Order Tracking model for enhanced Delivery Experience: a predictive approach

José Pedro da Silva Sá

Master's Dissertation

Supervisor: Prof. Isabel Horta

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Integrated Master in Industrial Engineering and Management

2019-07-01

Resumo

Um domínio fulcral no sucesso de plataformas *e-commerce* prende-se com a experiência providenciada ao consumidor, factor preponderante no grau de satisfação atingido. Sendo a entidade sob análise, a plataforma líder para o mundo online da moda de luxo, tal factor toma ainda proporções maiores. Desta forma, o prjeto descrito na presente dissertação tenciona avaliar e melhorar o atual paradigma da experiência de entrega reccorrendo a métodos analiticos e colocando sempre o consumidor no centro de cada análise e decisão.

De forma a suportar processos que acrescentem valor de forma inequívoca e consensual, a revisão literária de várias temáticas, consideradas relevantes bem como o seu enquadramento organizacional foram realizados. Com foco direccionado para o segmento pós-compra da *costumer journey*, a presente dissertação revela os alicerces da experiência de entrega, bem como métodos para aprimorar cada um deles. Para tal, um estudo estruturado aprofundado efetuado, envolveu os seguintes procedimentos – desenvolvimento de um framework de análise, definição do estado da arte, identificação de pontos de melhoria e por último, análise de benchmarking. Como resultado deste estudo, três linhas de ação foram implementadas com o intuinto de melhorar de forma impactante, a experiência de entrega. Por conseguinte, uma algoritmo predictivo inspirado em machine learning foi desenvolvido de forma a suportar um novo canal de comunicação proativa, capaz de detetar antecipadamente, eventos de atraso em dois pontos da rota de uma determinada parcela. *Naive-Bayes*, *Logistic Regression* e *Extreme Gradient Boosting* forma os algoritmos selecionados para treinar, testar, *tune* e por último, avaliar para cada modelo. De facto, os bons resultados de perfomance alcançados particularmente através do *Extreme Gradient Boosting*, demonstram a elevada capacidade preditiva dos modelos, construidos sob um conjunto de variáveis pré-determinadas e pré-processadas.

Em suma, o conhecimento extraido ao longo do projeto, bem como os fluxos de trabalho implementados contribuiram para aprimorar a experiência do consumidor, maioriatiamente no segmento de entrega da cadeia de valor. Adicionalmente, a presente dissertação serve como prova de que soluções analíticas concevidas recorrendo a grandes volumes de dados, podem mudar o estado da arte, criando novas formas inteligentes de interação, valorizadas pelos consumidores.

Abstract

Customer experience is a pivotal domain for e-commerce platforms success, directly impacting customer satisfaction. Being the company under analysis, the leading global platform for luxury e-tail business, such impact takes even greater proportions. The present dissertations aims to assess and improve current delivery experience paradigm adopting a customer-centric and data-driven approach.

In order to create the foundations for value-adding processes, a literature revision combined with framing the topic in a specific organizational context and reality were context. Focusing on the post-order segment of the customer journey, the present dissertation unveils the building blocks of delivery experience. A structured and in-depth study is performed comprising framework development, current scenario definition, improvement points identification and benchmarking analysis. As an output, three workstreams are established in a quest for the ultimate delivery experience. Going beyond traditional analytical methods, a machine learning-based algorithm was developed to support a proactive communication flow, determining order delay events across two route checkpoints. Naïve-Bayes, Logistic Regression, and Extreme Gradient Boosting were the selected classifiers to train, test, tune and lastly, assess for each model. In fact, good performance results achieved particularly using Extreme Gradiente Boosting classifier, indicate a high predictive ability built over a pre-determined and pre-processed set of variables.

Briefly, valuable insights gathered, and lines of action set up contribute to enhance customer experience on the last-mile delivery segment, an ever-changing supply chain element. In addition, the present dissertation proves that data-driven analytical methods can contribute to drive customer engagement towards new forms of interaction, in a seamless and meaningful way.

Acknowledgments

Throughout not only the present dissertation but also during my academic path, a vast number of people and entities contributed to my success. Five years of constant learning, commitment, and dedication passed by in a quest for something intangible and uncertain, a professional status that I would appreciate in place that reflects values an principles I believe in. Therefore, firstly, I would like to thank Farfetch for the opportunity of materializing my ambitions and reveal an amazing working culture. In addition, a huge thank you not only to *Faculdade de Engenharia da Universidade do Porto* (FEUP) but also to *Linköpings universitet*, for the enormous academic contribute which also shaped me as a human being.

Thereafter, I would like to show gratitude to my parents for being the major enablers of my academic success. Additionally, thanks to my sister and friends for endless support and help. Within Farfetch universe, I would like to show appreciation for the Delivery Operations team due to the amazing integration and for creating every day, a remarkable working environment full of companionship and mutual cooperation. To Pedro, André, Inês, Tomás, Isabelle, Vitor, Wilson, Daniela, Ivo, Rita, Rui, Diana and Hugo, a word of appreciation. I feel obliged to give a special thanks to Inês Carvalho for her unconditional and decisive support for present dissertation success. Additionally, I would like to highlight André Alves for his willingness to help whenever I needed. Lastly but not least, I would like to express my acknowledgments to Isabel Horta due to her successive contributes that helped me improve the academic content of the present dissertation.

Contents

Acronyms and Symbols

3PL	Third-Party Logistics
AWB	Air Waybill
AOV	Average Order Value
AOC	Average Order Contribution
AUC-ROC	Area Under the Receiver Operating Characteristic Curve
BDA	Big Data Analytics
CPO	Contacts Per Order
CNNIC	China Internet Network Information Center
CRISP-DM	Cross-Industry Process for Data Mining
D&T	Duties and Taxes
DDP	Delivered Duty Paid
DAP	Delivered At Place
EDD	Estimated Delivery Date
FN	False Negative
FP	False Positve
GMT	Greenwich Mean Time
IPO	Initial Public Offering
KPI	Key Performance Indicator
ML	Machine Learning
NYSE	New York Stock Exchange
NPS	Net Promoter Score
ODD	On-Demand Delivery
ROC	Receiver Operating Characteristic
RT	Retention Rate
SCM	Supply Chain management
SKU	Stock Keeping Unit
SLA	Service Level Agreement
TN	True Negative
TP	True Positive
XGB	Extreme Gradient Boosting

List of Figures

List of Tables

Chapter 1

Introduction

In an increasingly tech-savvy world, in which customers are constantly being empowered by new and game-changing technologies, building competitive advantages has become harder in retail. E-commerce has been disrupting the deepest foundations of retail and is expected to represent 17.5% of retail sales worldwide by 2021 (Statistica, 2018). The fashion industry is leading this growth with 36% of global retail sales being digital-enabled by 2022 (Meena et al., 2017).

From a business perspective, in an attempt to capture the online retail growth, e-commerce players are thriving to offer a wide range of products and services with meaningful value for customers such as shorter lead times, smaller and more frequent deliveries, pushing, even more, the complexity barrier of logistics. In addition, a platform-based thinking where several layers of value are created and offered to clients has been adopted by leading players resulting in a plethora of value-adding services. Therefore, in a constant quest to retain customers to its ecosystem, Farfetch has been embodying this customer-centric vision on its strategic planning. The business unit "Fashion concierge", based on a search and find app for luxury items or the 90-minutes delivery method reflect the mentioned approach. By fulfilling not only customers' needs and expectations, but also offering innovative services, a firm can boost its awareness and reputation, critical competitive advantages in the fashion industry. Despite that, "while many companies say they are customer-centric, few actually put the customer at the center of the innovation process." (Ramaswamy, 2009b). Include this mindset in organizations while adapting workstreams to it, represents an immense challenge for today's organizations, particularly on e-commerce world, where customer physical interaction is scarce.

1.1 Farfetch

The project under scrutiny was carried out in Farfetch, the leading global technology platform for luxury fashion industry. In 2007, José Neves, founder and current CEO, decided to merge two of his passions, technology and high-end fashion, taking this conservative and digital-skeptic market to the online world. In recent years, the company has been growing rapidly which resulted in its inclusion, through an Initial Public Offering (IPO) process, on the New York Stock

Exchange (NYSE), the world's largest stock exchange by market capitalization, under the ticker symbol "FTCH.". The company sold 44.2 million shares and raise $884 million in its first public investment action. At the date of this report, and using market capitalization as a metric of value, the company worth around 7 billion dollars.

Besides the most recognized business unit and major revenue source, the online marketplace for luxury fashion, the company has been investing in other complementary fields. In May 2015, Farfetch acquired Browns, an iconic British fashion and luxury goods boutique, which allowed not only being close to customers and understand their needs but also, as an important source of supply. In April 2017, it was unveiled the vision for brick-and-mortar retailing and "Store of the future" business unit was launched. By taking advantage of physical retail presence, the aim to this unit is developing technology for the physical environment. A technology-powered retail operating system has been driving the reinvention of consumer experience, through online and offline integrations. Moving forward, on October 2017, "Fashion Concierge" was acquired. This entity runs an e-commerce solution sourcing luxury items on behalf of Farfetch private clients, worldwide. Finally, "Black and White" unit consists of a white-label e-commerce solution for luxury fashion retailers built on top of Farfetch platform, providing the same capabilities and scale as the marketplace.

Through its extensive network, the company connects more than 1500 boutiques and luxury brands to customers scattered over near 200 countries. The supply side is more concentrated in historical fashion-enthusiast countries, with Italy playing a leading role, representing the origin of around 48% of all items by volume. Throughout the years, a strategy to diversify supply geographically has been adopted offering a truly global shopping experience. On the demand side, the most prominent markets are the following - the United States of America, Hong Kong, China, Japan, Russia, Australia, and the United Kingdom. Looking at the future, Amed et al. (2019) revealed on its latest report that China is expected to overtake the US as the largest fashion market in the world in 2019. In parallel, D'arpizio and Levato (2018) estimates that, by 2025, Chinese customers will account for 45% of personal luxury goods market which explains the growing importance of this market for e-tail players. Due to geopolitical reasons, it becomes difficult to penetrate in the mentioned market with uncountable western e-commerce companies struggling to capture its full potential. In addition, legislative actions and technologies enforced by the Chinese government to regulate internet domestically, commonly known as "The Great Firewall of China", leads to a mandatory local data-storage solution. As a consequence, Chinese and non-Chinese internal data flows are heterogeneous and hence, data-processing procedures differ which ultimately affects customer experience.

Considering another complexity layer, Farfetch's business model is marketplace-based, which implies that all items displayed online are owned and stored by upstream partners (boutiques and brands) with some specific exceptions. With this no stock policy, inventory-related holding costs are virtually nil, which enables the company to present a wide range of items to customers. This broader assortment of clothing encompasses one remarkable competitive advantage which differentiates Farfetch from all its direct competition. In contrast, by having less control over supply compared to an inventory-based model and a low number of SKU's, stockout exposure increases. Furthermore, with multiple origins and lack of visibility and control over the fulfillment process, it becomes hard and challenging to provide a consistent and coherent customer experience in what regards lead time and packaging. Logistics complexity also arises due to a tremendous number of possible routes and low transportation volume. To ensure that the transportation flow occurs smoothly, Farfetch works closely with its 3PL (Third-party Logistics) providers. Having multiple partners, the company is capable to offer several shipping methods worldwide. At the present time, the company offers the following methods - "Express", "Standard", "Same-day" (that could evolve to "Next-day" delivery), "F90" (deliveries in 90 minutes) and Click-and-Collect. The availability of the first two depend on the country of origin and destination and, represent logically, a vast majority of the order volume. A customer can opt for "F90" option in selected areas of 10 cities - New York, Berlin, London, Madrid, Milan, Paris, São Paulo, Dubai, Hong Kong and Tokyo. In the mentioned cities apart from San Paulo, and other 10 (Chicago, Dallas, Dayton, Los Angeles, Miami, Seattle, Barcelona, Manchester, Porto, Rome and Shanghai), "Same-day" delivery option is available. On this last two options, supply is limited to specific boutiques and customers can only order from a predefined coverage area. Lastly, Click-and-Collect option consists of a pick-up location from a customer perspective, available in a number of boutiques and an additional 20 sites in London enabled by a 3PL provider.

Due to the luxurious nature of business, Farfetch's customer base has a more demanding attitude than average consumer which requires a tailored interaction and responsive support. Therefore, to understand each customer segment traits and aligned with its globalization vision, company offices and workforce are spread around the globe in 13 cities, being London its headquarters. In what regards organizational structure, its core can be considered functional with some departments adopting a more matrix approach towards resources allocation to projects. Therefore, a clear chain of command as well as a considerable degree of specialization is present. Specifically, the project under analysis was carried out in Delivery Operations team which belongs to Logistics Department, an element of Operations division. The team aspires to streamline and optimize all processes involved since the order is shipped from a specific partners' stock point until delivered to the end-customer and the subsequent inverse logistics. Being specialized in the mentioned supply chain segment, a delivery-related customer experience project emerged, and it is the basis for the present dissertation.

1.2 Motivation and Goals

Amaze customers in all ways possible has always been a priority in Farfetch's business strategy. Several initiatives have been contributing to it across the customer journey since the moment of searching and finding an item until the refund resulted from a return process. In what concerns delivery experience, enhancements were achieved mainly on the pre-order segment such as providing an accurate estimated delivery date at checkout. By taking a problem-solving approach, the aim of the present project is to revamp post-order delivery-related customer experience, particularly, from the moment the item is shipped from a specific partner's stock-point until the shipping address indicated by the customer. With that in mind, the present dissertation has two complementary but individual purposes.

1. Present a holistic overview of post-order delivery experience drivers with the objective of clearly identify improvement points and deploy resources to materialize the proposed solutions;

2. Develop a data-driven analytical tool which addresses a specific area of opportunity and delivers real value to customers. This value proposition lies on a predictive model capable of supporting a proactive communication trigger, notifying customers if a certain order is delayed.

To achieve the desired goals in the most efficient and effective way possible, it is imperative to structure workload in logical steps, being the intended purpose of the next section.

