming Lost Parcels, and the results obtained.

Chapter 5 discusses the results and plans the initiatives implementation

Chapter 6 assesses the conclusions of the project and presents future research.

Chapter 2

Background Literature

The Background literature chapter provides the theoretical background that helps understanding the conditions surrounding the development of this project.

This overview analyzes both external environment of the business and the tools and methodologies that support the technical aspects presented further in the document.

2.1 E-commerce

According to Gregg and Ungerman (2020), during the first three months of the epidemic, e-commerce grew by a factor of ten, and following the COVID-19 outbreak, stay-at-home time rose while shopping frequency declined. However, the length of time spent at home then dropped, but the frequency of shopping trips rose. These, however, did not approach the pre-pandemic level. (Yuan et al., 2021)

To begin with, as a general tendency, customers understood the value of e-commerce shortly following the pandemic epidemic. Furthermore, this recognition did not drop considerably after that. Second, customers who modified their opinions regarding e-commerce stayed at home substantially longer than other groups. This suggests a positive relationship between the length of time spent at home and the need for e-commerce.

Furthermore, many respondents utilized e-commerce because they recognized its value, ease, and utility. As a result, even though teleworking and online classes did not become the norm following COVID-19, e-commerce usage is projected to increase in the future. As a result, future urban planners should address the necessity for adjustments in commercial facility placement. As a result, logistics networks had to evolve (Kawasaki et al. (2022)).

2.2 Luxury

Simply said, luxury has evolved into the luxurincation of the everyday. The term luxury is derived from the Latin word luxus, which implies sensuality, splendour, and pomp, and its derivative luxuria, which indicates excess, riot, and so on. The emergence of luxury in Western civilization is linked to rising prosperity and consumption. (Yeoman and McMahon-Beattie (2010)).

Atwal and Williams (2009) establish five perceived aspects for a brand to be considered a luxury goods seller: conspicuousness, uniqueness, quality, extended self, and hedonism.

The first three will produce goods that are immediately noticeable, have a strong emphasis on aesthetics down to the smallest detail, are as exclusive as possible, and are created using the best materials and techniques, all of which will ultimately raise their price. The latter two are personal considerations unrelated to the items.

Dubois et al. (2001) add two more crucial features: history and superfluousness. Luxurious items have some type of tradition or personal history behind them, such as pictures or jewelry from a noble family.

2.3 E-commerce Luxury Fashion Industry

With footwear and jewelry included, the global fashion business generated €1.64 trillion in annual revenues in 2019. Since 2013, the fashion industry has grown by 4 to 5 percent, with the luxury sector seeing the biggest gains. Part of this development may be attributed to Asia's resurgence in luxury spending. Despite the complexity and instability of the global environment in which luxury businesses operate, this positive growth persisted at around 4-5 percent in 2018 and 2019 and is anticipated to stay in the range of 3-5 percent through 2025. (Cabigiosu, 2020)

Many factors might impact the future of the luxury industry: the recent slowdown of economic growth in many countries, the recent adoption of protectionist policies in US and China, the digital revolution and the impact of technology on production systems, the influence of Millennials and Generation Z and the Brexit in the Eurozone.

According to Cabigiosu (2020), in this constantly changing, intricate, digital, and global environment, businesses require additional competencies and resources to be competitive. One billion euros is said to be the minimum amount needed for a worldwide luxury fashion company to thrive. Smaller businesses run the danger of becoming ensnared in a local competition that, for instance, does not result in success in nations like China or does not fully capitalize on e-commerce. Consequently, we have seen an increase in mergers and acquisitions since the year 2000. (Cabigiosu (2020)).

Every firm strives to have order-winning operational attributes due to the competitive environment imposed by customers and rivals. This is mostly manifested in quick order to delivery timeframes and a broad product selection. Giving the consumer what they want at the right time has a significant influence on his retention, which increases future sales. Besides, also the low return rates and low Lost Parcel incidence, are a significant metric that also takes part on the retention rate and the future sales.

In this approach, the luxury fashion sector positions itself as a high-potential market for suiting competitors, who must be very adaptable and have a strong emphasis on client centricity. To do so, one may take advantage of this data-immersive environment and continually use cutting-edge analytical techniques to provide the essential business insights. (Sheng and Liu (2010)).

2.4 International B2C Supply Chains and Risks for the Luxury Goods Market

According to Amed et al. (2022), Executives anticipate that supply chain difficulties, an increase in domestic luxury spending amid sluggish overseas travel, and the further evolution of digital channels will have the most influence on their organization in 2022.

The BoF-McKinsey study found that 84 percent of CEOs ranked global supply chain disruption as their top priority, as the instability of the previous two years — which included material shortages, traffic jams, and skyrocketing shipping prices — is anticipated to persist in the upcoming year.

As the logistics industry continues to expand and issues intensify, executives must pay particular attention to the transparency and control systems in their supply chains in order to meet consumer demand.

Profitability will unavoidably be impacted by supply chain challenges. 67 percent of fashion executives forecast that retail prices would increase by 3.2% on average in 2022 as a result of these cost inflation pressures, with 14% anticipating price increases of above 10%. Because many players will be concentrating on sales growth, price increases will be used to offset decreasing profits.

Despite this, 17 percent of executives say they will cut prices, with the mid-market anticipating the biggest reductions—possibly as a result of the segment's ongoing difficulty throughout the year.

The sustainability gap, which was cited by 15% of executives in the BoF-McKinsey report as one of their top three worries for 2022, is the second biggest concern on executives' minds after supply chain disruptions. The business benefits associated with minimizing their company's impact on the environment and society, however, may outweigh any costs or challenges they encounter, according to the 12 percent of CEOs who see sustainability as an opportunity in the upcoming year.

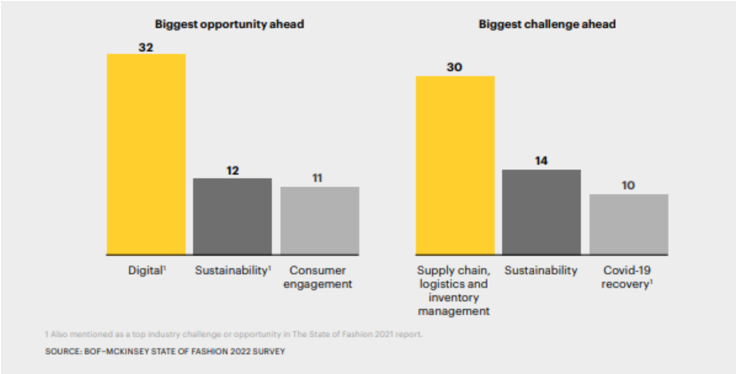


Figure 2.1: Biggest Opportunities and challenges of the Luxury Goods Market. Amed et al. (2022)

2.5 The importance of data and analytics to the business of fashion

According to Sandrine Devillard and Altable (2021), Data will be the key to unlocking the insights required to adapt to change and reengage clients in the upcoming months and years. Despite this, the outbreak due to Covid-19 has shown a large gap in data gathering and analysis across the majority of the industry. The gap between data leaders and laggards has widened; while some data-savvy luxury and fashion firms have seen their market values soar, others have been overtaken by rivals. In fact, the top 25 retailers account for more than 90% of the increase in the sector's worldwide market value during the epidemic, with the bulk of them serving as examples of the significant shift to digital, data, and analytics.

Simply said, it is better for fashion and luxury firms to learn how to use data as soon as possible.

Data is more abundant than it has ever been, and the COVID-19 situation has made the case for data capacity much stronger. Fashion and luxury brands that are expected to emerge from the crisis stronger than before are using data to keep ahead of the competition.

The applications of data and analytics are vast, diverse, and well-known, but knowing where to focus along the value chain isn't always obvious. Identifying where and how to integrate data into the business in a cross-functional manner, as well as developing the proper operating model to do so, is frequently a challenge. Fashion companies are more and more navigating and extracting value across the whole value chain over the last 12 months.

Fashion and luxury industries, in particular, have seen concrete gains from incorporating data into their planning, merchandising, and supply-chain activities. Stock and store optimization decisions based on data have increased sales by 10%. Furthermore, improving visibility throughout the supply chain has streamlined inventory management, enhanced returns forecasting, and optimized transportation networks, resulting in inventory cost savings of up to 15%.

Most importantly, fashion brands who have used data to tailor customer e-commerce experiences have seen a 30 to 50 percent increase in digital sales.

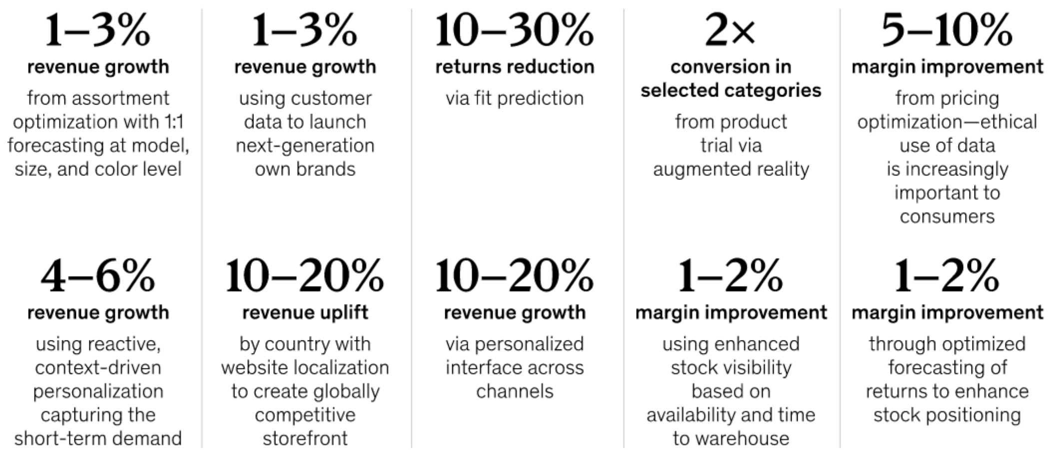


Figure 2.2: Example Impact of data and analytics adoption by e-commerce fashion companies (Sandrine Devillard and Altable, 2021)

2.5.1 Machine Learning Categories

Machine learning can be assigned to three categories, as seen in figure 2.3. Supervised machine learning, unsupervised machine learning and Reinforced learning. The difference between supervised and unsupervised learning is that when using supervised models, the user has created a predefined label with a set of characteristics (Batta, 2020).