1.3 Methodology

A thorough planning ensures that projects deliver real value against the business opportunity. Dvir et al. (2003) suggests project success is "positively correlated with the investment in requirements' denition and development of technical specications". For this reason, clear stages of the project were initially mapped as well as achievable milestones. Therefore, the present project can be segregated in six stages.

1. Diagnosis and theoretical framing
   Firstly, a high-level analysis of customer journey was performed trying to capture the main drivers of delivery experience. Afterwards, in order to frame concrete value-adding actions, a framework to assess post-order delivery experience was developed and company's vision was clearly mapped. Therefore, the foundations for future improvements were established, allowing not only the alignment of internal stakeholders involved but also the establishment of a straightforward vision.

2. Areas of opportunity identification
   To narrow the scope of the project in value-added activities, opportunities to delivery-experience enhancements were identified and prioritized. Moreover, the available datapoints were collected and analyzed in order to understand intrinsic relationships and impact among delivery experience elements.

3. Benchmarking analysis
   Afterwards, capturing relevant value propositions available at the market was possible by conducting a benchmarking analysis to delivery experience-focused providers. Pros and cons of external interventions and internal developments were weighted.

4. Workstreams definition

   After understanding internal and external potential and having in consideration required development requirements, workflows were established in parallel.

5. Development of a machine-learning-based customer communication solution

   As a basis for one area of opportunity previously identified and using carriers tracking data as the main data point, an analytical method was developed to support a communication-focused solution. By determining whether an order being is delayed across its journey, the model aims to support the creation of a new forward-looking communication flow, notifying consumers proactively and hence, positively impact customer experience and ultimately, satisfaction and retention.

6. Impact analysis

   On the customer side, the repercussion of the present project on customer satisfaction was studied. On the company point of view, potential savings enabled as well as the fitness degree on overall customer strategy was investigated.

1.4 Structure

The present dissertation was organized in order to clearly depict the evolution of the project under scrutiny as well as the several tasks of different nature conducted. In chapter 2, a literature review covering the most relevant topics, namely - fashion, e-commerce, supply chain management and logistics, customer experience, and big data analytics principles – is presented. Chapter 3 comprises the framing of the subject at Farfetch. Order flow, delivery experience current scenario and datapoints are presented, providing detailed insights that clearly support the pertinence of this project. Chapter 4 consists of a structured analysis comprising framework definition, opportunities identification, benchmarking analysis and lastly, workstream establishment. One of the workflows revealed is explored in chapter 5, with the development of a machine learning-based algorithm, thoroughly described. Lastly, Chapter 6 represents a conclusive reflection about results and future work that would revamp the value proposition of the analysis and algorithm developed.

Chapter 2

State of the Art

2.1 E-commerce meets Luxury Fashion

Shopping online has been changing society stakeholders' habits in a profound and decisive way. Studies project that worldwide retail commerce sales to hit \$4.9 trillion in 2021, which represents a growth rate of 256% since 2014, revealing a steady upward trend with no signs of decline (Statistica, 2018). Essentially, electronic commerce has been growing due to three main reasons - easy and unceasing access to personal computers, the expansion of broadband connections and the continuous development of the Internet (Du et al., 2010). The ubiquity dimension of an online channel allows the liberation of markets from being restricted to a physical space (Laudon and Traver, 2018). By extending time and space traditional boundaries, consumers have more power and control over their purchases. Information such as price and service details for specific products in different online platforms are easily accessible which makes the shopping process much more transparent (Clark and Johnston, 2006). Besides that, Falk and Hagsten (2015) suggests a strong correlation between labor productivity and e-sales activities. For sellers, it creates a new path to reach a much wider audience while reducing intermediaries throughout the supply chain. Recent literature has been revealing the power of decentralizing e-commerce processes in a peer-to-peer way. By establishing trust between two entities and removing intermediaries between them, efficiency could increase while diminishing overall cost (Alabi, 2017).

As a game-changing form of e-commerce, Mobile commerce (shopping online using mobile devices, also referred as m-commerce) is poised to become an even more profound economic and social phenomenon (Kourouthanassis and Giaglis, 2012). Chong (2013) showed that perceived usefulness, ease of use, enjoyment, and cost as well as satisfaction and trust have a significant inuence on consumers' m-commerce continuance intentions. Concerning intrinsic motivation for shopping on mobile devices, Ono et al. (2012) developed a model on six motivation predictors - value, role, adventure, social, gratification, and idea motivation. Kourouthanassis and Giaglis (2012) envisions m-commerce as a constantly changing phenomenon, moving towards a more social-minded reality, in what is called social commerce (s-commerce). S-commerce represents a subset of e-commerce that uses social networking sites for business-oriented interactions and user contributions to facilitate the online buying and selling of various products and services (Kim and Park, 2013). Xiang et al. (2016) present evidences that consumers on these platforms are prone to impulse buying behavior. Due to this impulsive nature, not only represents a lucrative shopping touchpoint but also an effective way to steer demand towards a more predictable reality. All these ways to reach end-consumer are being explored in luxury, a historically digital-skeptic market that it is adopting online channels in a fast pace. According to Ko et al. (2019), the luxury status of a certain brand is explained by five elements - high quality, authentic value, prestigious image within the market built on qualities such as craftsmanship, be worthy of commanding a premium price and be capable of inspiring a deep connection with the consumer. Olivier Abtan (2014) explains the latest shifting in luxury, in which consumers are moving towards a quest for immersive experiences instead of owning an actual luxury product and appreciate its intrinsic social prestige. Regarding common popularity effect, luxury fashion products can benet from popularity cues due to the sudden soaring of perceived quality (Yu et al., 2018). Naturally, e-commerce is a key channel to quickly capitalize on the desire for a specific product.

2.2 Supply chain management and Logistics

Today's customers have the ability to purchase whatever, whenever, wherever they desire and at the price they want, putting them in total control of the market (Kayikci, 2019). To achieve meaningful outcomes, supply chain strategy must be linked vertically to customers and horizontally to other parts of the enterprise (Jacobs and Chase, 2017). Exhibit 2.1. depicts all the required capabilities and flows to develop and manage an effective supply chain, capable of delivering five crucial competitive-enhancing factors - price, speed, dependability, flexibility, and quality. In addition, Billings and Ho (2017) declares innovation and sustainability as additional critical categories for supply chain long-term success. Concerning its integrating role, pursue a truly "Triple-A" paradigm (Lee, 2004) is crucial. That is, fostering agility, creating the right alignment with every stakeholder and adjusting rapidly to changes could mean a valuable boost in visibility and control across the value chain. Additionally, it helps companies to tackle the increasing variability throughout the supply chain, widely known as the "Bull Whip effect". On fashion apparel industry, challenges include tremendous product variety and very short product life cycles. In such environment, it is highly important to effectively manage trade-offs between variety benefits and inventory and/or other costs arising from variety increase (Mehrjoo and Pasek, 2014).

The growing adoption of online channels as well as emerging digital subchannels such as social media forces retail business models to change (Verhoef et al., 2015). Providing customer, seamless and consistent experience throughout different channels and devices is part of an omnichannel retailing approach. As a practical example, in the past, retailers built two types of distribution centers, one to manage store fulfillment, whereas another to manage purely e-commerce. Omni-channel distribution centers combine both e-commerce and traditional store distribution channels (Kayikci, 2019). Moving downstream, large e-commerce players have identified last-mile services as a key differentiator. Nearly 25% of consumers are willing to pay significantly more for delivery options such as same-day or instant delivery (Joerss et al., 2016). Considering this, Joerss et al. (2016) proposes three consumer delivery models that are likely to dominate the last-mile in the future depending on locale density and delivery options - autonomous ground vehicles with parcel lockers, drones, and bike couriers. With geographical barriers becoming much less evident on e-commerce, customers adopt a more global shopping behavior. According to Eurostat, 26% of online EU shoppers bought or ordered goods or services from sellers in non-EU countries. Therefore, for online players, it is critical to ensure that products smoothly flow between countries and customs-related issues are considered and predicted.

From an operational perspective, vital supply chain elements were mentioned, and the present chapter represents the basis for in-depth exploration of specific segments. Since the present dissertation has customer experience enhancements has its primal goal, it is imperative to fully capture customer point of view.

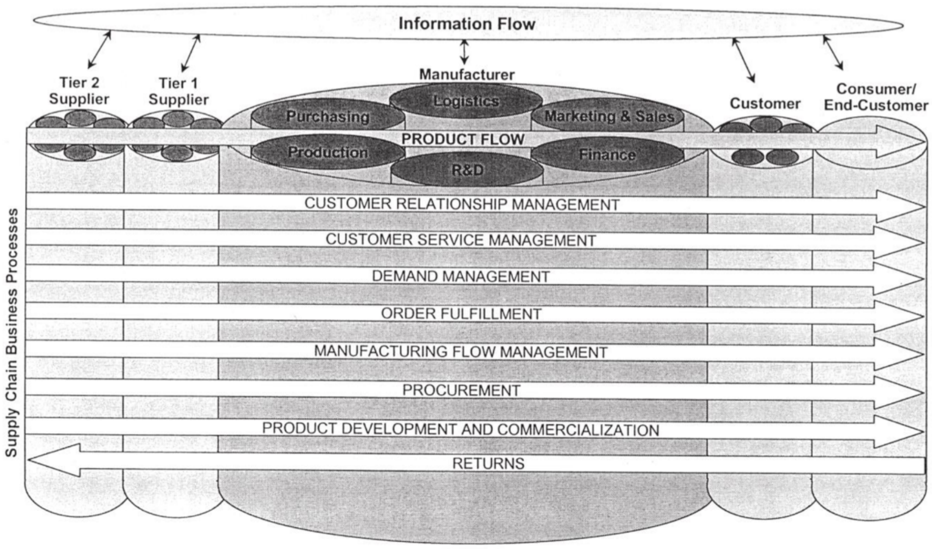

Figure 2.1: Supply chain management (Lambert et al., 1998)

2.3 Customer experience

An increasing customer-centric approach has been adopted by companies in trying to build meaningful competitive advantages. Broadly, customer experience can be defined as a "multi-dimensional construct focusing on a customer's cognitive, emotional, behavioral, sensorial, and social responses to a firm's offerings during the customer's entire purchase journey" (Lemon and Verhoef, 2016). With the intention to frame a customer experience analysis, it is important to clearly identify customer roles and their contributions. As an input to a transformation process, A. Lengnick-Hall (1996) argues that customers can be treated, not only as a resource, a source of valuable information and knowledge but also as a co-producer when it participates in the value-creating conversion activities. In addition, Ramaswamy (2009a) explains that co-creation of value with clients is possible by aligning "outside-in customer-to-employee experience with inside-out employee-to-customer experience". By adopting a downstream view, A. Lengnick-Hall (1996) states that the buyer role is influenced by perceived quality, determined by expectations and actual experience, as well as relationship with seller. Treated as users, customers create two important outcomes: determining customer satisfaction and develop a relationship with the producer. In the long run, the alternation between user and buyer is what defines customer perceptions and expectations. When picturing consumer as a product, figuring out in advance, change of behavior as well as influence and drive how they will respond to shopping stimulus are the ultimate goals of any system output.

In what concerns the categorization of quality-enhanced attributes, the theory of the attractive quality explains the relationship between objective performance and customer satisfaction with an attribute (Witell et al., 2013). As depicted in figure 2.2., must-be quality represents a certain requirement that customers expect to be fulfilled, hence its absence leads to dissatisfaction. Contrarily, an attractive quality provides satisfaction when achieved fully, but do not cause dissatisfaction when not fulfilled. Additionally, the achievement of a one-dimensional quality could have a positive and negative impact on satisfaction. To revamp its competitive dimension, a company should thrive to fulfill must-be attributes, while trying to achieve the one-dimensional ones provided by the market leader, as well as, include a certain number of attractive attributes enough to delight customers and improve awareness and reputation. Furthermore, an indifferent quality represents attributes whose impact on satisfaction are insignificant and could be, in theory, suppressed. When a high degree of achievement results in dissatisfaction, a reverse quality is observed. Witell et al. (2013) suggests that successful quality attributes are assumed to follow a life cycle in which a quality attribute starts by being "indifferent", then moves to 'attractive" through to "one-dimensional" and finally, end up as 'must-be". Despite that, across literature, alternative life cycles are proposed in an attempt to identify patterns for specific regions or markets. For instance, Löfgren et al. (2011) identified the reverse movement, that is, taking "a step backwards in the life cycle of successful quality attributes through, for example, a change in design".

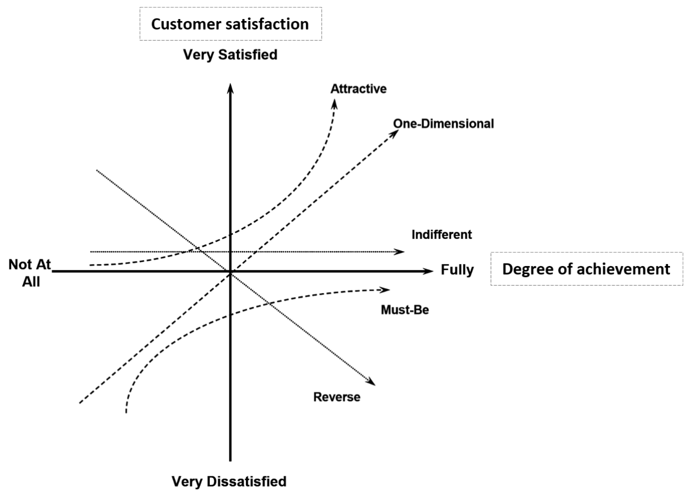

Figure 2.2: An overview of the theory of attractive quality (Witell et al., 2013)

Regarding online retail, Molla and Licker (2001) pictures e-commerce customer satisfaction as a "dependent variable to e-commerce success". In particular, Lin et al. (2016) argues that e-service quality and logistics service quality are unequivocally linked to customer satisfaction. Widely acknowledged, one of the pillars of competitiveness and customer retention is related to how responsive a company is in terms of time-to-ship offerings. Rangel (2018) detected a drastic drop in satisfaction when items arrive after the time of intended use. Therefore, the author states that supply chain capabilities should be allocated to avoid late arrival instead of focusing only on delivering as soon as possible. Despite that, e-commerce intense efforts towards increasing delivery speed are evident with Amazon, the biggest e-commerce player by market capitalization, being able already to ship to 72% of United States population within a day (Kim, 2019). It is worth to point out the underlying risk perception of an online sale and its effects on consumer decision-making process. Rangel (2018) states that luxury consumers are more concerned with risk awareness for virtual purchase processes in comparison to physical ones. With that in mind, Lee (2007) shows that "consumer risk averseness is negatively related to consumer attitude". In addition, the author pictures risk awareness as a multidimensional topic linked to - performance, financial, safety, social, psychological and time/opportunity dimensions. The psychological and social risk, due to the inherent exclusivity and high social status present on high-end fashion, are naturally mitigated. Highly important in online platforms, a visible transaction guarantee as well as offering all price-related information needed may decrease or even avoid the financial risk perception dimension for the luxury e-consumer (Broillet et al., 2019). From a consumer perspective, Kini and Choobineh (1998) argues that trust in an online system is influenced by characteristics of individuals making the transaction, the online system itself, the task for which the system is being used and lastly, information environment.