Whereas the unsupervised algorithm, it interprets the data set by clustering the data into different classes based upon the relation it has recognized between different records.

Reinforcement learning is moderately different, the user provides the algorithm with an environment and the algorithm makes decisions within that environment. It is continually improving itself with each decision based on the result of the previous decision.

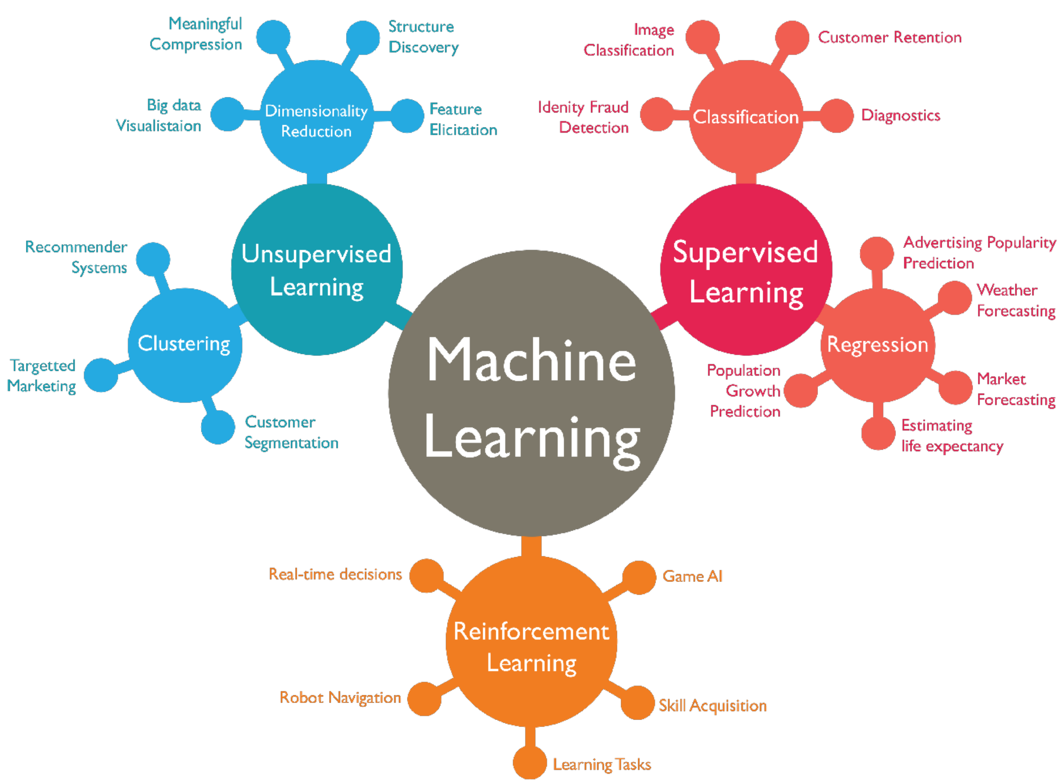


Figure 2.3: Machine learning and its three main categories Techleer (2017).

2.5.2 Gradient Boosting Algorithms

Gradient boosting is an ensemble technique that attempts to create a strong learner from a given number of weak learners, i.e models that only perform slightly better than random guessing. (Natekin and Knoll, 2013)

An ensemble is a set of predictors, all trying to predict the same target variable, which are combined together in order to give a final prediction. The use of ensemble methods allows for a better predictive performance compared to a single predictor alone by helping in reducing the bias and variance in the predictions, this way reducing the prediction error.

This process is done by creating a model from the training data, then creating a second model which will try to correct the errors from the first model and so forth, until a high accuracy value is achieved or the maximum number of models is added. In the end, it combines the outputs of every weak learner in order to create a strong learner which is expected to improve the prediction performance of the model.

In the case of classification problems, the results of the weak classifiers are combined and the classification of the test data is done by taking a weighted vote of each weak classifier's prediction. The main principle of boosting is to iteratively fit a sequence of weak learners to weighted versions of the training data. After each iteration, more weight is given to training samples that were misclassified by earlier rounds.

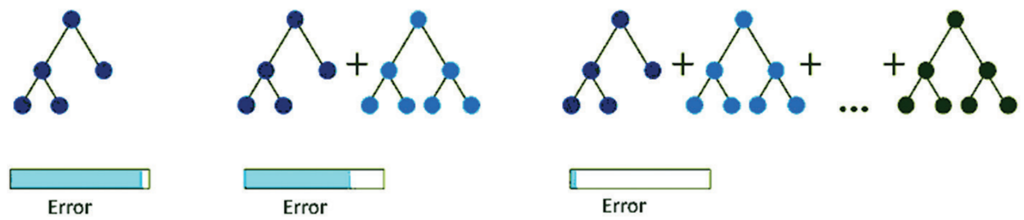


Figure 2.4: Gradient Boosting Principle

At the end of the process, all of the successive models are weighted according to their performance and the outputs are combined using voting for classification problems or averaging for regression problems, creating the final model.

2.5.3 Catboost

Introducing CatBoost, a novel gradient-boosting algorithm proposed by Prokhorenkova et al. (2018), and it successfully uses categorical features while causing the least amount of information loss. Compared to other gradient boosting methods, CatBoost is unique. To address the issue of target leaking, it first employs ordered boosting, an effective version of gradient boosting methods. Additionally, this approach works well with small datasets.

Third, CatBoost has the ability to manage category features. Typically, this treatment is finished during the preprocessing stage and entails changing the original category variables to one or more numerical values. Moreover, Prokhorenkova et al. (2018) found that CatBoost can be successfully applied to diverse types and formats of data.

CatBoost has recently been used in a variety of industries, including finance, and with a variety of data formats, including time series data. A new binary feature is added for each category in place of the old variable. Another benefit of the approach is that it avoids overfitting brought on by conventional gradient boosting algorithms by performing random permutations to estimate leaf values when choosing the tree structure. Binary decision trees are used by CatBoost as its primary predictor. Following is a description of the estimated output provided by Dorogush and Gulin (2018):

$$Z = H(x_i) = \sum_{j=1}^{j} c_j 1_{x \in R_j} \tag{2.1}$$

where $H(x_i)$ is a decision tree function of the explanatory variables x_i, and R_j is the disjoint region corresponding to the leaves of the tree.

Assuming we observe sampled data:

$$D = (x_j, y_j)_{j=1,\ldots,m} \tag{2.2}$$

Where

$$X_j = (x_j^1, \ldots, x_j^m) \tag{2.3}$$

is a vector of n features and response feature $y_j \in \mathbb{R}$ which can be binary (i.e yes or no) or encoded as numerical features (0 or 1). Samples (X_j, y_j) are independently and identically distributed according to some unknown distribution p(,).

The goal of the learning task is to train a function $H : \mathbb{R}^n$ n which minimizes the expected loss

$$L(H) := EL(y, H(x)), \tag{2.4}$$

where L(,) is a smooth loss function and (X, y) is a testing data sampled from the training data D.

2.6 Model Validation

Predictive models have a goal of being able to generalize outputs from inputs that they have never received. To effectively validate the results from the model, one has to divide the data-set in three components.

The training set is the data from which the model will learn and iterate its best fit. The validation set is used to evaluate the training results and to tune parameters.

Finally, the test set is the data the model receives once it is fully trained, to understand its behavior in unforeseen conditions.

The training set consists of the majority of the data set, and the validation and test sets are divided equally from the remaining data. If this is not done, the model will be evaluated using data it was trained on, giving false conclusions about its performance. This split is done in a sequential order, in order to capture the temporal relationships between the observations.

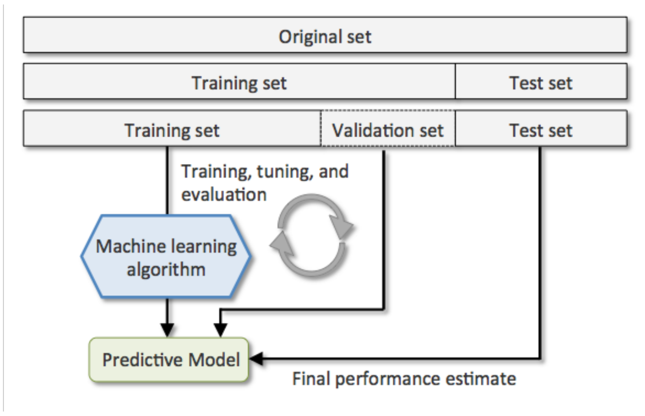


Figure 2.5: Model Validation Principle and flow

2.7 Model Evaluation

When classifying data, the goal is to predict the class labels. Two possible output classes are the only ones available in binary categorization (i.e., Dichotomy). In multiclass classification, there may be more than two potential classes. The project profile dictates that only the binary ones will be discussed. (Shah et al., 2021) and (M and M.N, 2015)

The effectiveness of classification can be evaluated in numerous ways. Some of the most used metrics include accuracy, log-loss, and AUC-ROC.

Simply put, accuracy reflects how frequently the classifier predicts correctly. Accuracy is determined by dividing the total number of forecasts by the proportion of true predictions.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{2.5}$$

Evaluation metrics like recall and precision operate differently to have different effects. Precision and recall are frequently traded off.