Concerning the fashion segment, changing consumer preferences force companies to adapt and reinvent quickly. A recent report (Amed et al., 2019) suggests "mobile obsessed, platforms first, start-up thinking" as new trends. Differences in demographic cohorts' habits are evident with Millennials being almost twice as likely as baby boomers to prefer up-and-coming designers. In this fast-changing industry, the ability to capture preferences' volatility and hence, tailor customer experience enhancements initiatives can result in fruitful outcomes. In a contingency perspective, understanding how to recover from service failures is crucial to ensure that customer never leaves the company ecosystem. (Singh and Crisafulli, 2016) highlights that perceived justice conveyed by online service recovery strategies restores customer satisfaction. A common service failure when shopping online is delivery delays. In fact, in November 2017, 41% of United States online shoppers who had experienced a late or non-delivery of an order experienced it on e-tail business, which reflects the importance of this topic to the industry. For this specific "pain point", (Chang and Wang, 2012) identified compensation, response speed, apologies, and contact channels as key attributes to boost service recovery and customer retention.

Summing up, multiple dimensions were covered always by adopting a customer point of view. Presenting a holistic overview while grasping key customer experience elements on both brick-and-mortar and e-commerce ecosystem, was always a concern.

2.4 Big data Analytics

With more than 4 million orders and around 20 tracking events for each order analyzed, the present project becomes part of the realm of big data. As Stiller et al. (2010) stated, "big data" is when the size of the data itself becomes part of the problem, and hence, cannot be processed effectively with the traditional applications that exist. In addition, McAfee and Brynjolfsson (2012) states that big data analytics (BDA) seek to glean intelligence from data with three key differences, namely, volume (amount of data generated), velocity (speed at which vast amounts of data are being generated, collected and processed) and variety (different forms of unstructured data). Moreover, Opresnik and Taisch (2015) appends veracity and value, completing "the five V's", consensual across literature as pivotal factors that determine the pertinence of a big data problem. As CRISP-DM model (figure 2.3) demonstrates, a structured way of thinking is needed to extract meaningful knowledge from data (Johnson and Wichern, 2007). The initial stage focuses on understanding the project objectives and requirements from a business perspective and hence, select the proper data for analysis. The second stage reflects the discovery of detailed insights about data followed by the process of making the data more suitable for any modelling which includes activities such as - data cleaning, transformation, normalization, and others, further revealed. Finally, modeling techniques are selected and applied, and certain parameters calibrated, which allows the emergence of a model ready to be assessed and fine-tuned and hence, be deployed to a specific context.

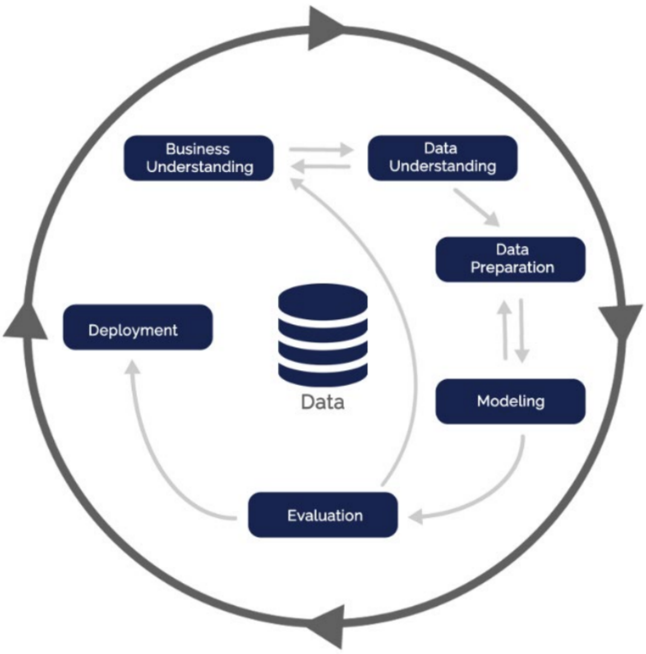

Figure 2.3: CRISP-DM model (Otaris, 2018)

As a consequence, Data Mining could be defined as the analysis of observational datasets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner (Hand et al., 2001). A computer science filed which plays a big role in BDA and derives from Data Mining is Machine Learning. Briefly, Machine learning (ML) consists on the science (and art) of programming computers so they can learn from data. Applications associated with ML are immense from document classication, fraud detection to medical diagnosis.

Firstly, it is imperative to clearly separate two learning principles: supervised and unsupervised. Regarding the first group, the training data (observations, measurements, etc.) are accompanied by labels indicating the class of the observations. Test data is classified based on the training set, which allows assessing algorithm performance. Two main forms of data analysis arise from this principle based on the target variable - classification and regression. Both are used to extract models describing important data classes or to predict future data trends. However, when the target attribute is categorical, a classification problem is present while a numeric variable leads to a regression problem. On the other hand, on unsupervised learning, the aim consists of establishing the existence of classes or clusters in the data. By partitioning data into homogeneous groups, clustering reveals a potent data-mining process. To build the model, a critical stage consists of splitting the dataset into a training set, used to study relationships between features (input) selected and its labels (output), and testing set to assess predictive algorithm performance. Regarding methods for estimating a classifier's accuracy, two techniques stand out.

1. <u>Hold out</u> - It splits the data into training data and test data. Afterwards, it builds a classifier using the train data and tests it using the test data. Due to the efficiency and easiness nature of this method, it is usually used when a dataset contains thousands of instances, including several hundred instances from each class.

2. <u>Fold Cross Validation</u> - As revealed on figure 2.4, the data is split into k subsets of equal size. In the k-iterations, one subset is used for testing and the remainder for training. (Géron, 2017) suggests 10-fold cross-validation as a common practice. It represents a powerful preventative measure against a common Data Mining issue, named overfitting. The mentioned problem happens when a model learns the detail and noise in the training data to the extent that it negatively impacts the performance of the model on new data. This means that the noise or random fluctuations in the training data is picked up and learned as concepts by the model.

```
Data: training dataset
Result: Estimate of performance of models with different parameters
initialization;
foreach Parameter set do
    Load classifier with new parameters
    Split into K-fold validation sets
    for k ∈ K do
        Train on train set
        Test on validation set
        Calculate performance measures
    end
    Calculate mean and variance of the performance measures in the K
      validation runs
end
```

Figure 2.4: Cross-validation underlying logic

2.4.1 Data pre-processing

Data Science experts often quote that data preparation consumes around 80% time of overall time of an analytics project, known as the "80/20 dilema" (Ruiz, 2017). One of the classical data preparation tasks involves outlier detection. In addition, Johnson and Wichern (2007) defines an outlier as an observation which seems to be inconsistent with the remaining dataset. Likewise, Sola and Sevilla (1997) argues that an exact definition of an outlier often depends on hidden assumptions regarding the data structure and the applied detection method. A multivariate analysis includes more than one independent variable and hence, a univariate outlier analysis can be performed within groups. Boxplot, invented by Tukey in 1941, presents a straightforward and clear way to represent the median, first and third quartiles of continuous univariate data. Outliers are considered points outside of 3/2 times the interquartile range and extreme ones outside of 3 times. Sola and Sevilla (1997) demonstrates the importance of data normalization as an input for any Data Mining activity.

While addressing data-mining problems, data set balancing is a key topic to achieve actionable outcomes. Abu-Mostafa et al. (2012) shows that when the data is highly unbalanced, algorithms tend to degenerate by assigning all cases to the most common label. Two major techniques to handle imbalanced data consists of under- or over- random sampling. On the one hand, when one target label is the underrepresented minority class in the data sample, over -sampling techniques may be used to duplicate these results. Conversely, in case of an overrepresented majority class, under-sampling may be used. A cost-sensitive approach could also be followed by increasing the weight of mistakes on the minority class.

2.4.2 Performance metrics

After the modelling process, forms of assessing techniques performance are explored in this section. At the core of binary classification problems, the concept of confusion matrix depicts the predictive ability of model with several metrics deriving from it (figure 2.5).

- True Positive (TP) - predictions which were correctly classified as 1;
- True Negative (TN) - predictions which were correctly classified as 0;
- False Positive (FP) - data entry is wrongly classified as belonging to class 1, which leads to a type I error;
- False Negative (FP) - data entry is wrongly classified as belonging to class 0, which leads to a type II error.

An extensively used metric consists of on the total number of correct predictions over the total number of observations, accuracy. This metric could reveal misleading for imbalanced scenarios. In other words, if the dataset contains 80% of a class 0, a model trained with that data and having an accuracy of 80% could mean that is it always predicting 0 and not finding relationships between variables. Therefore, it is imperative to consider other meaningful metrics used across

		PREDICTED CLASS	
		Class=Yes	Class=No
ACTUAL CLASS	Class=Yes	a	b
	Class=No	c	d

a: TP (true positive)
b: FN (false negative)
c: FP (false positive)
d: TN (true negative)

Figure 2.5: Confusion Matrix for classification purposes

literature. Sensitivity (or recall) measures the completeness level by determining the true positive recognition rate. Contrarily, the true negative recognition rate is commonly named specificity. In addition, precision measures the exactness level by computing the proportion rate of relevant within positive predicted. Finally, F-score is based on the harmonic mean of metrics precision and recall. Therefore, values close to 1 indicates a high predictive ability.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{2.1}$$

$$Recall = \frac{TP}{TP+FN} \tag{2.2}$$

$$Sensitivity = \frac{TN}{FP+TN} \tag{2.3}$$

$$Precision = \frac{TP}{TP+FP} \tag{2.4}$$

$$F-score = \frac{2*Precision*Recall}{Precision+Recall} \tag{2.5}$$

Equally important is the receiver-operating characteristic (ROC) curve, which represents an evaluation of the classification accuracy of a certain test under various conditions. The mentioned curve can be determined by plotting the true positive rate against the false positive rate as depicted in figure 2.6. Originally ROC methods were developed to help evaluate the accuracy of radar operators during World War II as part of the development of signal detection theory. Later, psychologists used the ROC in evaluating experiments in sensory detection and now, the applications are endless. A widely used indicator as a general measure of a classier performance is the Area Under the Curve (AUC-ROC). In fact, it comprises a robust, direct way to assess the predictive ability Han (2011).

Lastly, it is crucial to understand the most important domains when comparing models predicting ability. Extensively covered across literature, a comparative analysis should be performed on a multi-dimensional level.

- Performance – classifier accuracy, consisting of predicting class label;
- Speed – time to construct and use the model (training and prediction time;
- Robustness – ability to handle noise and missing values;

- Scalability – capacity to construct the classifier or predictor efficiently given large amounts of data;

- Interpretability – understanding insights provided by the model.

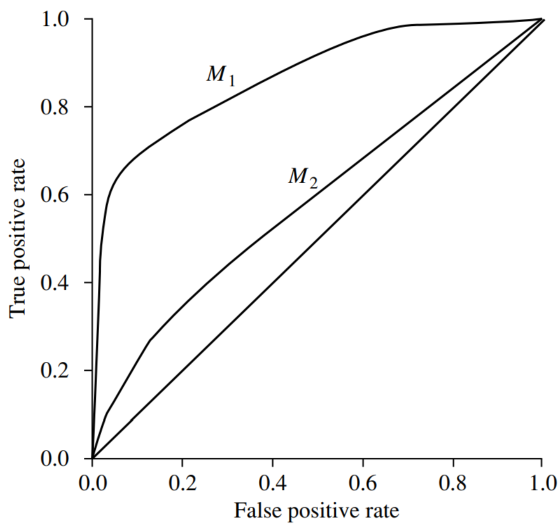

Figure 2.6: ROC curve for two classifiers (Han, 2011)

Adopting an end-to-end approach, all domains that impact the present dissertation content were properly framed in literature not only conceptually but also, analytically. Thereby, narrow the analysis to Farfetch context and fully demonstrate the high pertinence and impact of the present dissertation on the organization, represents the purpose of the subsequent chapter.

Chapter 3

The Challenge

3.1 Order processing

With order volume rising in a faster pace every year, combined with the worldwide nature of shipping operations, internal processes must be as streamlined and automated as possible. From the moment an order is placed, a six-steps process ensures that an item is delivered and represents the operational foundations of the current project. From a delivery-focused perspective, lead time comprises the order cycle time span composed by the mentioned steps, whose statistical parameters are displayed in figure 3.1.

Step 1 - Boutique checks stock availability

Firstly, the allocated partners for a specific order confirms its fulfillment ability. Not having a real-time stock-related data synchronization system, an in-store sale could happen at the same time on an online order or boutiques' clerks may fail to register all sales occurring in the brick-and-mortar environment. In case of no stock, the order can be cancelled, and the customer is refunded, or a similar item is suggested. If the suggestion is accepted, the item is swapped, and hence, the order continues its normal flow.

Step 2 - Payment approval

Meanwhile, within Farfetch boundaries, a fraud recognition tool categorizes orders in three levels according to the likelihood of being fraudulent which leads to its automatically acceptance or rejection for most orders processed. For cases in which payment authenticity is uncertain, a manual investigation process is triggered.

Step 3 - Boutique decides packaging

When an order is confirmed, Farfetch recommends the most suitable standardized box for a specific product category. The specifications of boxes are defined by Farfetch and delivered through a 3PL provider. Despite that, partners have full control over the process of allocating a specific box to an order and can add personalized features.

Step 4 - Boutique creates a shipping label

Once the order has been packed, an internal tool creates, automatically, an AWB (air-way bill), document that identifies a parcel when shipped to end-customer. For a low volume of orders, in which specific carriers' information system is not fully integrated with Farfetch back-office system, the mentioned document must be created manually using external tools. In case of customers have not filled shipping-related information correctly, a manual correction process must be carried out as well.

Step 5 - Send Parcel

At this stage, the order is ready to be picked up by the carrier staff. Depending on the order volume, pick-ups can be performed on a daily basis or upon necessity. From the moment the parcel is scanned by carrier, the shipment is considered to be in-transit and the carrier is liable to deliver it on a time window in accordance with service level agreements established.

Step 6 - Parcel in transit

Presently, the delivery process is mainly assigned to DHL and UPS, which perform around 98% of all shipments (based on 2018 data) with standard and express being delivery options primarily used. With expansion to new markets on both demand and supply sides, the need to diversify carriers used emerged, which resulted in agreements with ECMS in China or Correios in Brazil, for instance. Despite that, their contribution to overall order volume is still low. On a higher granularity level, routes are divided in cross-border in case of a customs clearance process occurs, domestic whenever order flows within a given country and intra-community if different European Union countries are involved. For delivery options "Same-day" and "F90", mostly local carriers are used to ensure that shorter time-to-deliver offerings are fulfilled. In case of cross-border shipments, in which a customs clearance process is undertaken, they are all exclusively handled by DHL. Moreover, for a specific origin and destination country, the route is not static, being altered in accordance with carrier's logistics network. As a crucial stage to customer experience-enhancement solutions, this step will be analyzed thoroughly across this dissertation.

Table 3.1: Steps statistical parameters for 2018, in days

Steps	1	2	3	4	5	6
Mean	0.583	0.108	0.196	0.021	0.592	3.200
Median	0.3	0.0	0.0	0.0	0.3	0.9
Std. Dev	0.57	0.55	0.55	0.42	0.55	2.17

Broadly, the first five steps encompass an important KPI commonly referred as speed of sending. Alongside with the percentage of no-stock orders, represent the main metrics to assess responsiveness degree of partners. An incentive plan is set into place to assure that boutiques have visibility over admissible values of responsiveness and can optimize their own fulfillment operations. Being the major contributor to lead time by far, step 6 (referred also as transit time) reveals a vital improvement point in which enhancements have a high overall impact. For that reason, delivery operation teams strive to smooth operations in straight collaboration with carriers, customs authorities and any other contributor. When analyzing order flows, it is important to distinguish three granularity levels, depicted in figure 3.1. At a higher level, a portal order consists on the set of items associated with one checkout stream (commonly referred as "basket" or "shopping bag"). Logically, a portal order contains at least one order from the same boutique, denoted as boutique order. From a delivery operations perspective, since that items shipped from the same boutique are packaged together, a boutique order is commonly named parcel. In sequence, a boutique order comprises at least one item. On the current project, given the fact that following analysis performed is delivery-focused, boutique order represents the most suited level.

Figure 3.1: Order granularity levels

3.2 Customer experience

At Farfetch, customer experience is treated as a multi-dimensional topic. Thereby, each team has its own initiatives to incrementally enhance customer interaction in multiple touchpoints across platform. Being the leading global online platform for the booming market of luxury goods, Farfetch has a unique opportunity to capture demand and ensure that customers remain in its boundaries, building a solid and powerful position in the market. A key element to gain customer trust and loyalty consists of removing any kind of friction during the customer journey, turning it as smooth and immersive as possible. Breaking down the customer journey in a holistic view, five main stages stand out as depicted in figure 3.2. When customers enter Farfetch ecosystem, a search process begins which may result on selecting a certain item materialized with its addition to the "shopping bag". An option to save items for later is also available in the form of a "wishlist". This feature creates an important snapshot of customers' way of thinking, crucial in fashion industry. For instance, trends can be easily identified, and merchandising tactics optimized to capitalize on them. In addition, knowing the origin of website visits is also crucial not only as an important key performance indicator for marketing initiatives but also to understand which "entry points" are more appropriate and profitable given customer traits. Therefore, monitoring cleverly and in a business-oriented way website traffic, while guaranteeing that the platform remains as much user-friendly and compelling as possible are tremendous challenges faced on a daily basis. When buying decision-making process ceases, customer goes through an order placement flow consisting on the following three phases: (i) choosing shipping address or a Farfetch collection point (enabled by boutiques brick-and-mortar points), (ii) selecting the payment method and, (iii) indicating the delivery method and review all order-related details.

Regarding payment methods, for the time being, Farfetch supports 19 modalities depending on the customer country, including 10 credit cards, 5 e-wallets and 3 banking operations-related. As stated before, express, standard, next-day and same-day delivery methods, as well as deliveries in 90-minutes are available depending on the shipping address for a specific order. Once an order

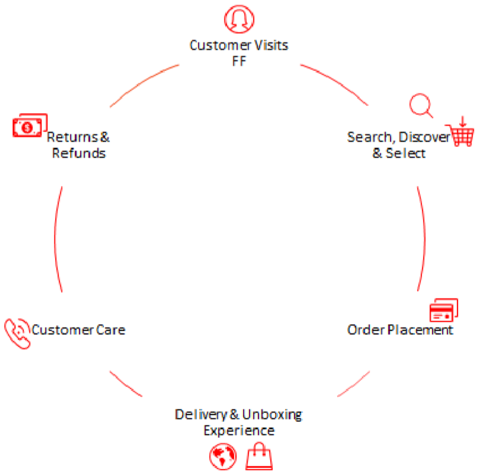

Figure 3.2: Customer journey capsule

has been placed and steps 1 and 2 are surpassed, partners' order fulfillment speed combined with responsiveness and efficiency of 3PL providers determine the time-based dimension of delivery experience. By being dependent on external entities, lack of visibility and control arises. Along with these, it is the fact of stock points being highly geographically scattered, delivering a consistent delivery experience consists a major challenge. Once customer interacts with an item for the first time, an unboxing experience takes in, which decisively influences customer engagement. Currently, customer has no control over packaging options or some sort of personalization such as customized gifts messages. Furthermore, a key competitive advantage embed entails returns free-of-charge together with refund of the total amount spent (including DT involved). As expected, fitting-related reasons consist of the major returns source making up around 60% of the total volume processed. With current return rate of around 18% (by volume), slightly inferior to other e-commerce players that offer a similar service level, the impact of this competitive advantage is tremendous.

Whenever a customer faces a challenge across the journey, it often results in contacting Customer Support team. Therefore, contact reasons reveal a powerful datapoint that helps to build the foundations for most customer experience analysis. By analyzing the number of contacts ranked by category, half of the 8 major reasons are delivery-related as shown in table 3.2. "Initiate a Return", "Tracking information", "Order delay" and "Lost parcel" represent around 36% of total volume combined.

By exposing the relevance of delivery information and being a massive source of customer contacts, improvements can drive an intertwined effect of savings and service level increments. Consequently, it is imperative to clearly map the current status of key elements that impact delivery experience taking an end to end approach. A lack of perception and stakeholder's alignment on this topic resulted in inefficient and separated workflows. Thence, the project under analysis also seeks to unify individual contributions and organize value-adding layers.

Table 3.2: Contact Reasons ranking by number of contacts

1. Initiate a Return
2. Refund status
3. Tracking information
4. Cancelled by customer
5. How to place phone order
6. Order delay
7. Faulty item
8. Lost Parcel

3.2.1 Time-related compliance

As exposed in section 2.1., knowing precisely when an order will arrive has a positive impact on customer satisfaction and retention. By increasing time-to-deliver awareness and accuracy, meaningful improvements could arise. Some internal initiatives allowed to present a dynamic Estimated Delivery Date (mentioned as EDD from now on) across the order cycle. A predictive algorithm was recently developed and trained using historical data to provide an accurate EDD at checkout. From that moment on, EDD accuracy rose from 70% to 85% and the time window become dynamic and 1.1 days lower. Therefore, the customer knows not only more precisely when an item will arrive but also the promised date interval is fulfilled more often. The mentioned value remains in personal user area until the order is shipped. Afterwards, EDD predicted by carrier accessible through tracking log overwrites the previous value. This post-shipment value consists on a one-day time window and is successively updated by the carrier as the order gets closer to the endpoint. Depending on carriers' ability to control their own operations, the value could evolve to shorter timespans. For instance, DHL is able to indicate a 4-hour time window when a specific order is out-for-delivery. Therefore, on post-shipment segment, one single EDD is provided to customer regardless of the platform used to acquire tracking information allowing a consistent delivery experience. Such touchpoints could be internal such as user area and transactional emails or external such as carriers' multiple tracking platforms. By achieving consistency on this topic, customer trust in Farfetch tracking information increases and hence, reliance on external platforms decreases. Despite that, while analyzing EDD accuracy from edge-to-edge, a drop of 35% was detected on the moment orders are shipped (figure 3.3) which indicates that carriers' initial prediction is less accurate than the pre-shipment Farfetch one. Evidently, lower carriers' time windows are harder to be fulfilled but other reasons are constantly being noticed. Farfetch logistic network traits such as high transit time variability justified by multi-lane and low volume nature of the marketplace reality combined with high cross-border flows clearly affect the predictive ability of carriers. In fact, carrier staff members acknowledged this inefficiency.

In addition to precision, it is also important to reflect how EDD should be presented to customer. Studies show that what users really care about is not shipping speed (that is, time window in days) but rather the delivery date interval (Appleseed, 2017). In case of using calendar dates, the format becomes a critical issue to ensure that correct timespan is immediately perceived by customers. With that in mind, checkout EDD was recently changed to accommodate this nuance.

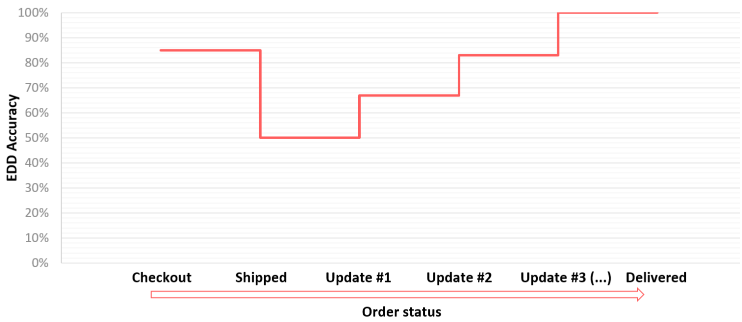

Figure 3.3: EDD accuracy evolution

Post-shipment EDD end-points, not explored in detail until the present project, were mapped and some format inconsistencies found.

Briefly, pre-order current time-related elements are capable of proving a robust and enhanced delivery experience. Since the order has been dispatched, efforts to ensure the initial promise is met while fine-tuning EDD across the journey in a consistent manner must be put into place.

3.2.2 Tracking experience

When shopping online, a core element consists on the degree of awareness related to order status. As stated before, tracking information represents the third major pain point when considering customer contacts which reflects the immense room for improvement. Providing full visibility over tracking leveraged by meaningful and tailored communication flows represent a vital milestone. Internally, the customers have access to tracking information through their personal area (via app or browser) and transactional emails triggered whenever a relevant change in order status is detected. Externally, once the order has been shipped, customer can access to carriers' tracking platforms, whose tracking log contains more shipment details than internal touchpoints. At the time of this dissertation, customer has visibility over four states within Farfetch ecosystem.

1. Order placed - This state remains while step 1 and 2 are not completed. Regarding tracking information embed in transactional emails, this stage is named "Reviewing" while via in-app or web-based user area, the corresponding term is "Placed";

2. Order confirmed - From this moment on, the customer becomes aware that order fulfillment has started. "Preparing your order" and "In progress" are the designations attributed to this stage to user area and transactional emails, respectively;

3. Order shipped – For a specific lane and delivery option, a pre-determined carrier is allocated to an order. From the moment parcel is picked up, depending on the carrier selected, tracking experiences could diverge. Some incumbents present a proper tracking platform supported by some communication flows such as DHL. Others such as Correios (used in Brazilian domestic standard shipments) offer poor tracking experience accessible through a nor user- neither mobile-friendly interface and inefficient customer support in case of incidents. The

stated problem is mitigated by the high prevalence of DHL on overall shipments of around 82%;

4. Order received – With the growing trend of customer reliance on other person or entity (e.g. neighbor or doorman or office personnel) to receive a given order as nomadic lifestyle adoption increases, it is imperative to clearly notify the final transition. Delays on capturing it could inhibit the ability to initiate a return request which harms the downstream return experience for instance.

At first glimpse, it is clear that current order status view does not hold enough tracking information, which prompted the customer need to acquire more details in external sources or contacting Customer Support team. Highly relevant pieces of information such as the access to carriers' receiver experience enhancement service, is only possible through external touchpoints. In parallel, a deeper analysis on tracking data flow from collecting, storing, processing and acting on it, allowed the detection of a high latency level which leads to late transitions between order states, the basis for communication triggers revealed on the next section.