Recall increases if Precision is lower, and decreases if Precision is higher. Equation (2.6) illustrates how precision, also known as the Positive Predictive Value, assesses how well the positive cases were predicted out of all the positive cases:

$$Precision = PositivePredictiveValue = \frac{TP}{TP+FP} \tag{2.6}$$

Recall is one of the most crucial evaluation measures used in identifying fraudulent credit card transactions. It is also referred to as True Positive Rate (TPR) and Sensitivity. Its significance depends on its capacity to identify successful cases.

The detection of fraudulent activities increases as the recall value rises. As a result, even while the suggested model may acquire a respectable amount of FP, it should not be at the expense of

FN as much as feasible. It is crucial to obtain the higher Recall value as much as possible to avoid missing any incidents of fraud. Recall is expressed in Equation (2.7):

$$Recall = Sensitivity = TPR = \frac{TruePositive}{TotalPositives} = \frac{TP}{TP+FN} \tag{2.7}$$

Although vital, recall is more crucial than precision.

The actual TP cases of all expected positive instances are what precision is concerned with. However, as long as the proposed model does not sacrifice FN cases as much as possible, making a reasonable number of FP instances using it is acceptable.

Recall is therefore more crucial than Precision in the detection of credit card fraud. F1-Score, on the other hand, combines the Precision and Recall scores to assess the model's performance.

The classifier with the highest F1-Score must be picked since the F1-Score is one of the evaluation metrics taken into account when comparing two or more models.

Equation (2.8) is the expression of F1-Score:

$$F1 = 2 \times \frac{TP}{TP+FN} \tag{2.8}$$

Finally, one of the key measures to evaluate the effectiveness of a classification problem is LogLoss, also known as Logarithmic Loss or Cross-entropy Loss.

For a single sample with true label $y \in (0,1)$ and a probability estimate $p = Pr(y = 1)$, the log loss is:

$$LogLoss(N = 1) = y \times log(p) + (1-y) \times log(1-p) \tag{2.9}$$

2.8 Final Considerations

The investigation presented in this chapter has a supportive role in the realization of this project. Understanding the e-commerce luxury fashion industry reveals the importance of optimizing the lost parcels rate, as to meet the customer's requirements.

After reinforcing the inherent motivation, the study of project management was used in the diagnosis of the order cycle processes, aiding in data gathering, manipulation and visualization, followed by having more sense of the timeline of the project. This way, there is a comprehensive baseline to enable accurate insight generation.

Finally, the investigation on predictive models is useful to complement the big data processes previously described.

Chapter 3

Lost Parcel Problematic in the US market Diagnosis

When the core business of a company implies shipping worldwide, Lost Parcels are problematic to defeat since they represent direct and indirect costs. The next chapters insights and inferences will be based on data from July first of 2021 until January thirty first of 2022.

The main market for Farfetch is the US, representing almost 30% of the total sales. Due to that, this market deserves special attention, especially as the rate of Lost Parcels is higher than in the rest of the world .

This rate was representative of more than 60% of Farfetch's global Lost Parcel CCC, which is the Customer Compensation Cost due to Lost Parcel claims, representing the sum of all the refunded orders value, and that can also be traduced into 1.4 Million dollars loss in just 2 months (January and February 2021).

Given this, this chapter will introduce the used methodology from the beginning of the project, its state of the art, the conducted analysis and the making of the predictive model alongside the results of each approach.

3.1 Project's Presentation and methodology

The Business Operations team approach to the problem was to first understand the hike in Lost Parcels and what led customers to ask for Lost Parcel tickets and also which of them were refunded. For that, several steps were taken:

- Assess the Lost Parcel root causes (the reasons behind the LPs tickets and when in the order's lifecycle is the LP case created)

- To quantify LPs that fall under the main root causes (the weight of each of the root causes have on the total LP CCC, and what does It tell us about the customers, their pain points and intentions)

- To understand the outcomes on LP from Fraud and Data Science (the investigation process for each team, and how many times are the outcomes contradictory)

- Finally, to align an action plan and to see what initiatives can be implemented to address LPs' root causes.

3.2 Order Cycle and the Lost Parcel claims

The order cycle starts when the customer checks out successfully online while buying an item, and ends when it has been delivered at the first or subsequent attempts. Then, the customer may keep the order, or return it and receive a refund from Farfetch.

Lost Parcel claims have several initiation moments, and are all triggered by the Client. The first situation is when the LP claims are created is when the carrier still didn't deliver the order, but the period time is still inside the estimated delivery date, which are naturally very few (3% of all the LP claims). 10% of the total LP claims happen when the carrier still didn't deliver the order but this time, after the estimated delivery time.

The most important claims (and the ones around which this project is centered) are created when the order was already supposedly delivered by the carrier, but the Client claims not to have received the order. This type of claims represent 71% of the total Lost Parcel claims.

It's relevant to say that 24% of those are refunded, and represent a cost for Farfetch, when looking at the year of 2021. Besides that, LP claims also happen when the order has been returned- these situations represent 9% of LP claims in the period analyzed.

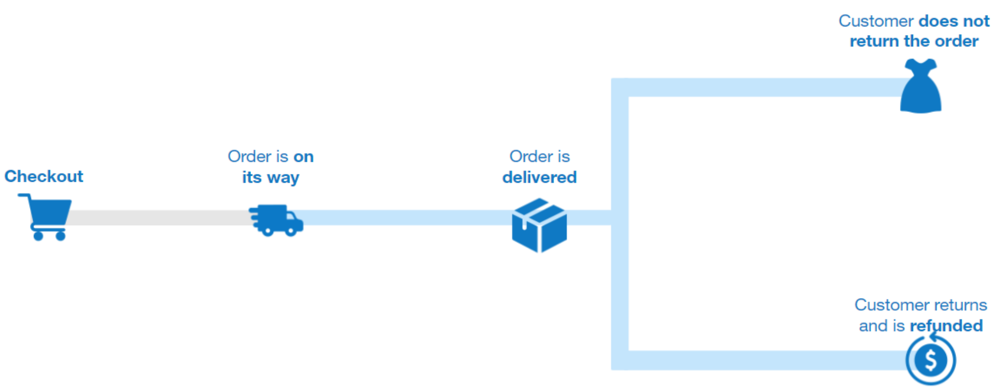


Figure 3.1: Order Cycle Flow

One of the most important findings in the initialization analysis was the overall incidence of accepted claims per carrier, since UPS was definitely the worst with a 1.4% incidence while DHL, for instance, only had a 1% incidence.

For this, the causes for this discrepancy were studied and the main inference was that this difference, from a carry to another, was happening when the order was “ supposedly delivered by the carrier”.

For a better definition of Lost Parcels, a division through the order cycle was made as shown in the figure 3.4.

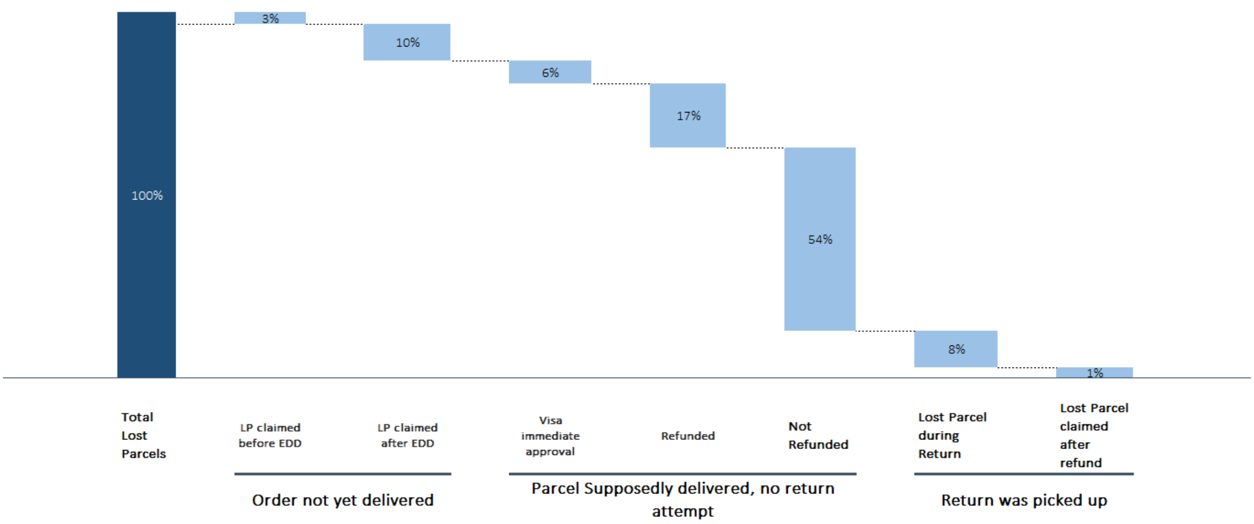


Figure 3.2: Division of Lost parcel claims through the order cycle.

Together with this, the fact that UPS didn’t count with a Direct Signature Service (DSS), which is a proof of receipt from the customer when he receives the order happened to be the main trigger to study the viability of applying this service to UPS, once FedEx already had it, and the LP incidence when the order was “supposedly delivered” was lower.

In fact, the LP incidence per type of delivery options was also studied, and when the order was delivered with the customer present (which in the case of having DSS, is mandatory), the incidence was lower than If it’s delivered at a neighbor or in a “safe place”.