3.2.3 Communication flows

In an increasingly digital aware world in which individuals are constantly being targeted with large amounts of information via several communication channels, it is critical to establish a meaningful communication strategy. Being part of high-end fashion market, communications flows enabled by marketing initiatives are a powerful way to maintain Farfetch aura of luxury and exclusivity and improve brand awareness. The company has been creating innovative shopping stimulus capable of generating revenue. For instance, in-app push notifications promoting the launch of a set of items from a specific brand are commonly set in motion. Additionally, a constant concern on content is present with teams focused on tailoring the right language and tone of voice to consumer segments keeping them engaged. In light of this, the company has been striving to ensure consistency on how to approach a customer via direct contact such as a customer contact call or through automated channels such as transactional emails. However, excluding marketing flows, order-related communication strategy has remained static in terms of triggers and channels as well as disperse with several entities interacting with customers. A clear example consists of the way certain carriers are feeding customers with valuable information in several order states transitions with their own language, not in accordance with Farfetch standards. In addition, customers are forced to leave Farfetch immersive and customer experience augmented ecosystem. To understand the extent and severity of the mentioned problem, the main internal and external communication streams were clearly identified.

As depicted in figure 3.4, four main triggers are currently responsible for at least five communication streams for every boutique order that goes through all steps and it is delivered. Since triggers are based on the boutique order level, for cases in which several boutique orders compose a single portal order, flows are five-fold the number of boutique orders. With this incapability of aggregating portal order communications by postponing, for instance, boutique order triggers, the customer is targeted too frequently. Concerning channels, email is primarily used by Farfetch and external stakeholders. Besides that, some carriers such as DHL, contact the customer via text message in countries which has a local number. Due to the growing global scope of the marketplace, additional channels must be provided to satisfy specific customer segments' needs. With an email penetration rate as low as 38%, the Chinese online users (according to CNNIC, entity responsible for operating and administering China's domain name registry) is a critical example of the need to revamp the mentioned problem.

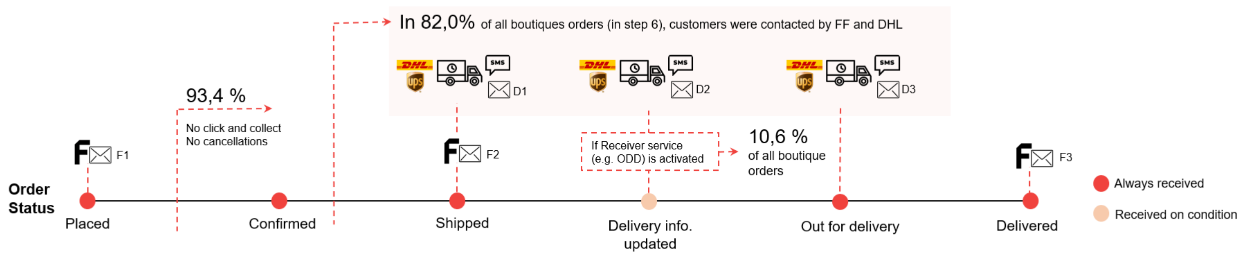

Figure 3.4: Order status and order-related communication flows

Concerning communication trigger accuracy, high and variable latency exposed on previous section leads to inadmissible delays. From a consumer perspective, the mentioned problem is evident in cases where redundancy exists. As a matter of fact, when order is shipped, a time-lag between Farfetch and carrier e-mails is easily perceived by customers. In an attempt to capture delivery enhancement elements, present on carriers own communication strategy that precludes its suppression, each flow content was mapped. Concerning awareness related to the out-for-delivery state via communication, it is only enabled by carriers. In addition, the access to receiver services, further explained in detail, is only possible through a smart link that customer receives firstly on shipped parcel DHL announcement and in further external streams. In case of activating a certain receiver service capability, explained in detail in the following section, a new communication flow is triggered. Additionally, an on-demand communication tool is available through an external provider to guarantee some flexibility in some untypical customer contact-required occasions. Such events are typically unpredictable such as order processing errors.

To achieve a more tailored and frictionless customer interaction, communicate uniquely when main transitions occur, becomes scarce. Due to high transit time variability emerging from a high logistics network complexity, opportunities to feed valuable information to customers are immense. For instance, more than 70% of all parcels go through a clearance process on destination country customs, responsible for unexpected delays. The stated problem is exacerbated for high-valuable items such as others sensitive to import regulations such as exotic materials-based products. In an even simpler approach, repetitive events such as holidays and weekdays in countries where parcels go through can easily be identified and are not being considered on the current communication strategy. Some spontaneous events such as adverse weather conditions or strikes, despite its predictability being low, could also reveal powerful datapoints. Therefore, adding a proactive approach towards communication triggers capable of handling not only delays and its root-causes but also exceptions that might occur such as damaged or lost parcel, could reinvent current experience. Such initiative would enable the service recovery, highlighted in section 2.3 as an important driver of customer trust and loyalty. Tracking data as an enabler of these flows is at Farfech disposal for most orders due to carriers' information system integrations in spite of being unstructured and noisy at the present moment. Furthermore, creating knowledge layers over tracking logs could also drive important internal streams. Customer support teams are acting on a purely reactive approach solving issues triggered by customer contacts at the present. For specific reasons, customer contacts are almost certain that will occur and are not being flagged as action-required on a proactive approach. In light of this, it is highly relevant to clearly understand carriers' role on customer support. For some address-related issues, for instance, contacts are handled in extent majority by carriers' own Customer Support team. In spite of representing savings, it does not fit in overall Farfetch strategy of controlling all communications flows.

Summing up, the order-related communication strategy is not in accordance with the company standards and ambition. Reach your customers at the right time, on the right channel, with the right message regardless of external providers interactions, represents a big challenge. Having said this, it is imperative to convert current static reality to a dynamic and flexible one.

3.2.4 Other drivers

As reflected in section 2.2., last-mile delivery has been disrupted in an intense way with e-commerce players rushing to develop game-changing capabilities. Thence, on a regular basis, people are confronted with innovative receiving experience opportunities. For instance, in-car and in-home delivery service already enabled by Amazon illustrates it. Empowering customers in a way that they can adjust delivery settings to upcoming and changing needs is crucial. To allow that, some e-commerce players are thriving to increase their control and visibility over downstream operations. Working in straight collaboration with carriers and leverage bargain power build upon increasing revenue streams is crucial to achieve tangible results. Currently, receiver services such as DHL "On-Demand Delivery" (ODD from now on) provide a wide range of different capabilities.

1. Change delivery date - Schedule a new delivery date and time window. The new delivery date must be up to 7 days following the original date;

2. Release or Leave signature – Customer either authorizes DHL to leave the shipment without a signature or uploads it on an online form. Clear instructions of a safe place at customer's address and specific date are also requested;

3. Collect from service point – Enables parcel redirection to a selected collection facility. These locations can be either DHL owned or 3rd parties (such as 7-eleven convenience stores)

4. Leave with neighbor – Permits redirection to nearby home address (up to two houses either side of the original address or directly opposite). In case of theft or lost parcel, customers are legally liable.

Exploring this datapoint, it was unveiled that DHL service presents a worldwide coverage with some exceptions for specific capabilities. Sweden and Finland were identified as the only relevant countries in which ODD is not fully deployed. It is important to mention that, from the beginning, Farfetch disabled another capability that allowed the customer to change shipping address along the way, due to fraud reasons. Regarding usage, 11% of all boutique orders activates at least one capability (with growing trend) which clearly shows the high customer appreciation for the service. Breaking down, as illustrated in figure 3.5, “Release or Leave signature” as the most valued one by far, followed by “Change delivery date”, “Collect from service point” and “Leave with Neighbor”. Linking time-to-ship dichotomy of delivery convenience against speed analyzed in section 2.3, convenience weight seems to be rising based on this datapoint.

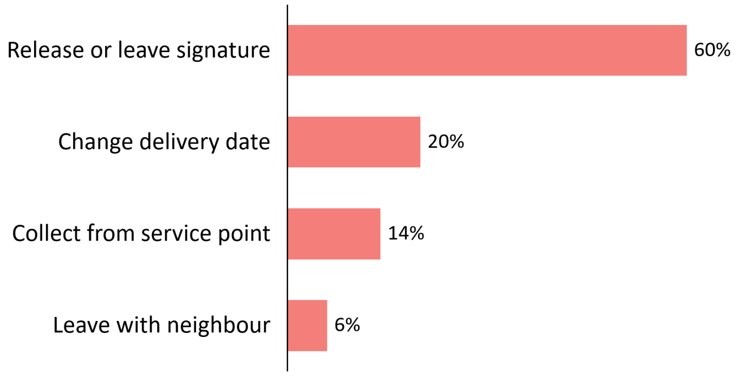

Figure 3.5: ODD capabilities usage

Moving towards customs-related drivers, it is important to reflect about a major competitive advantage Farfetch has been building in straight collaboration with carriers, associated with import and export regulations. According to International Commercial terms (named incoterms) nomenclature, the company operates on a DDP (Delivered Duties Paid) modality. DDP consist of a commercial rule in which the seller bears the responsibility, risk, and costs (mainly, DT) associated with transporting goods until the buyer receives it. From a customer standpoint, lower transit time, smother clearance process and eliminate customs authority’s touchpoint generates an unprecedent frictionless experience. All charges involved are simply incremented to the final price presented online. Currently, DDP method is applied to the extent majority of destination countries while the remaining ones due to specific customs or carrier constrains, an opposite rule is deployed, named DAP (Delivered at Place).

Succinctly, it is important to capture and have visibility overall delivery experience pieces provided by multiple entities to exactly understand post-order customer satisfaction degree. Moreover, it is essential to be aware of different capabilities available through other carriers for instance and understand whether or not might interest Farfetch clients. With the current post-order delivery experience analyzed, it is important to understand the analytical potential of current project by identifying the most pertinent data sources.

3.3 Nature of data

With a goal in mind of developing a data-driven analytical solution towards enhancing delivery experience and frame other less challenging improvements, datapoints at disposal were identified. Time-related data of the main operational steps were easily accessible. Despite that, the ability to know the exact location and time of a certain order since it has been shipped required more detailed information about step 6. Therefore, order tracking logs, available in Farfetch database for express and standard delivery options, was the first datapoint explored. Tracking log consists on a record of diverse nature events that are generated every time an AWB is scanned. Due to the fact of being an unstructured data source with several inconsistencies, it was not used on its full potential as input for knowledge discovery processes. In addition, tracking data of multiple carriers was not normalized when captured, equivalent data was stored in different ways. Besides that, data quality and size differ by carrier. For instance, DHL tracking logs present a clear event categorization while other carriers such as UPS ones offer poor visibility over specific checkpoints where parcel goes through. Location-related data was available at a standardized country and city level. Since that most current analysis are performed at a country level due to noisy nature in lower granularity levels, this datapoint reveals a powerful resource. Additionally, in demand peaks periods such as Black Friday, tracking logs seemed to reveal an abnormal amount of inconsistencies such as timestamps in the wrong order or wrong locations. It is also important to highlight that manually generated data such as creating an AWB on step 4 using external tools, is considered a noisy data source and is eliminated for most studies.

Concerning items intrinsic features, category and value could affect profoundly its flow speed particularly when involve a customs clearance process. Others such as brand, size or color are considered irrelevant while performing this kind of analysis. In theory, sales volume could also affect time-to-deliver offering since that peaks could disrupt carriers network resulting in delays and more exceptions. Route traits such as customs regulation and exception days (weekends and holidays) experienced in several countries a parcel goes through must be deeply analyzed to understand the inner dynamics.

Thereupon, data foundations for the envisioned value-creation process seemed to be consistent despite some limitations. Being an uncharted terrain in Farfetch, data-driven post-order initiatives represent a promising topic to decisively impact customer experience. It was consensual that it represented an essential battlefront to conquer even more customer loyalty and trust.

Chapter 4

Delivery Experience Study

4.1 Delivery experience framework

As explained across the present dissertation, delivery experience needs to be revamped mainly after the checkout process without losing holistic awareness. Understandably, transit time consist of a decisive and evident driver. Solutions to lower it and variability involved are constantly being addressed by the Delivery Operations team. However, delivery experience goes well beyond this value and hence, it is crucial to understand relationships and contributions among other drivers. In order to achieve fruitful outcomes, it is essential to separate a set of ideas into subsets that are mutually exclusive and collectively exhaustive (MECE principle). With that in mind, a framework emerged focused on 3 high-level domains. By adding conceptual value to processes, it was possible to accelerate the alignment of internal points of contact, as well as to understand the impact of potential areas of opportunity. Breaking down organizational silos and hence, boost information flow was always a concern.

1. Compliance level

When considering the sense of expectations fulfillment, delivering when promised is a critical issue. Therefore, the compliance level is impacted by six intertwined sublevels. Primarily, accuracy degree measured as a percentage of orders in which lead time is within EDD borderlines. The second sublevel consists of value consistency, that is, to maintain the same value across internal and external touchpoints for each time instance. The third dimension is related to how internal systems are capable of transmitting logistics operations knowledge (such as delays or other exceptions) in EDD across the order cycle by providing the most up-to-date value, in short, update frequency. Additionally, it is equally important to update the date interval in a consistent manner ensuring that each update is within the previous one, obtaining update consistency. Throughout successive updates, narrow time window without compromising accuracy as soon as possible allows to fine-tune the initial promise. Finally, the sixth-dimension focus on the way EDD is presented to customer. Format should be selected based on how easily the time element is perceived by customer.

2. Awareness level

Tracking seamlessly the order journey in its multiple stages and locations creates a sense of awareness essential to keep the customer engaged and nerveless. Regarding tracking experience, it important to understand not only the order states required to achieve a high awareness level but also how to display it in various channels such as iOS and Android app or different web browsers. Along with high tracking visibility, triggering efficient customer communication flows throughout automated channels, in relevant milestones is essential. Communication-enhancing initiatives typically address at least one of three sublevels - triggers, content and channels. Triggers could be divided in the ones responsible to notify main order status and others liable for proactive interaction such as notifying whether an order is delayed. The first group activation is more crucial than the second one since that embodies a must-be quality. Apart from automated channels, direct customer interactions through Customer Support should also be considered on this point.

3. Empowerment level

Once the order has been shipped, customer deliver requirements can fluctuate due to changing needs. Tools and capabilities that empower customers in what concerns control over the delivery process are contemplated on this point. Typically, such means are enabled by carriers and are not being captured by farfetch platform. Receiver experience-enhancing services such as DHL ODD are the most significant example of that. Future improvements in post-order delivery experience should be analyzed in accordance with the framed levels and sublevels. While assessing delivery experience-related initiatives impact, several key performance indicators must be considered.

- Retention Rate – percentage of customers who had a tier (customer loyalty level) 12 months ago and still have a tier now;
- Conversion Rate – percentage of visitors that successfully made an order. A visit is considered a conversion only when the customer successfully places the order and sees the confirmation page;
- Express Service Level – percentage of boutique orders served by express service delivered in less or equal than 4 days;
- EDD Fulfillment – percentage of orders which not arrive within dynamic promised date (before and after) across order status;
- First Attempt Rate – percentage of orders delivered on first attempt:
- Net Promoter Score (NPS) – mathematical classification and measurement of customer satisfaction and brand loyalty powered by a follow-up survey. Only delivery-related questions should be contemplated on this point in this context;

- Contacts Per Order (CPO) – Customer contacts motivated by a delivery issue.

In parallel, it is crucial to monitor the operational side. That is, oversee lead time and particularly, transit time is crucial to detect abnormal deviations and prevent carriers network inefficiencies. The idea of conceiving a multi-module, data-driven and smart model capable of positively impact three-dimensional subject matter set the tone of subsequent workstreams. Aligned with a platform thinking rooted in the organization, customer experience increments should evolve towards a single point of contact reality in which customer does not need to leave farfetch environment to access information of any nature. Increase ecosystem retention and decrease external reliance represent in fact, a vital aspect of holistic customer strategy. Thereafter, considering the developed framework and decreasing granularity level, it was possible to identify specific opportunity points and give rise to associated requirements.

4.2 Areas of opportunity

Comparing the expounded framework against section 3.2., improvements become clear and their individual contribution for delivery experience can be unveiled. For assessing compliance level, EDD performance current status explored above represented the starting point. On the post-shipment segment, low overall accuracy represents an immense concern for the delivery operations team. However, any improvements would comprise consistency between internal and external stimulus, currently, almost perfectly aligned. The only misalignment is related to when the order is out for delivery, moment when the client is presented with a narrower delivery window (typically a four-hour one) on external touchpoint and hence, not being captured by Farfetch ecosystem which only presents the value in date short format, essentially, one-working day window size. In addition, internal dissimilarities could be found between transactional emails and profile account in what concerns date format, unveiled above, which contributes to the incongruity phenomenon. On the first subdomain of EDD update, frequency, some analysis concluded that some updates were being suppressed by back-office services. The second relevant factor, successive update consistency is not being accomplished because each update represents a different date apart from the first post-shipment EDD when compared with checkout time window. But even in that transition, a high percentage of mismatch cases were uncovered.

Adopting a customer-centric approach, several interventions were planned to revamp time-related elements reliability and relevance. Over the short-term, the following interventions were planned and set in motion.

- Fully capture EDD updates regardless of carriers;
- Standardize EDD format across all touchpoints to an alphanumeric one (such as "29 Mar 2019") to increase perceptibility;
- Postpone EDD post-shipment updates by suppressing the first and/or second one depending on accuracy by route;

- Incorporate in Farfetch environment a narrower time window from the moment an order is out for delivery;

- Work closely with carriers back-office teams to ensure that better post-shipment accuracy is granted on their side.

With that in mind, an end-to-end customer experience upgrading could be obtained by improving five of six subdomains previously mapped in section 2.1. The only parameter implicated is value consistency on the first or two updates suppressed, whose effect is inversely proportional to ecosystem retention, the long-time goal of the company. Besides that, updates importance increases as orders get closer to the customer which also mitigates the effect of an initial suppression. Over the long-term, when ecosystem retention is being achieved in its fullness, an intelligent layer should be created powered by a predictive algorithm that calculates a dynamic time window throughout parcel journey. Not only the predictors that feed EDD at checkout should be considered, but also other route-related variables that can be collected as orders get close to the end-point. In fact, this algorithm is a variant of the one developed on the scope of the present dissertation, further revealed.

To strengthen awareness level, initiatives focused on order status and communication flows were set up. Regarding order status, a consensual must-be quality consists of five states - placed, confirmed, sent, out-for-delivery and delivered. Following Kano model logic unveiled before, out-of-delivery must be added on the tracking interface. Apart from that, order states names standardization and its proliferation throughout all internal touchpoints are needed to deal with naming incoherencies detected. Apart from this simple view enhancements, other features were considered attractive qualities. Presenting a detailed tracking view (full log) is a good example of that. Additionally, significant improvements could be made on visual-friendliness of the tracking element on interface. For instance, in “F90” and same-day delivery options, a way to trace every move of the courier would have an undoubtedly positive impact on customer experience. On express and standard options, a good looking and dynamic map with main checkpoints could have a similar result. Regarding external touchpoints, each carrier presents a wide range of tracking platforms with some being more either user or mobile-friendly than others. These platforms were mapped and the one considerate most appropriate for each carrier was identified. As a result, any element that redirects the customer to carriers’ platforms were updated.

Regarding communication flows and its current static nature, workstreams were created based on three main sublevels described in section 4.1 - triggers, content, and channels. Concerning must-be quality-related triggers, the one informing out-for-delivery should be implemented. Moreover, a difference-maker proactive flow was envisioned, and represent the basis for the following chapter of the present dissertation. At the core of this flow should be an intelligent layer build on top of historical and real-time tracking data capable of predicting delays and flagging exceptions that might occur across parcel journey (loss or damage for instance). The exposed layer drives two vital triggers, one automated with higher reach, notifying the customer, and other internal, requiring customer support intervention for extreme cases. Such action-required cases should be based not only on excessively higher delays but also on customer tiers. Upper tiers require delicate treatment, a more tailored and thoughtful experience than a simple automated email. Within the confines of attractive qualities, the ability to merge triggers in a single flow tackling the high number of emails for multi-boutique order "baskets" identified was studied for future implementation. Another example exposed internally was the ability to adjust communication trigger to social hours, timespan in which customers demonstrates more willingness to be targeted and naturally, benefits could arise from instant peer recommendations taking advantage of socialization activities.

For the content dimension, Farfetch already communicates in a wide range of languages and in a proper tone of voice for the targeted audience. Beyond that, in the course of the present project, changes were proposed based in an attempt to include valuable information on flows such as

- Accommodate carriers self-service capabilities on the Farfetch-labelled email sent when order is shipped materialized with inclusion of a smart link;
- Embed items recommendations already presented on portal communication flows to uplift cross-selling (action of selling an additional product to an existing customer);
- Embed smart and effortless feedback to capture customer impression and judgment.

Equally important to what is how to reach customers, using the right channel for the right cluster. With this in mind, several channels were identified, and their potential assessed. Besides traditional channels such as short message service (SMS) and electronic mail, others should be considered.

- Freeware and cross-platform tools – Whatsapp, Facebook Messenger, Wechat, Twitter, among others;
- RCS (Rich Communication Services) alike flows – RCS consists of a communication protocol between mobile-telephone carriers and between phone and carrier, aiming to replace SMS. Similar and even more powerful capabilities are already supported by some manufacturers, particularly Apple with for example, through business chat application;
- In-app or PWA (Progressive Web App) server push notifications – delivery of information on the Web that is initiated by the information server rather than by the information user or client, as it usually is. An interesting topic regarding this communication channel is content size. Studies show that push notifications with fewer words have higher click-through rates (Blair, 2018);
- IVA (intelligent virtual assistant) enabled flows - software entity which can perform tasks for an individual based on verbal commands.

Obviously, integrate the mentioned channels to communication infrastructure and should follow a certain priority according to needs of most representative markets. Set as a top priority was channels addressing Chinese customers, a flourishing market with specific traits as exposed. Therefore, in parallel to the present project, Wee Chat integration was performed revitalizing overall Chinese customer experience. Externally, the SMS consists the only additional channel powered by some carriers for specific countries. Therefore, deeper analysis on this topic was conducted to understand the development requirements of such channel. When external and internal content, trigger and channel improvements are fully deployed, communication flows of carriers become worthless and disposable. These mentioned flows must be suppressed, removing redundancy and getting close to the single point of contact reality. Despite that, to accomplish this reality and deploy most of awareness level enhancers, demanding internal development processes would need to be carried out. Lastly, it is important to define needs for the third delivery experience driver, empowerment level.

In short, by identifying in a structured way the vast room for progress while adopting a holistic approach, significant improvements on specific points are constantly being achieved on delivery experience. However, with such a plethora of improvement points, a benchmarking analysis was conducted to understand how delivery experience-focused external providers could help tackling some problems detected particularly those that required more development efforts. It is also important to reflect on tracking data processing delay which affects almost any customer experience drivers and represent the foundations for many improvements. In parallel to the present project, latency has been reducing with some API's (Application Programming Interface) integration improvements.

4.3 Benchmarking

In recent years, due to rapid market expansion and penetration, Farfetch stakeholders early realized that deploy all technological developments internally required in short periods of time would be unfeasible. With that in mind, outsourcing as a way to capture high-value propositions and deliver successive increments on experience became a recurrent practice. In order to correctly assess external providers in full scope, three main layers were set as a starting framework to support the decision-making process in accordance with the customer-centric framework exposed in section 4.1. Firstly, the integration layer represents the foundations responsible for generating standardized real-time tracking data from multiple sources. This layer must feed a second one, responsible for extract required knowledge from data, named intelligent layer. Lastly, an action layer allows the creation or reinvention of touch-points and hence, deliver value to customers. Essentially, requirements for this last layer are the ones mapped on the last section that affect compliance, awareness and empowerment level. Briefly, five-dimensional EDD solution envisioned, communication enhancement related to trigger, channel, and content as well as betterments on tracking interface, and include self-service capabilities within Farfetch boundaries represent key end-goals. Decreasing granularity degree for the first two layers, specific requirements were mapped.

Integration layer

- API integration with a pre-determined list of carriers and services

- Independence in gathering tracking data
- Standardization of multi-language and time zone dissimilarities tracking events
- Average latency no longer than 20 minutes
- Ease on adding new carriers or new delivery options of existing carriers
- Support in case of incidents

Intelligent layer

- Capable of delivering EDD solution envisioned and hence, improving five dimensions exposed
- Predict delayed shipments in several checkpoints and fine-tune predictions as orders get close to customers
- Identify delay root-causes for specific order clusters
- Flag other exceptions that might occur
- Transit time pattern recognition, capable of detecting deviations for specific lanes

Based on technical requirements and solutions offered in the market, three delivery experience-focused third-parties were identified as potential partners, named from now on as Company X, Y, and Z. In order to understand each value proposition, a trial was set up using real-time Farfetch orders tracking data and solutions analyzed against requirements. Concerning the integration layer, all seemed to meet the requirements. Lower latency than current back-office Farfetch system reality stand out as a notable feature. However, when considering knowledge-gathering processes, none of them fully met the conceived capabilities. Only company Z was capable of delivering a robust EDD solution by filtering the most accurate ones for each route. Regarding delay prediction ability, all used too simplistic underlying heuristics. These heuristics involved determining whether a shipment is stalled for more than a certain period of time in a given logistic hub. Despite that, all of them were able to categorize accurately tracking events to some extent. Therefore, some delay root-causes were identified (such as unpredictable customs clearance delay) but other resulting from more spontaneous events such as bad weather or strikes were not displayed. In addition, other action-required exceptions were detected and correctly classified such as lost or damaged parcel, customer refusal, address issues or parcel not collected.

In respect to the action layer, as expected not all requirements were met but clearly one third-party standout. Compliance level improvements due to the underlying intelligent nature to achieve it were analyzed on the previous layer. In what concerns awareness level, critical for targeting customer, it is important to clarify what each provider offer for either communication or tracking experience. Starting with channels dimension of communication, X and Z presented solutions that go beyond traditional methods. Both offered the ability to contact the customer through at least one channel of each category described previously. Regarding content, X, Y, and Z could embed carrier self-service capabilities through smart link incorporation but none of them presented a white-labelled interface for use it due to carriers' API limitation. Y was the only one not capable of embed smart feedback and even more important, available languages did not fit on current farfetch diverse customer base. Another drawback of Y was the low customization degree compared to other providers. Z was already able to include cross-selling features while others did not have such ability on the roadmap. Concerning triggers, the ability to communicate within social hours was supported by X and Z. The envisioned internal trigger requiring direct contact through Customer Support team, was enabled by all providers by integration with current in-use customer support software. The integration solution proposed was webhook-based, which consists of an API concept that allows applications to interact with real-time information. With respect to tracking experience, X, Y, and Z presented a solid tracking white-labeled interface with all order states defined as a priority. Broadly, no empowerment level improvements could be achieved in collaboration with providers under scrutiny since none of them focus on increasing control degree over the delivery process, being that more linked to carriers' base.

Briefly, Y seemed to present a scarce value proposition compared to other delivery experience enhancers. Between X and Z, both fulfilled a considerable number of requirements with third-party Z being slightly more resourceful on the interaction layer and presenting more robustness on the intelligent layer.

4.4 Workstreams

With all information to support the decision-making process collected from end-to-end, workstreams were defined. Various improvement points defined in section 4.2 could easily be addressed allocating internal resources while others require more development efforts. Clearly, Z provider could revamp delivery experience on a high number of touchpoints. Despite that, proactive communication flow enabled by highly accurate predictive trigger was not possible with provider's elementary heuristics. Therefore, three intertwined workflows, one external and two internal were launched.

1. <u>Deploy effortless short-term improvements</u>
   Based on recommendations unveiled in chapter 4.1, tasks that require low development must be conducted sequentially, as they were uncovered. The process of assigning internally the mentioned tasks to the specialized teams was conducted and initiatives status constantly updated.

2. <u>Outsourcing solution</u>
   In straight collaboration with third-party Z, creation of awareness and compliance level enhancers flows were proposed. In the first stage, communication triggers not encompassed by current Farfech order-related communication strategy should be generated quickly proving as a complementary tool such as out-for-delivery or ODD activation announcement email.