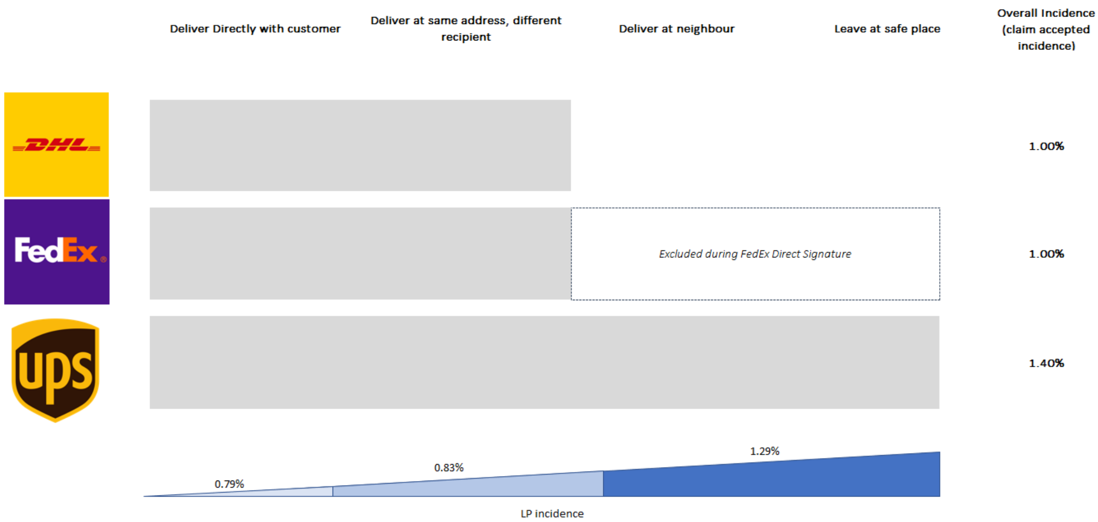


Figure 3.3: LP incidence per carrier and type of delivery service.

3.3 Project's State of the Art

Since the beginning of the increase in the rate of Lost Parcels and also in the rate of refunded volume for Lost Parcel reasons, the Business Operations team has been brought in to help resolve the issue through an independent stance. Lost Parcels are a cross functions issue shared by Customer Service, Delivery, and Fraud operational verticals.

Right on the onset of the Covid-19 pandemic, fraudulent activity and delays in supply chains became increasingly strong headwinds to Farfetch's global delivery business model. As an indicator sensitive to both these headwinds, Lost Parcel rate in the United Stated jumped from 1.1% to 1.5% from May '20 to Jun'20. Then, in November 20 this value increased again from 1.5% to 1.8%, alongside with the rate of refunded orders, which increased in about 3%. The increase of the rate in Q4 '20 was a consequence of adding a new carrier and a new delivery type to the United States market, which caused disruption.

This led to an alarming net benefit loss which was reflected on the Business Operations Team main work focuses. The initial approach was to decrease LP levels covering the entire customer journey, starting with the order checkout on the marketplace, and ending in the return of the order, or not.

One of the first proposed actions was to change the parametrization of payment refusals to decrease fraud levels, which would be a responsibility of the fraud team.

Fromthe beggining until the order is on its, but not yet delivered, proactive communications initiative was applied. This measure consisted in communicating with the customers if the order does not move or the delivery is delayed, to increase responsiveness from CS to carriers or customs' requests, and finally, to give customers the option of waiting for the parcel, or to be refunded via RTO.

In the end of the order cycle, in case of the customer not returning the order, applying DSS to UPS and continuing the trial with FedEx were the measures to study. This initiative consisted in applying DSS to one of the carriers with the worst LP incidence. This way, as shown before, the fact that the order is delivered to and only to the client would decrease the LP claims, and Farfetch would benefit from a proof of receipt in case of this claims.

In case the customer does return the order, pushing partners to accept returns to consequently unblock refunds was also something to study.

3.4 Features for Net Benefit Analysis

Based on the rate of Lost Parcels for a carrier with DSS option, and another with no access to that service, we conclude that Direct Signature Service is a fundamental service in order to reduce the number of Lost Parcel claims, mainly when the order is supposedly delivered.

It's also important to mention that FedEx, which has the DSS service only deals with international orders (from any other country, except the US, to the US). On the other hand, domestic orders are the ones which only work with UPS, that doesn't use DSS.

This way, It was necessary to understand if It would be viable to apply DSS to domestic orders since the price of DSS for UPS is not the same as the price of DSS for FedEx. This was accurate in a preventive way, since the conclusions of this analysis were that the breakeven cost of applying DSS to domestic orders would be 21% lower than the DSS cost.

	Cross-Border or International *with Direct Signature*	**Domestic** *without Direct Signature*
Share of orders sent	77%	23%
LPCCC cost per order		153% higher than international orders
CB non fraud net cost per order		70 % higher than international orders
RTO non covered cost		90% lower than international orders
Breakeven Direct Signature Cost	-	21% lower than the DSS cost

Figure 3.4: Project and Business as usual relation with time.

Once the cost of applying DSS is higher than the breakeven cost, the only way of becoming viable would be to find a way of only applying this service to orders that are most prompt to become Lost Parcels and Refunded, in a way that the cost of becoming a Lost Parcel with an associated refund surpasses the cost of applying DSS.

This way, the plan was to focus on features that at first sight, could have an impact on the Lost Parcel incidence directly and indirectly. For instance, It wouldn't make sense to analyze the time an order occurred since It is, at first sight, independent from the probability of being refunded, being a lost parcel or the value of the refund.

Nonetheless, there are features that make sense to analyze. Regarding those, an analysis was made to the different cities in the US (city). Also an analysis of the payment method type was made because of the different services some have and that people could take advantage of in order to be refunded. Furthermore, an analysis of the customer himself, like If It's a New or existing customer, If It has a Farfetch tier or if It doesn't.

Parcel features, on the other hand, are the ones that have a direct relationship with the products of an order itself, like the average order value and the category of the product. The order value can be a good criteria since It impacts directly the Refunded Value of orders (for instance, if an order is valued in 3000$, the totality or almost the totality will be refunded, which will represent a higher cost for Farfetch). On the other hand, there are several categories that have, on average, a bigger value, and that are more prompt to be refunded or to become a lost parcel.

Besides the simple criteria analysis, Data Science had a prepared model where the output were probabilities of a certain order becoming a Lost Parcel. These probabilities were based on previous data. The model was returning a high incidence of Lost Parcels relatively to the total number of returned orders. For instance, in a total of 20% of the orders, 2.5pp of these were Lost Parcels, when there's a 1% incidence for these, normally. The problem of this result was that, although it returned a good volume of Lost Parcels when applied to a probability threshold, financially, It had a negative effect on the net benefit, or savings. Meanwhile, the simple criteria analysis had an important role in improving the model, as It clarified the biggest area of opportunity.

Having arrived here, the communication between the Business Operations team and the Data Science team was very important, since it was necessary to influence and explain to the DS team the whys of making sense that the new model focus on the Value of the Refunded Orders and not on the Volume of these.

Nevertheless, due to the low rate of Refunded orders in a big sample with lots of independent features makes it hard to be accurate.

Below, there are the features that were necessary to collect from the databases from Farfetch through SQL on Big Query. The code for that will be available on the Appendix C. After that, through excel, an analysis was conducted on each of the independent variables (except the Order Code ID). The dependent variables that the analysis focused on were the binary one "Lost Parcel?"

which tells us if the order had a Lost Parcel claim, the binary “Refunded?” which tells us if the order was refunded or not and the continuous one “Model Probability”, which is the Model’s dependent variable, telling the probability of an order to become refunded. Also an analysis to the model’s probability was driven and will be our main focus in the subchapter 4.2.

Table 3.1: Dataset Features and their description.

Type	**Field Name**	**Description**
Independent Variables	*Order Code ID*	Order Identifier
	Order Value	Value of the order with shipping costs to the customer
	Guest/ Not Guest	Whether the customer is in a guest profile or not while buying
	Existent/New	Whether the customer is existent or new
	Category	Category of the products in the order
	City	Representative City of the Shipping address
	Tier	Tier of the customer (FF-Bronze, FF- Platinum, FF-Private Client, none)
	Payment Method	The type of payment method (Klarna, Apple Pay, etc.)
Dependent Variables	*Lost Parcel?*	Whether the order became a Lost Parcel or not
	Refunded?	Whether the order was refunded or not
	Chargeback Status	The value of the chargeback if asked
	Model Probability	Model’s Output giving the probability of becoming a Refunded Lost Parcel

Chapter 4

Net Benefit Diagnosis and Results

Financial loss was the main driver of this case. This way, net benefit is the most important metric of the whole analysis and It was also the main focus during the whole case, although some of the proposed actions were driven by for instance, customer satisfaction index.

For this, net benefit is the difference between the costs of applying DSS and the costs of leaving the case AS IS. In the AS IS scenario, the LP CCC remains the same, as well as the CB cost and the RTO cost. On the other hand, the "with DSS" scenario has a lower LP CCC (It was assumed that with DSS applied, the domestic orders would behave like the international orders in terms of LPCCC, CB Cost, and also RTO costs). Chargeback costs are direct refunds that some banks have in the contracts with Farfetch, but that won't be an analysis goal for this thesis.

For this last scenario, It is, naturally, necessary to add the DSS cost. It's also important to say that It was assumed that the CB costs and the RTO costs variation was disposable, since in fact the differences were too small to biase the results from the LPCCC variation. With this analysis, we could predict the net benefit of applying DSS to domestic orders, assuming that they would behave like international orders (orders with DSS). For that reason, the next equation was applied to all of the features to compare the different Net Benefits:

$$\text{Net Benefit = Total Cost}_{withoutDSS} + RTOcost_{withDSS} - RTOcost_{withoutDSS} + CBCost_{withDSS} - CBCost_{withoutDSS} \tag{4.1}$$

in, which:

$$TotalCost_{withDSS} = LPCCC_{internationalorders} + CostofDSS \times VolumeofOrders \tag{4.2}$$

$$TotalCost_{withoutDSS} = LPCCC_{domesticorders} == \sum_{i=1}^{n} RefundedValueoforderi \tag{4.3}$$

where $i \in DomesticOrders$

$$LPCCC_{internationalorders} = \sum_{i=1}^{n} RefundedValueoforderi \tag{4.4}$$

where $i \in InternationalOrders$

4.1 Simple Criteria Analysis

At Farfetch, there's a programme called selective allocation through which there's the possibility of assigning a specific service per order.

This means that in order to apply DSS to specific orders only, It would be necessary to use the programm.