In case of successful outcomes based on metrics revealed on 4.1, a thorough analysis should be undertaken to understand if it is beneficial to assign all post-shipment communication flows to Z or even a complete take-over. In addition, EDD available intelligence should be reflected on new communication flows and in a near future, on all touch-points. At last, tracking interface improvements powered by a white-labeled solution should be fully deployed.

3. Internal solution for intelligent layer module
   For driving a proactive communication flow, the internal solution conceived consists of a machine learning-powered predictive algorithm responsible for foreseeing whether an order will arrive after the initial promise (EDD presented at checkout) in several route checkpoints. A checkpoint could be a logistics hub or customs facility for instance, and the algorithm should be able to support a proactive trigger on each of them as accurate as possible. Thereupon, since that data of required nature was reachable as exposed in section 3.3, an analytical tool was developed and is described in the next chapter. It is important to highlight that this analytical solution emerged from a concrete business need on a quest to enhance the post-order delivery experience paradigm.

Chapter 5

Classification Algorithm and Results

The present chapter comprises the development of a machine learning-based tool whose necessity arose as an output of the previous problem-solving approach towards delivery experience enhancement methods. The aim of this advanced analytical solution is to provide solid and intelligent foundations to the creation of a proactive communication flow. By creating an accurate and tailored trigger, it becomes possible to feed customer, via automated or direct channel, with valuable information about delays, revamping post-order delivery experience and generating customer support-related savings. Therefore, adopting CRISP-DM methodology described on section 2.4 as a mean to structure development, the value-adding process began. Due to the fact that the dependent variable is categorical, classification consist of the most suitable method of supervised learning to be applied.

The data was retrieved using the programming language SQL (Structured Query Language) through Microsoft SQL Server Management Studio, version 17.7. Data understanding, and preparation processes were conducted not only in SQL but also in R through RStudio while the modelling, evaluation and deployment were, naturally, only performed in R. Multiple R packages were used from end-to-end to prepare, process and visualize data. In addition, Tableau was used for more advanced data visualization procedures.

5.1 Data foundations

5.1.1 Data collection

As exposed in section 3.3, the starting datapoint to collect valuable data consists of carriers tracking data available at Fafetch database. The sample was gathered for the year of 2018 not only to capture volatility involved across the year, but also to deal with a stable dataset with all order flows ceased. Exploring it, some initial insights allowed to steer data gathering towards routes in which high delayed orders were a painful reality. As depicted on figure 5.1 based on the top 30 routes by order volume, 14.29% of orders that cross at least two countries and a customs clearance is involved (named cross-border routes), were delivered later than EDD presented at checkout, much higher than domestic or intra-community shipments. Despite that, the mentioned chart displays that domestic delayed orders rate is also significant. Recently, carriers service upgrading in both United States and Brazil allowed to mitigate this issue, decreasing delayed orders rate. In addition, cross-border orders represented around 75% of the entire order volume and only one carrier service is responsible for shipping, “DHL Express”. Thereupon, it was obvious that cross-border data should represent the foundations to extract a tailored dataset.

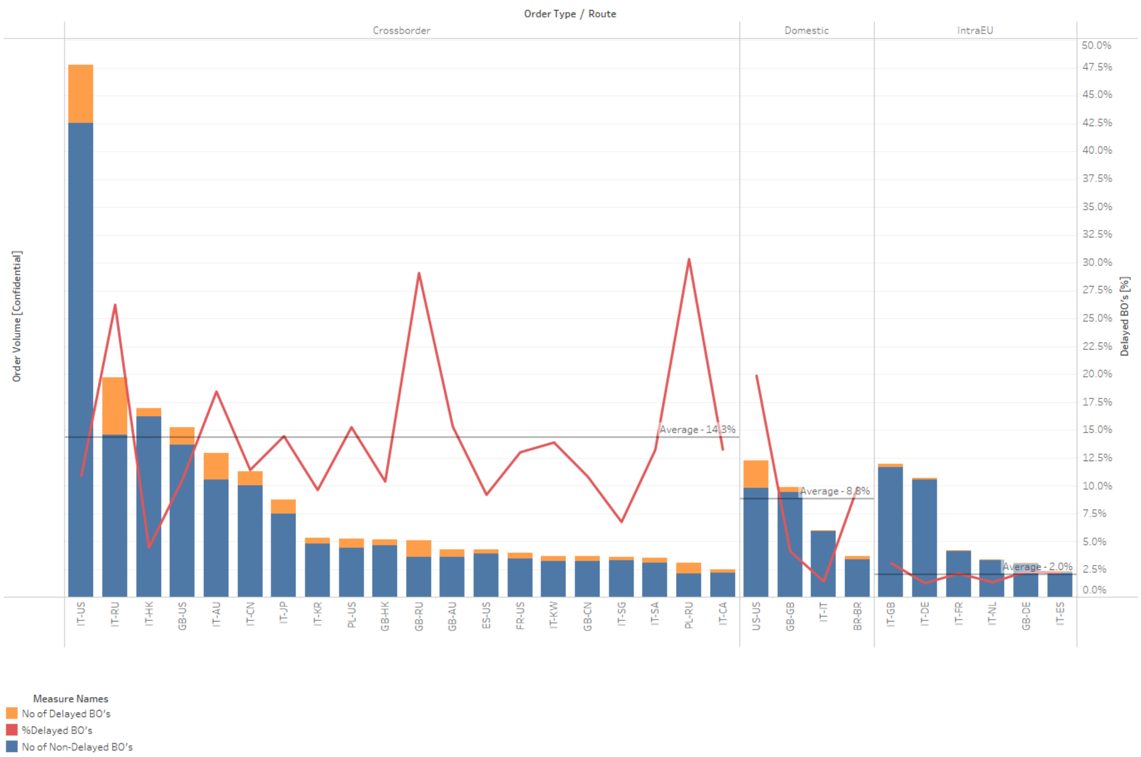

Figure 5.1: Delayed orders for top 30 routes by order volume

While exploring the mentioned datapoint and aiming to feed a communication trigger accurately, it was crucial to determine a homogenous moment in time and location across orders journey, until which all tracking available information at the time is gathered and the dependent variable predicted. After a thorough route analysis, two checkpoints were identified as the ideal timing to trigger the data collection as depicted on figure 5.1, resulting in two different models. Firstly, when an order arrives to the second logistic hub (named **first model**) and secondly, when it arrives to the destination country customs (named **second model**). The idea was to create a trigger neither too close to upstream neither downstream and potentially affect as many orders as possible. On the one hand, notifying customer of a delay occurrence right after it has been placed or shipped is inconsistent. On the other hand, informing too close to the end-point is less valuable from a customer perspective and more easily achieved with simpler heuristics.

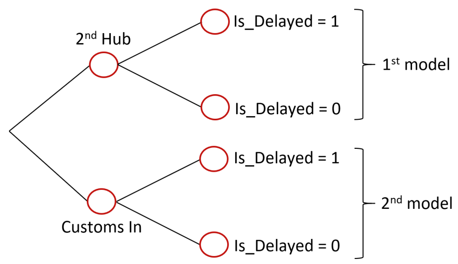

Figure 5.2: Model variants

With that in mind, locations of several checkpoints a parcel goes through as well as time between each of these facilities were collected. Besides Step 6 partitioning (until the second hub and until destination arrival), a variable that comprises the time spent from operational steps 1 to 5 (summarized in table 3.1) was also added to the model. Aware of volatility involved in transport and customs due to a number of reasons identified, other variables considered potential predictors were added. As exposed on section 3.3., items category and value affect transit time, especially if they experience a customs clearance process. Therefore, a binary attribute was created for parcels which contain at least one item of the following categories – jewelry, watch or exotic materials-based. These categories are usually targeted by import regulations and a more rigorous clearance process is performed. In addition, parcels whose items value combined exceeds 2500 USD were also flagged. This value represents a common threshold from which a different and more time-consuming clearance process is carried out. In order to capture demand peaks that could disrupt carriers' network, a binary attribute was created. Based on pre-determined order volume threshold, 11 weeks (out of 52) were labelled and every parcel that was shipped within these periods was flagged. Lastly, based on destination country, a boolean variable was engendered to reflect the DDP and DAP commercial rules already explained. Additionally, easily determined events such as weekends and holidays clearly affect transit time experienced by customer. In light of this, a categorical variable was created that embodies the number of non-working hours between order creation date and maximum value of the EDD. Holidays in countries where parcel goes through apart from origin and destination country were ignored because intermediary logistic hubs are operational in these periods in extent majority despite less order flow capacity. In addition, day of the week in which order was shipped was considered pertinent and added to the dataset.

In Table 5.1, the mentioned attributes are summarized and understandably, each dataset instance is at parcel (or boutique order) level. Pre- and post- data preparation categorical levels are also displayed. Concerning the dependent variable, a delay label is applied for every order whose lead time (time between order creation and first attempt of delivery) is higher than maximum value of EDD checkout (initial promise). Logically, labelling precisely the target variable is crucial. With that in mind, time zone, and weekend days corrections were performed ensuring that lead time experienced by customers and EDD actually seen by them at checkout were comparable. In fact, this procedure was escalated to other variables and consist one of the pre-processing tasks further explained in detail.

Table 5.1: Summary of dataset variables

Name	Type	Description
FromCountry_2digit	Categorical	Country of origin on 2-digit format (47 / 29 levels)
FromLocation	Categorical	Name of origin city (217 / 64 levels)
Hub_2_2digit	Categorical	2^{nd} logistic hub country on 2-digit format (64 / 17 levels)
Hub_2	Categorical	2^{nd} logistic hub city name (251 / 20 levels)
Hub_3_2digit	Categorical	3^{rd} logistic hub country on 2-digit format (70 / 9 levels)
Hub_3	Categorical	3^{rd} logistic hub city name (205 / 17 levels)
Gateway	Categorical	Customs gateway name (306 / 49 levels)
ToCountry_2digit	Categorical	Country of destination on 2-digit format (177 / 48 levels)
ToLocation	Categorical	Name of destination city (825 / 78 levels)
sent_date_dw	Categorical	Shipped day of the week (7 levels)
Exception_hours	Categorical	Expected number of non-working hours (6 levels)
Step1to5	Continuous	Time comprising operational steps 1, 2, 3, 4 and 5
Step6_Ori_Hub2	Continuous	Time between origin and 2^{nd} logistic hub arrival
Step6_1stLeg	Continuous	Time between origin and destination country arrival
Has_exception	Binary	Label for sensitive items
Is_peak	Binary	Label for demand peak period
Is_DDP	Binary	Label for commercial rule applied
Is_HighValue	Binary	Label for high-valuable parcels
Is_Delayed	Binary	Label indicating a delayed shipment

5.1.2 Exploratory analysis

Before advancing into classification learning algorithms, it is crucial to understand the impact of potential predictors on delay occurrences. With that in mind, a thorough analysis of independent variables against the dependent one was conducted.

Location (*FromCountry_2digit,Hub_2_2digit,Hub_3_2digit, Gateway, ToCountry_2digit*) As described on figure 5.3, when analyzing cross-border delayed orders rate by route (origin and destination country), high variability is evident. On the one hand, the three worst performing routes have Russia as costumer country. Such phenomenon is explained by the fact that DDP commercial rule is not applied in Russia, which leads to higher time spent in customs. On the other hand, two best performing routes have asian countries as destination country, specifically, Hong Kong and Singapore, which indicates a streamline reality on these routes. Moving on the order journey and analyzing second logistic hub country (moment for first data collection process and communication trigger) and destination country, interesting insights arise (figure 5.3). Germany started to surface due to fact of being the country where the carrier centralizes most of its European flows before shipping internationally. The same understandings about Russia and Hong Kong can be extracted from this datapoint. When comparing evolution of a specific route from figure 5.2 to 5.3, some predictive ability emerges. For instance, order shipped from Italy to United States that remained in the origin country until the second hub, delayed orders rate rises from 10.9% to 12.7%. In addition, being the second crucial instant when parcels arrive to destination country, a performance analysis by gateway was conducted. As depicted in figure 5.4, high volatility is present even in gateways of the same country. For instance, "East China Area" and "Beijing" (China) present a 14.8% discrepancy, respectively in regard to delayed orders rate. Such fact confirms that gateway represents the most appropriate granularity level to include on the predictive model.