In order to make the allocation system work, several steps must be taken into account. Firstly, the technology team needs to develop the connection between the service and the selective allocation programme. After that, the product team needs to agree with UPS about what is about to be done, and if it doesn't go against their politics. For this specific case (the application of DSS to UPS), the criteria that the system can digest are just simple criteria, or an order flag. This order flag can come from different fonts. In this case, the flag would be triggered by the built model that will be explained next .

This analysis consisted in applying the formulas above and trying to predict how much Net Benefit would Farfetch have saved if DSS was applied in the domestic orders, for the period of Jul'21 to Jan'22. If the savings are positive, that means, the orders that contribute for that positive net benefit savings would be flagged in the selective allocation and would have DSS when delivered.

Due to confidentiality terms, the absolute values of the total number of orders of each criteria, of the total and per order LPCCC as well as the CB costs and RTO costs. This way, the most visual way to show the results is through the variation of each of the costs, both before and after the DSS application.

4.1.1 Payment Method

Payment method is the bank, app or card that was used to pay for an order. There are several types of payment in Farfetch's marketplace such as credit card, debit card, Apple Pay, Klarna, AlyPay, etc.

This was the first criteria to be analyzed, in which Klarna Pay Later was definitely the one in which Lost Parcel claims reduced and with the best net benefit savings.

As mentioned before, the CB lost cost per order and the RTO unrecovered cost variation are disposable, but the LPCCC per order variation decreases the enough to be viable to pay for it.

Due to the pay later service, customers tend to try to make more claims because of the fact that they still didn't pay, they already have the article and they still want a refund. This way, by applying DSS to the orders that have this type of payment would give Farfetch the benefit of having the receipt proof, and not accepting claims from customers.

The table below is also a good example of how It would not be beneficial to Farfetch to apply DSS to all the payment methods, once the increment on the costs because of the DSS application to all orders surpasses the benefit that the DSS brings on the decreasing of the LP CCC cost.

Table 4.1: Case if DSS was applied to all the orders bought with KlarnaPayLater

		Klarna Pay Later	All tiers
1. Value			
	LPCCC as %GTV	1.3%	0.6%
	Share of GTV	9.3%	100.0%
	Share of LP CCC	19.4%	100.0%
2. Volume			
	Share of Orders Volume	12.5%	100.0%
	LP incidence (%)	2.4%	1.5%
	Share of Refunded Volume	19.4%	100.0%
3. Cost per order variation if DSS was applied			
	LPCCC / order	-66.2%	-60.0%
	CB lost / order	Disposable	Disposable
	RTO unrecovered	Disposable	Disposable
		plus DSS cost	
4. Total Cost Variation		-17.0%	6.8%

4.1.2 Category

Another important criteria that stood out was category, as we see below. Category is a feature which tells which type of item the order had.

They can be shoes, clothing, accessories, jewellery, etc.

However, bags were the category that obtained the best results of applying DSS. This might have 2 reasons behind.

First of all, bags are a medium expensive article and consequently has a higher refund value, representing a loss due to refund and due to the article loss, in fraud cases or in case of being damaged, for example.

Table 4.2: Case if DSS was applied to all the orders with bags

		Bags	All Categories
1. Value			
	LPCCC as %GTV	0.6%	0.6%
	Share of GTV	15.2%	100.0%
	Share of LP CCC	13.6%	100.0%
2. Volume			
	Share of Orders Volume	7.1%	100.0%
	LP incidence (%)	1.5%	1.5%
	Share of Refunded Volume	13.6%	100.0%
3. Cost per order variation if DSS was applied	LPCCC / order	-66.0%	-60.0%
	CB lost / order	Disposable	Disposable
	RTO unrecovered	Disposable	Disposable
		plus DSS cost	
4. Total Cost Variation		-30.9%	6.8%

4.1.3 City

The feature city represents the city where an order's shipping address belongs to.

This feature was also an important one, not only for the simple criteria analysis but also for the model, once at first sight It didn't seem to be an impactful feature in the Lost Parcel incidence.

However, looking at the results there were some cities that had not only a higher incidence on Lost Parcels, but also a higher average order value. Glendale is an example of those, which in fact is one of the wealthiest zones in the US.

It was also important to guarantee a data balancing in order to compare the domestic orders with the international orders, and also to make the assumption of the behavior with DSS applied. That is, the assumption of the domestic orders behaving like international orders in case of the application of the DSS.

Although there were other cities like Los Angeles and Brooklyn, in New York with also positive Net Benefit Savings, they weren't summed up to the Glendale case once the savings were lower, in a way that variance could have a negative effect on the savings.

This also happens in the category features and in the average order value one, where there are other criteria that have a positive impact on the Net Benefit, but their impact is so little, that in other time period or in other external environments, these could represent a negative impact on savings due to variance.

Table 4.3: Case if DSS was applied to all the orders with Glendale as destiny.

		Glendale	**All cities**
1. Value			
	LPCCC as %GTV	2.8%	0.6%
	Share of GTV	0.7%	100.0%
	Share of LP CCC	2.9%	100.0%
2. Volume			
	Share of Orders Volume	0.6%	100.0%
	LP incidence (%)	3.2%	1.5%
	Share of Refunded Volume	3.9%	100.0%
3. Cost per order variation if DSS was applied	LPCCC / order	-89.3%	-60.0%
	CB lost / order	Disposable	Disposable
	RTO unrecovered	Disposable	Disposable
		plus DSS cost	
4. Total Cost Variation		-76.5%	6.8%

4.1.4 Customer Tier

Focusing now on direct related features with the consumer, the login form was studied and the background behavior of the client as well.

At Farfetch, clients are divided into tiers, as a way of recognizing and present them in several ways to the ones who buy more, as are the cases of tiers like Private Client, Platinum and Gold. The other ones are Bronze and Silver.

These recognitions and gifts are sometimes used in refund cases, where Farfetch concedes refunds whatever the reason behind is. That's why Platinum, which is the tier for the best clients alongside with Private Clients have the best cost variation when DSS is applied.

Also in this case, it happens that there are tiers like the gold one where the savings are positive, but It wouldn't be viable to sum them up with the Platinum ones, once the variance could affect them negatively in a ordinary or extraordinary situtation.

This way, fraud behavior is prevented as well as Lost Parcel claims when the EDD was still not reached, for instance.

Table 4.4: Case if DSS was applied to all the orders where the customer belongs to the Platinum tier

		FFACCESS-PLATINUM	All tiers
1. Value			
	LPCCC as %GTV	0.5%	0.6%
	Share of GTV	6.6%	100.0%
	Share of LP CCC	5.1%	100.0%
2. Volume			
	Share of Orders Volume	5.3%	100.0%
	LP incidence (%)	0.9%	1.5%
	Share of Refunded Volume	4.0%	100.0%
3. Cost per order variation if DSS was applied	LPCCC / order	-63.9%	-60.0%
	CB lost / order	Disposable	Disposable
	RTO unrecovered	Disposable	Disposable
		plus DSS cost	
4. Total Cost Variation		-8.9%	6.8%

4.1.5 New and Existing Customers

The “New or Existent” type of customers was also studied in order to visualize if, for instance, an existing customer was more prompt to ask for a Lost Parcel than the new, or the reverse. In this feature, nothing was concluded since both of them did not have savings.

Table 4.5: Case if DSS was applied to a new or an existent customer

		Existing	New
1. Value			
	LPCCC as %GTV	0.5%	0.8%
	Share of GTV	77.4%	22.6%
	Share of LP CCC	69.5%	30.5%
2. Volume			
	Share of Orders Volume	68.2%	31.8%
	LP incidence (%)	1.4%	1.8%
	Share of Refunded Volume	64.9%	35.1%
3. Cost per order variation if DSS was applied	LPCCC / order	-64.8%	-61.9%
	CB lost / order	Disposable	Disposable
	RTO unrecovered	Disposable	Disposable
		plus DSS cost	
4. Total Cost Variation		2.3%	13.6%

4.1.6 Guest or Logged Customers

Also, the “Guest or not guest” was seen to figure out, in the same registry as the “new or existent”, and just like before, nothing was concluded since both of them did not have savings with the application of DSS.

Table 4.6: Case if DSS was applied to a guest or not guest customer

		Not Guest	Guest
1. Value			
	LPCCC as %GTV	0.6%	0.7%
	Share of GTV	90.2%	9.8%
	Share of LP CCC	88.9%	11.1%
2. Volume			
	Share of Orders Volume	86.8%	13.2%
	LP incidence (%)	1.5%	1.5%
	Share of Refunded Volume	86.6%	13.4%
3. Cost per order variation if DSS was applied	LPCCC / order	-64.7%	-59.3%
	CB lost / order	Disposable	Disposable
	RTO unrecovered	Disposable	Disposable
		plus DSS cost	
4. Total Cost Variation		3.4%	24.0%

4.1.7 Average Order Value

´ Finally, and the one with best results was the average order value feature, as said before, and also the one that really reverted the way Data Science Team was approaching the problem, once the savings were disruptive and also brought the model disruptive results.

As we can see, If Farfetch applied DSS to all the orders with an AOV from 600 to 3000 $, the savings would be way better than what happens in the other criteria and would benefit a lot from that. However, the problem with simple criteria is just the fact that it’s not improved at its maximum once, as we saw, there were other features with very good results like bags in category and klarna pay later in payment method.

That’s what the next section will focus on by showing that the combination of all those variables can give really good results.

Table 4.7: Case if DSS was applied to all orders with a value from 600-3000$.

		600-3000$	All ranges of AOV
1. Value			
	LPCCC as %GTV	8.5%	0.6%
	Share of GTV	50.9%	100.0%
	Share of LP CCC	58.9%	100.0%
2. Volume			
	Share of Orders Volume	27.6%	100.0%
	LP incidence (%)	22.4%	1.5%
	Share of Refunded Volume	28.8%	100.0%
3. Cost per order variation if DSS was applied	LPCCC / order	-54.8%	-60.0%
	CB lost / order	Disposable	Disposable
	RTO unrecovered	Disposable	Disposable
		plus DSS cost	
4. Total Cost Variation		-17.8%	6.8%

4.2 Model's Output Analysis

The model's first iteration consisted of giving as an output the probability of a certain order to become a lost parcel. After the simple criteria analysis, It was noticed that the target of the model should be orders that are most prompt to be refunded, and not the ones that are most prompt to become lost parcels.

This happened to be accurate since there are orders that are Lost Parcels, and that are not refunded, which don't represent a cost for Farfetch. Since the analysis is made to a net benefit target level, and only the refunded orders represent an involved cost in the total LP CCC, then the model was not being optimized to the right purpose of the analysis.

This way, the model was rebuilt with the help of the Data Science team, in order to optimize the right target which happens to be the Net Benefit formula. For this formula, as described before, only the orders with refund costs, as well as the CB and RTO costs are valued.

For this new model, the characteristics are described below:

Table 4.8: Model Parametrization Characteristics

Model	**CatBoostClassifier**
iterations	1000
early_stopping_rounds	40
loss_function	Log Loss
random_state	10
auto_class_weights	Balanced

Regarding the Training, Validation and Test Sets, these were different from month to month, in order to use the most recent data. For instance, if we wanted to test the model in the month of January 2022, the validation set would be from the start date of the trial period minus 75 days until the start date of the trial period minus 45 days (completing one month time period for the validation set).

The training set, meanwhile, is always from the 1st of June of 2020 until the first day of the month in analysis minus 75 days. For the training set, the period January of 2021 until June of 2021 was not utilized since a problem with a carrier was increasing the refund rate by more than 10% which would bias the model.

The 45 days are the necessary time gap in order to make sure that the Lost Parcel cases are integrated in FARFETCH's databases. This happened for the months of the analysis, which were from the month of July in 2021 until January 2022.

Running the model with the specifications above we obtain some evaluation metrics about which we can infer some conclusions. Starting with the data distribution analysis, we can tell that the target observations are very reduced (we are trying to target only almost 0.6% of the total observations, which represent the ones that are refunded). Due to this, accuracy is not a good metric to use. Consequently, looking at precision and recall which measure the model's capacity of predicting correctly the orders that will be refunded:

As it should be expected, given this data distribution, the precision for a threshold of 0.5 is very low. Nonetheless, to the same threshold, the recall value rounds the 73%, which signifies that ¾ of the total Refunded orders by utilizing 0.5 as threshold.

By utilizing the parameter auto_class_weight="Balanced", the model is adjusted during the training to avoid that the class unbalancing has a strong impact. If this parameter was not utilized, the model would classify the great majority of the samples as Class 0- once this is the class with 99% of the observations, and the accuracy would almost be, if not, 99%.

Nevertheless, this accuracy value does not mean anything to our goal.

Using the parameter, the model makes predictions for the class1, which represent the refunded orders, making the accuracy decrease to 33%, reflecting a great existent class unbalancing but allows the model to have a better precision once It's making more predictions for the class 1.

Table 4.9: Evaluation Metrics Results.

Evaluation Metrics	**Value**
LogLoss	0.68
Precision	1.1%
Recall	73%
Accuracy	33%

However, the best evaluation metric for this problem is our goal: the net benefit savings. And in fact, although not comparing models, we can compare the model to the simple criteria analysis, and see that in terms of net benefit savings, the model is the best option once it saves more net benefit than the other options.

The used threshold was chosen by doing the same analysis as the simple criteria analyisis. This analysis consisted in applying DSS to orders with a probability above a certain threshold.

In this case, the best results were the ones where the probability of becoming a lost parcel claim and a refunded order was above 0.53 until the end (0.59).

The results are in the table below:

Table 4.10: Case if DSS was applied to all orders with a probability above 0.53.

		0.53-0.59	**All Probability ranges**
1. Value			
	LPCCC as %GTV	0.9%	0.6%
	Share of GTV	24.1%	100.0%
	Share of LP CCC	37.6%	100.0%
2. Volume			
	Share of Orders Volume	24.5%	100.0%
	LP incidence (%)	2.2%	1.5%
	Share of Refunded Volume	37.2%	100.0%
3. Cost per order variation if DSS was applied	LPCCC / order	-76.4%	-60.0%
	CB lost / order	Disposable	Disposable
	RTO unrecovered	Disposable	Disposable
		plus DSS cost	
4. Total Cost Variation		-27.5%	6.8%

Another important output of the model and that launches another comparison study is the features importance for the propensity of being refunded orders.

Firstly, and to describe the features utilized in the model, It is presented a brief description of each of them, as well as their name in table 4.11.

The results of this model gave importances to the different features following a Gini Index, presented in the Appendix A. The comparison between the results obtained in the model and in the simple criteria will also be discussed next chapter.

Table 4.11: Model's Features Description

Feature Name	**Description**
carrier_rolling_360_day_lost_parcel_rate	Lost Parcel rate of the order carrier in the 360 days before the order date
carrier	The carrier company that dealt with the order
total_price	Total Price of the Order
customer_zipcode_cleaned	Zip Code of the shipment address of the order
brand_rolling_360_day_lost_parcel_rate	Lost Parcel rate of the order's brand in the 360 days before the order date
card_type_provider_cleaned	Card Type Provider of the order
customertier	The order's customer tier
quantity	The quantity of articles the order had
category_rolling_360_day_lost_parcel_rate	Lost Parcel rate of the order's cateogry in the 360 days before the order date
ip_rolling_lost_parcel_rate	The IP address the customer used while making the order
user_rolling_lost_parcel_rate	Lost Parcel rate of the order's user in the 360 days before the order date
time_since_last_order	Time between the order in study and the last order of the same customer
lost_parcel_attempt	Whether the order's customer already tried to make a Lost Parcel Claim
is_vip	Whether order's customer is VIP or not
has_previous_refund	Whether the order's customer has had a previous refund
isnewcustomer	Whether the order's customer is new or existent.

Chapter 5

Results Discussion

This chapter is focused on the interpretation of the obtained results specifically on the simple criteria analysis results and the model's output results. As explained in chapter 3, the best results of each criteria were presented and will now be compared.

The aim is to opt for one of the criteria or the model's output as a flag in order to put into practice the chosen initiative.

The process putting into practice will also be described in order to understand what changes would the chosen criteria had to pass through.

5.1 Discussion

The presented results in chapter 3 are in a rate format because of the confidentiality terms required by Farfetch. Due to that, the Costs variation between the As-is case (without DSS) and the case where DSS is applied cannot be compared as absolute metrics. In order to present a viable comparison of the obtained results, It is necessary to explain each case and to order the different criteria by absolute net benefit savings.

The best criteria was the model's output with a total cost variation of -27.5%. Although It's not the criteria with the most positive cost variation, It is the one that brings more Net benefit savings (in absolute values), since it also has a big share of orders volume with 24.5%. As we can see, the 0.53-0-59 catches almost 38% of the total LP CCC, and it only represents 24.1% in terms of GTV delivered, which means that by targeting 24% of the total GTV, It is possible to target 38% of the total LP CCC.

This also means that by applying the DSS to orders that have a probability of becoming refunded orders above 0.53, the total LP CCC for these orders decreases 76.4%. By adding the cost of DSS, the Costs with LP CCC decrease 27.5% relatively to the cost wihtout DSS.

Regarding the Average Order Value criteria, by targeting 50.9% of the total GTV, It's possible to catch almost 59% of the total LP CCC. In fact, with the AOV criteria It would be possible to target more LP CCC than in the Model's output case. However, It would also be needed to apply DSS to more orders, that would not be a Refunded Order case.

The average order value was an important feature to analyse since It was the turning point of this project, due to the fact that It allowed to understand that the pain point of it were not the Lost Parcel cases, but the refunded orders cases that contributed to the increase of the total LP CCC. With that information, the model's target was not anymore the orders that were most prompt to become lost parcel cases, but the ones that were most prompt to become refunded order cases.

Looking now to the category criteria, we can see that in case of applying DSS to the bags category, and by targeting 15.2% of the total GTV, 13.6% of the total LP CCC would be caught. With this, the LP CCC would decrease by 66%, and by adding the cost of DSS to the 7.1% of the total orders, the total cost would decrease by 30.9% relatively to the cost without DSS. Although it decreases more relatively when compared to the model's output and to the AOV criteria, It is not better, once the absolute difference of the cost without DSS and the cost with DSS is lower. ´ Glendale, in the category of city has 0.77% of the total GTV delivered. However, it is possible to catch almost 3% of the total LP CCC.

By applying DSS to the orders that have a shipping address to Glendale, It was possible to reduce the total LP CCC cost by 89.3%, and with the cost of DSS applied to the 0.7% of the total orders this city represents, the total costs would decrease 76.5%, which is one of the best decreases of all, but in absolute terms is only the fifth in the list.

Regarding the Klarna Pay Later type of payment, It's possible to catch up to 19.4% of the total LP CCC, by applying DSS to 9.3% of the total GTV delivered. This allows the LP CCC to decrease by 66.2%, which by adding the DSS cost to he 12.5% of the total orders this type of payment represents, gives a total decrease of 17% of the total cost relatively to the cost without DSS.

The last criteria that brings net benefit savings is the customer tier one with the FFAccess-Platinum tier which represents 5.2% of the total GTV delivered, and targets 5.1% of the total LP CCC. By applying DSS to this tier, the LP CCC would decresase almost 64% but the appliance of the DSS cost only makes this variation decrease 8.9%.

Passing now to the features in which didn't bring any positive impact of applying DSS: the New/Existing Customers and the Guest/Not Guest type of Login. Both criteria had relevant decreases on the total LP CCC, but with the appliance of the DSS cost, It would not be viable to do it.

Regarding the feature importance as a model's output, It can also be inferred that It kept up with the simple criteria analysis output. On the one hand, the model was able to verify and prove the subsequent analysis of the carriers since, It confirms that the carrier is the most important feature regarding the lost parcel rate and the refunded orders rate.

In this case, It proves that our starting point of the DSS implementation analysis is actually true: The carrier is one of the most influent features on the refund rate, and the mitigation actions should be applied to it in order to decrease the refund rate of the UPS orders.

On the other hand, It also proves the fact that several features like the total price (or as in the simple criteria case: the average order value), the zip code, a more detailed version of the feature city, the brand, the category, and the customer tier, are in fact important and impactful features in the lost parcel claims rate, refund rates, and consequently, net benefit savings.

Finally, and as seen in the simple criteria analysis, with the features importances, it is possible to infer and prove, once again that the fact that the customer is new or It has a previous order, and If he is registered in the platform or not are not relevant features when the goal is to understand to what orders should the DSS be applied.

5.2 DSS initiative Implementation

Based on the last results discussion and the choice of the criteria that brings better absolute results in terms of net benefit savings, we can move on to the explanation phase of the implementation process of the DSS initiative.

Although there are several criteria that would be beneficial if DSS was applied, not all could be implemented simultaneously since some could conflict with others, and that would decrease the performance of the analysis made.

The model's output was the one with better absolute results and that would bring more net benefit savings to Farfetch, if DSS was applied to orders with a probability of becoming Refunded orders of above 0.53, as seen in chapter 3.

Besides, this model joins the different features in a way that they don't conflict with each other, and so by opting for the model's output, we're also opting for an attempt of doing a perfect combination of the features that make the order more prompt to become a refunded order.

For the implementation of this measure, there are always 3 important steps: the first one is to prepare a business case to present to all the stakeholders, which in this case were the Business Operations team, the Data Science team, the product team and the delivery team, with the goal of persuading the product team put it on the road map.

This would trigger the Delivery team to negotiate the fee with UPS and finally the tech team to develop the solution through which the “needs direct signature flag” can be added order by order, and in this case, based on the probability of each order to become a refunded order. For instance, right after the customer makes the order, the model would predict the probability of that order to become a refunded one, and if the probability is 0.53 or more, soon after would be flagged with "needs direct signature" .

The tech team would also work on integrating this solution with UPS once they would need get communicated on their side order by order on whether Farfetch wants direct signature or not.

As usual in Farfetch, for anything to be applied, this initiative would firstly have to pass by a testing period, to see if the model is flagging the way it should, if UPS was proceeding with the service as agreed, and also, if the rate of refunded orders was in fact decreasing as predicted.

5.3 Mitigation Plan

With the DSS initiative ready to implement, a review of the whole plan to understand if there’s still improvement margin and if all the proposed goals in the beggining of the projects were being tackled.

First of all, It is important to know if the already implemented strategies are working and then to see if the newly created initiative, the DSS case, is going to be well positioned in the whole plan.

The customer journey starts when the order is checkedout on the marketplace. In here, the parametrization of payment refusals to decrease the fraud levels, and consequently, the lost parcel claims that happened to be fraud attempts is the main character. Now, the DSS strategy also starts right from the beggining with the model giving a probability to the order and flagging it in case of being higher than 0.53. Right after, and when the order is on its way until It’s delivered, proactive communications strategy will continue to happen and It will be important to manage the customer’s expectations when the order is delayed or it doesn’t move for days. By doing this, Farfetch will be increasing responsiveness from carriers, from customers and finally, to give option of waiting for the parcel, before the customer asks for a Lost Parcel claim, unnecessarily.

In the end of the order cycle, DSS will act, making the LP claims that lead to refunded orders decrease by 27.8% and also gives Farfetch more advantage to negotiate possible refund cases where the customer signs the receipt proof.

This way, the order cycle has mitigation initiatives working all the way. Regarding the ones that were already implemented, It is possible to say that they had a very good impact in the lost parcel claims rate and in the refunded rate, onde It decreased the peak of these metrics to the normal rates, but still not enough, since one of the greatest improvement opportunities was still not being tackled.

With the implementation of the DSS, all of the cycle components would be tackled, bringing several benefits financially and to the customer satisfaction index.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

This project's main goal was to create a mitigation plan to the lost parcels problematic in the US market. With this, several challenges came up: the building of the model that identified orders most prompt to become refunded lost parcels, the identification of the triggers and mitigation actions to be activated in the plan, and also to ensure the technical and economic viability of implementing the initiatives.

The first step was first to assess the LP root causes along with starting the business case preparation. The subsequent exploratory analysis quantified the LPs that fell under the main root causes, helping to understand what could be the most important features to look at while making the further analysis on the net benefit. After that, understanding the outcomes on LP from Fraud and DS teams to position the team in the right place to communicate and interact with both teams.

With this, It is inferred that most of the LP claims happen when the order is supposedly delivered representing 71% of the total Lost Parcel claims. It is also inferred that the most important and impactful features on the net benefit analysis were the category one with a decrease of 31% in total cost variation when DSS is applied to bags, the payment method when the DSS is applied to Klarna Pay Later decreasing 17% in the total costs and Glendale in the feature city also decreasing 76.5%. Ultimately, the AOV feature which was one of the core discoveries for the project was the feature with more savings, with a decrease of 17.8% and that was a really good help to understand that the main focus of the problem shouldn't be the Lost Parcel incidence that any criteria catches, but the sum of all refunded values, or LP CCC.

This way, the analysis of the model's output, the probabilities, was another important inference, since It's the one complaining with all the simple criteria and that improves the net benefit savings the most, with a decrease of 27.5% of the total costs.

6.2 Future Work

Along the development of this project, some limitations of the proposed approach were identified, and are interesting to be tackled in further investigations.

Firstly, some of the features used in the simple criteria analysis and in the model were related to articles of an order and not to the order itself, which makes the results not as accurate as if the analysis was made to the article level. However, the comparison of the net benefit savings would not be as accurate as the as-is case.

Another improvement opportunity is the escalation of this project to other markets. Although FedEx has a much lower DSS cost per order and this service is allocated to all orders, making Farfetch benefit from it, there are lots of customers that don't want to sign, or sometimes are not home, and consequently would prefer that the order is left in the nearby coffee shop. Besides that, the chances of returning the order are much higher, once Farfetch'd guidelines together with the carrier guidelines reduce the number of delivery attempts before return in a way that Farfetch could lose customers' satisfaction on average.

This way, by expanding this analysis to other markets, the savings would be even higher, the customers' satisfaction index would increase and the return rate would decrease, which would also increase savings by a good margin.

Bibliography

Amed, I., Berf, A., Balchandani, A., Hedrich, S., Jensen, J. E., Straub, M., Rolkens, F., Young, R., Brown, P., Merle, L. L., Crump, H., and Dargan, A. (2022). The state of fashion 2022.

Atwal, G. and Williams, A. (2009). Luxury brand marketing - the experience is everything! *Journal of Brand Management*, 16.

Batta, M. (2020). Machine learning algorithms - a review. *International Journal of Science and Research (IJ*, 9.

Cabigiosu, A. (2020). An overview of the luxury fashion industry.

Dorogush, A. V. and Gulin, A. (2018). Catboost: Unbiased boosting with categorical features. volume 2018-December.

Dubois, B., Laurent, G., and Czellar, S. (2001). Consumer rapport to luxury : Analyzing complex and ambivalent attitudes. *Les Cahiers de Recherche Groupe HEC*, 33.

Kawasaki, T., Wakashima, H., and Shibasaki, R. (2022). The use of e-commerce and the covid-19 outbreak: A panel data analysis in japan. *Transport Policy*, 115.

M, H. and M.N, S. (2015). A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining Knowledge Management Process*, 5.

Natekin, A. and Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics*, 7.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., and Gulin, A. (2018). Catboost: Unbiased boosting with categorical features. volume 2018-December.

Sandrine Devillard, Holger Harreis, N. L. and Altable, C. S. (2021). Jumpstarting value creation with data and analytics in fashion and luxury.

Schröer, C., Kruse, F., and Gómez, J. M. (2021). A systematic literature review on applying crisp-dm process model. volume 181.

Shah, V., Golmohammadi, M., Obeid, I., and Picone, J. (2021). Objective evaluation metrics for automatic classification of eeg events.

Sheng, T. and Liu, C. (2010). An empirical study on the effect of e-service quality on online customer satisfaction and loyalty. *Nankai Business Review International*, 1.

Yeoman, I. and McMahon-Beattie, U. (2010). The changing meaning of luxury.

Yuan, X., Li, C., Zhao, K., and Xu, X. (2021). The changing patterns of consumers' behavior in china: A comparison during and after the covid-19 pandemic. *International Journal of Environmental Research and Public Health*, 18.

Appendix A

Model's Output

This appendix represents the feature importance as an output of the model created together with the Data Science team, which can be found in Figure A.1.

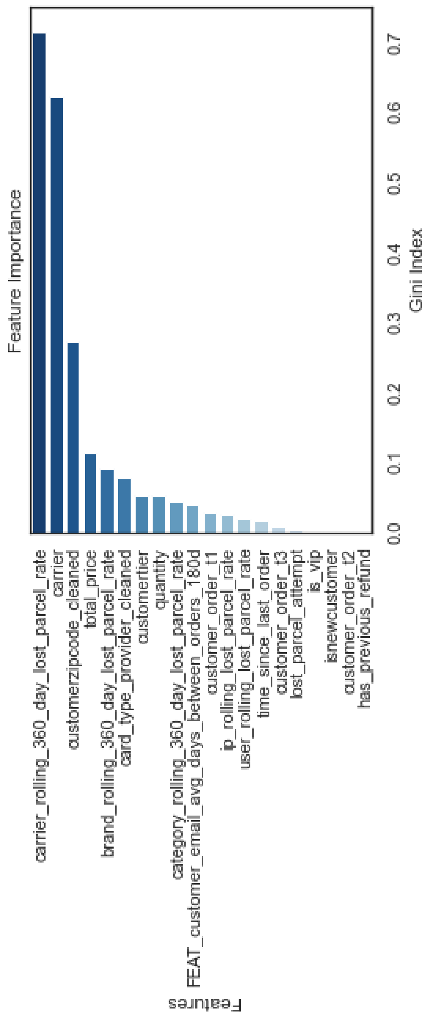


Figure A.1: Feature Importance for the order's probability of becoming a refunded lost parcel.

Appendix B

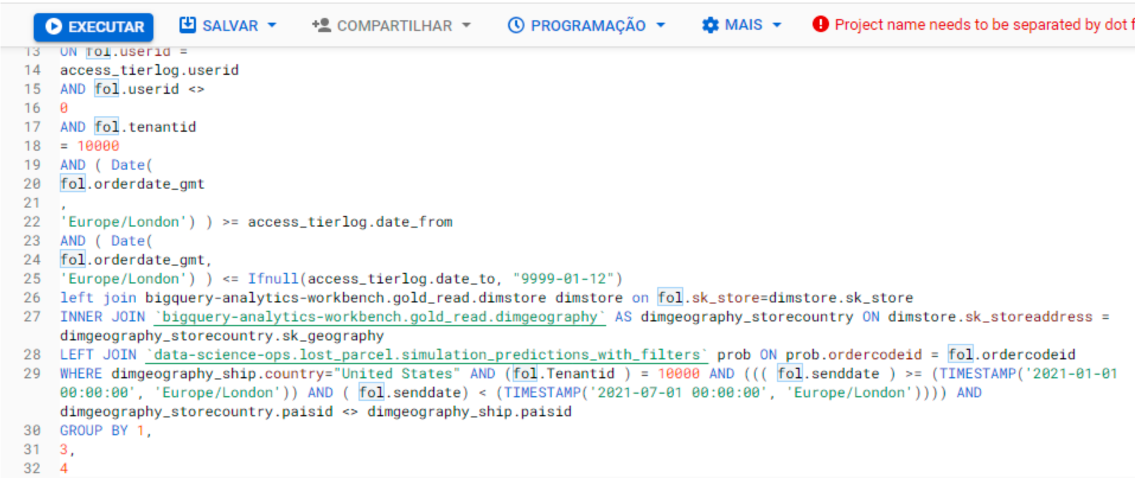
SQL Code for finding international and domestic orders

This appendix aims to give a better idea of the background work of the data mining process. The process of getting into Farfetch's Databases and retrieve the necessary data from them. The figures don't have all of the used codes, since there is no added value on showing them all. They are all very similar, except for some features that require more treatment. It's the only part of the process that is possible to show due to the confidentiality terms of the company.

EXECUTAR   SALVAR   COMPARTILHAR   PROGRAMAÇÃO   MAIS   Project name needs to be separated by d

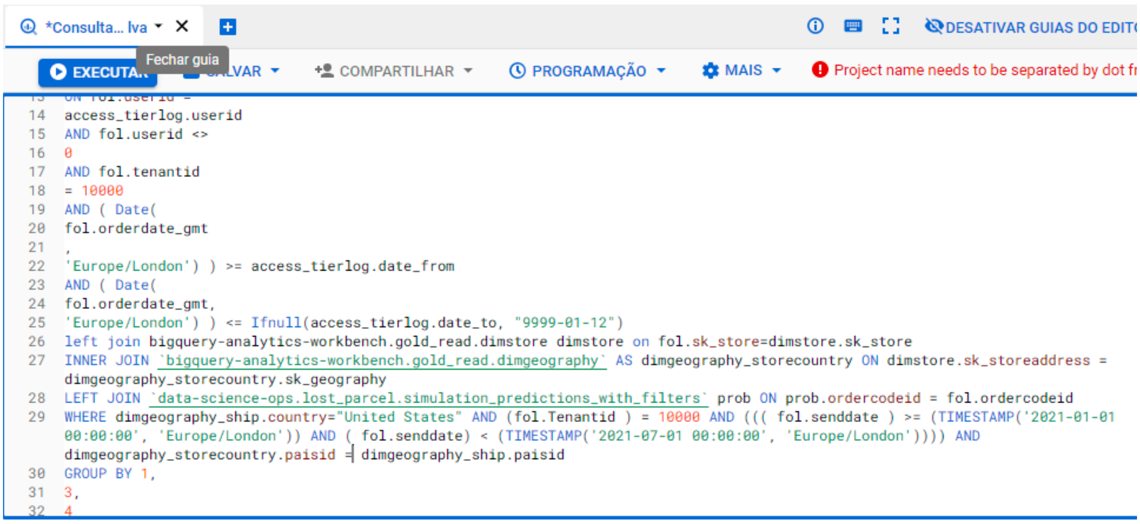
```
SELECT fol.ordercodeid,
sum(valuewithoutship_user_usd
+ valueship_user_usd) AS order_value,
lp_refund_usd_ccc,
probability
FROM `bigquery-analytics-workbench.gold_read.factorderslines` fol
LEFT JOIN `bigquery-analytics-workbench.http://silver_read.lt_final_all_vtest` lp
ON
lp.ordercodeid =fol.ordercodeid
left join bigquery-analytics-workbench.gold_read.dimgeography dimgeography_ship on fol.sk_geographyship=dimgeography_ship.
sk_geography
LEFT JOIN `ff-marketing-analytics.core.access_program_status_log` AS
access_tierlog
ON fol.userid =
access_tierlog.userid
AND fol.userid <>
0
AND fol.tenantid
= 10000
AND ( Date(
fol.orderdate_gmt
```

Figure B.1: First part of the SQL code to retrieve international orders data from Farfetch Databases.




```
ON fol.userid =
access_tierlog.userid
AND fol.userid <>
0
AND fol.tenantid
= 10000
AND ( Date(
fol.orderdate_gmt
,
'Europe/London') ) >= access_tierlog.date_from
AND ( Date(
fol.orderdate_gmt,
'Europe/London') ) <= Ifnull(access_tierlog.date_to, "9999-01-12")
left join bigquery-analytics-workbench.gold_read.dimstore dimstore on fol.sk_store=dimstore.sk_store
INNER JOIN `bigquery-analytics-workbench.gold_read.dimgeography` AS dimgeography_storecountry ON dimstore.sk_storeaddress =
dimgeography_storecountry.sk_geography
LEFT JOIN `data-science-ops.lost_parcel.simulation_predictions_with_filters` prob ON prob.ordercodeid = fol.ordercodeid
WHERE dimgeography_ship.country="United States" AND (fol.Tenantid ) = 10000 AND ((( fol.senddate ) >= (TIMESTAMP('2021-01-01
00:00:00', 'Europe/London')) AND ( fol.senddate) < (TIMESTAMP('2021-07-01 00:00:00', 'Europe/London')))) AND
dimgeography_storecountry.paisid <> dimgeography_ship.paisid
GROUP BY 1,
3,
4
```

Figure B.2: Second part of the SQL code to retrieve international orders data from Farfetch Databases.




```
ON fol.userid =
access_tierlog.userid
AND fol.userid <>
0
AND fol.tenantid
= 10000
AND ( Date(
fol.orderdate_gmt
,
'Europe/London') ) >= access_tierlog.date_from
AND ( Date(
fol.orderdate_gmt,
'Europe/London') ) <= Ifnull(access_tierlog.date_to, "9999-01-12")
left join bigquery-analytics-workbench.gold_read.dimstore dimstore on fol.sk_store=dimstore.sk_store
INNER JOIN `bigquery-analytics-workbench.gold_read.dimgeography` AS dimgeography_storecountry ON dimstore.sk_storeaddress =
dimgeography_storecountry.sk_geography
LEFT JOIN `data-science-ops.lost_parcel.simulation_predictions_with_filters` prob ON prob.ordercodeid = fol.ordercodeid
WHERE dimgeography_ship.country="United States" AND (fol.Tenantid ) = 10000 AND ((( fol.senddate ) >= (TIMESTAMP('2021-01-01
00:00:00', 'Europe/London')) AND ( fol.senddate) < (TIMESTAMP('2021-07-01 00:00:00', 'Europe/London')))) AND
dimgeography_storecountry.paisid = dimgeography_ship.paisid
GROUP BY 1,
3,
4
```

Figure B.3: Second part of the SQL code to retrieve domestic orders data from Farfetch Databases.

Appendix C

Used metric for the simple criteria analysis

In this appendix, It is shown the excel template for the simple criteria analysis. Different metric had to be used in order to understand what were the main contrasts between features and what in fact was an important indicator of a good net benefit savings. The results shown during the whole thesis would be presentend following the template presente in here, but the confidentiality terms didn't allow it.

1. Value
1.1. Total GTV deliv. ($)
1.2. LP CCC ($)
1.3. LP CCC as %GTV
1.4. Share of GTV*
1.5. Share of LPCCC*

2. Volume
2.1. Total Order Volume (units)
2.2. LP Order volume (units)
2.4. Share of Volume*
2.5. %LP
2.6. Share of Refunded Volume*

3. Per order assumptions (based on hist. perf.)
A. Without DSS
3A.1. LP CCC / order ($)
3A.2. Cost of DSS / order ($)
3A.3. CB lost / order ($)**
3A.4. RTO unrecovered / order ($)

B. With DSS
3B.1. LP CCC / order ($)
3B.2. Cost of DSS / order ($)
3B.3. CB lost / order ($)**
3B.4. RTO unrecovered / order ($)

4. Costs per AOV ranges (crp. to the order vol. under each range)
A. Without DSS
LP CCC ($)
DSS ($)
CB lost ($)
RTO unrecovered ($)
TOTAL ($)

B. With DSS
LP CCC ($)
DSS ($)
CB lost ($)
RTO unrecovered ($)
TOTAL ($)

5. Net Benefit ($) (Without DSS - With DSS; i.e. As Is - To Be)
Net Benefit in M USD

Figure C.1: MS Excel Template for the simple criteria analysis