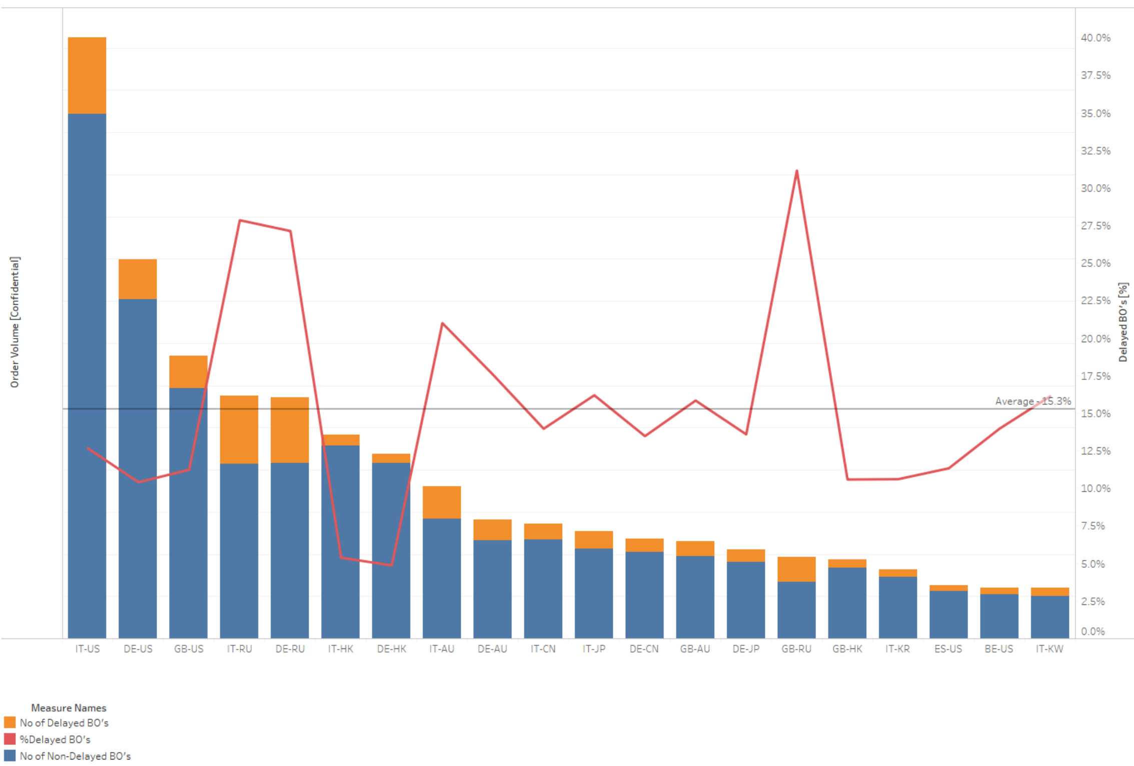

Figure 5.3: Delayed orders for top 20 cross-border routes (2^{nd} Hub - Dest.) by order volume

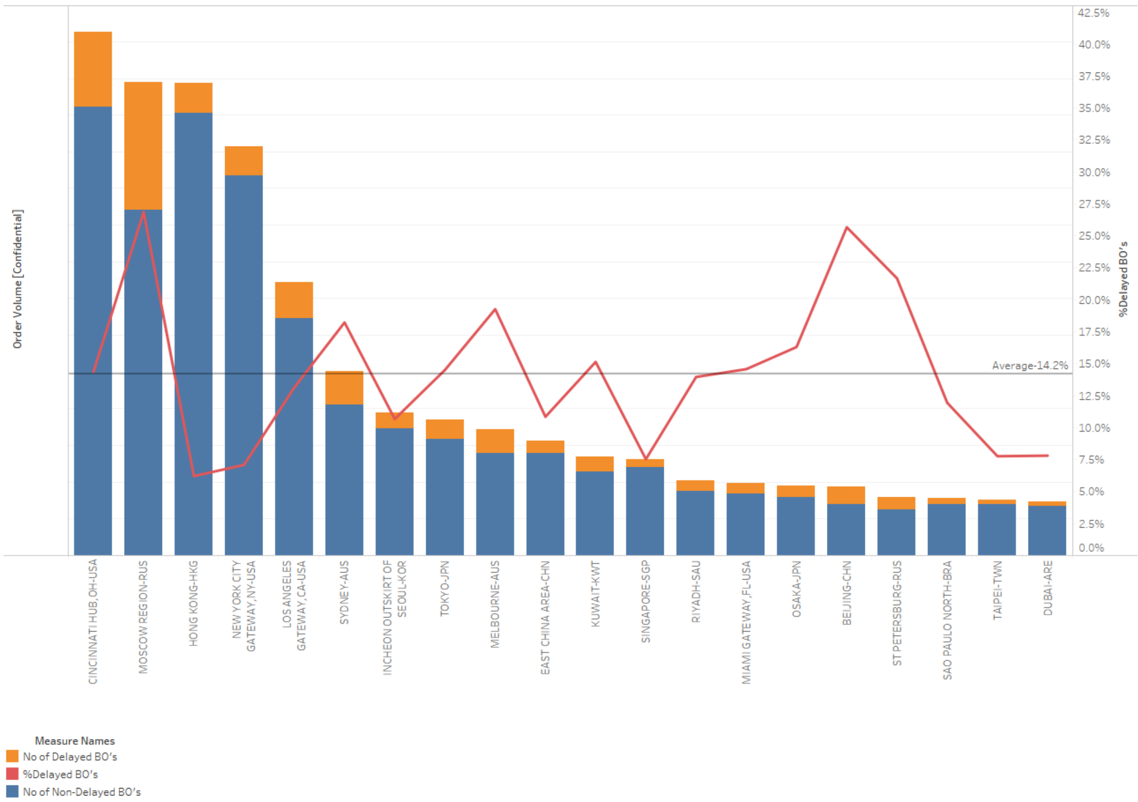

Figure 5.4: Delayed orders for top 20 gateways by order volume

Time between checkpoints (*Step1to5, Step6_Ori_Hub2, Step6_1stLeg*)

Figures 5.5, 5.6 and 5.7 depict how the three continuous variables are distributed by displaying a boxplot and histogram of each one and the difference of the grouped mean by "Is_Delayed" label (represented by dashed lines in histogram). Clearly, all of them behave differently when comparing against two levels of the target variable, being the fulfillment-related variable ("step1to5") the most representative one. A high number of outliers on this sample is also revealed.

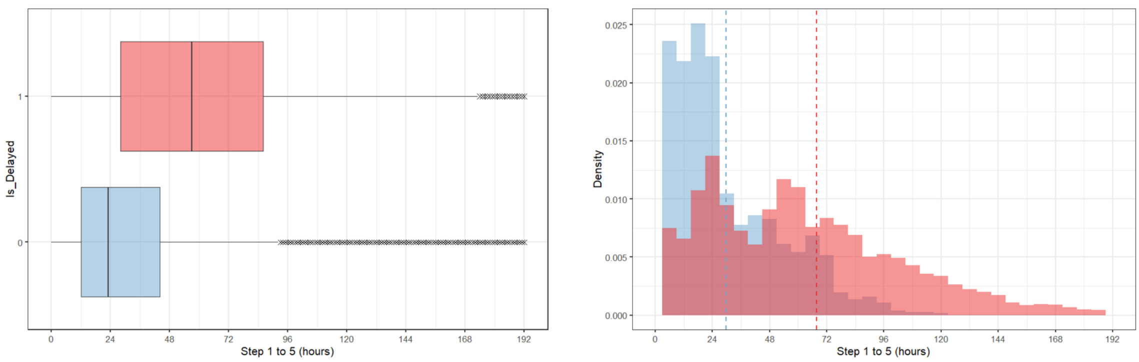

Figure 5.5: Statistical distributions for *Step1to5*

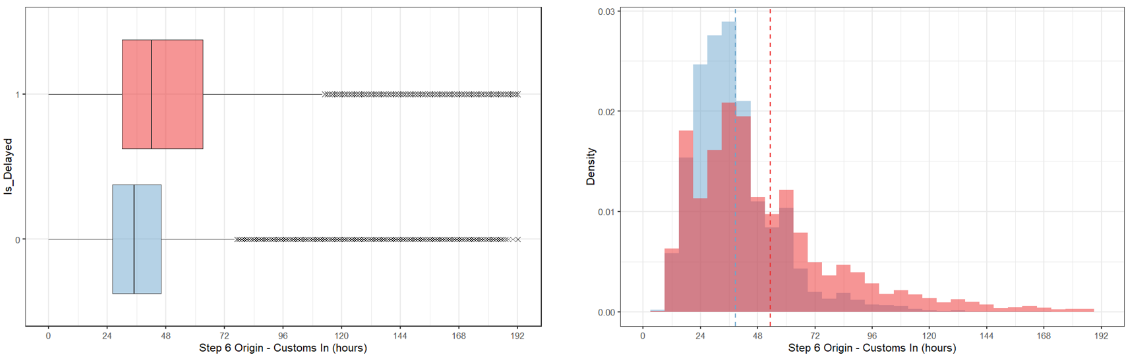

Figure 5.6: Statistical distributions for *Step6_1stLeg*

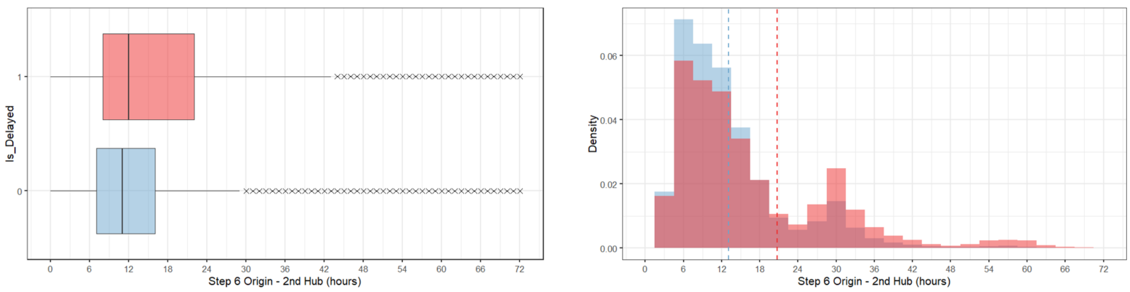

Figure 5.7: Statistical distributions for *Step6_Ori_Hub2*

Flags (*Has_exception, Is_peak, Is_DDP, Is_HighValue, Is_Delayed*)
Regarding binary attributes, insights exposed on figure 5.8 justify its creation. Boolean indicator of demand peaks presents the lowest impact on delayed shipments with a delayed orders rate difference (between peak season and non-peak season) as low as 0.8%. Such fact means that carrier network can accommodate high order volumes in short periods of time without compromising performance in a profound manner. In fact, these specific periods are planned in collaboration with carriers, which contributes to this effect. Regarding parcels whose combined item value was considered high, a low order volume (around 10.000 units) are affected. Nonetheless, the impact on delays is notorious with high-value delayed orders rate more than double compared to non-high-value rate. As explained previously, such effect is caused by a different and more rigorous customs clearance process for parcels whose value is above the considered threshold. In respect to commercial rules applied, parcels that are shipped to DDP-based countries have a delayed orders rate 6.4% lower than DAP ones. Such effect resides on the operational dissimilarities of the two methods already uncovered. Lastly, boutique orders containing at least one exception (exotic-based item, jewelry or watch) caused 2.4% more delay events than boutique orders containing non-sensitive items.

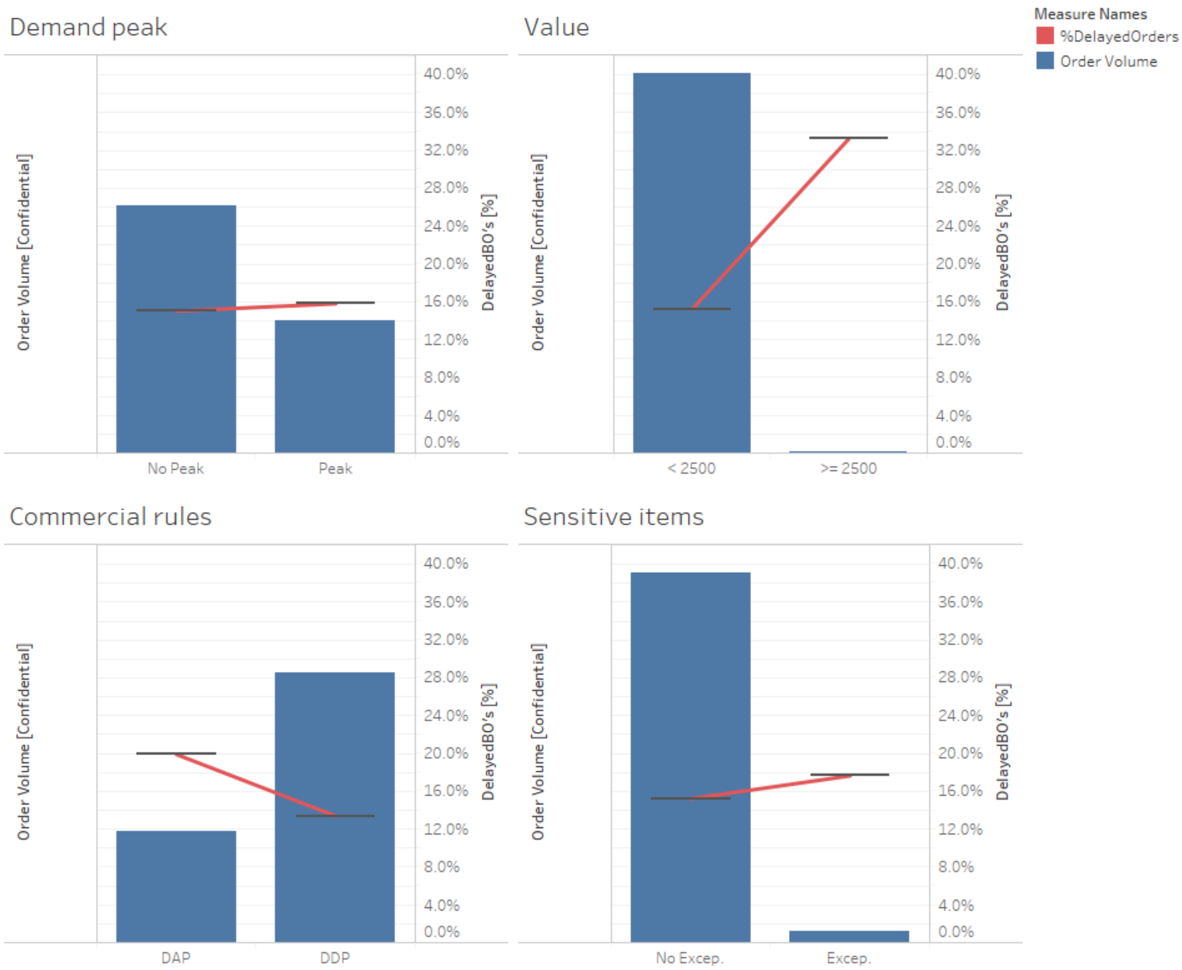

Figure 5.8: Binary attributes analysis

Other categorical attributes (*Exception_hours, sent_date_dw*)

As depicted on figure 5.9, a high number of orders are affected by two non-working days, which is understandable considering that lead time fluctuates around 4.5 days, in average. It is important to pinpoint that weekends are considered on EDD presented at checkout while holidays are ignored. This last fact is what explains the high delayed orders rate for one non-working day affected orders, mostly holidays. As result of backlog during weekends, a high order volume is present on Mondays. In addition, the residual order volume during Sunday and Saturday is explained by the fact of a low number of specific boutiques that remain operational and others that experience a weekend in other days of the week such as in some middle east countries. In what concerns the delayed orders rate on Tuesday, in could be explained by backlog effect during the weekend that was not completely absorbed on Monday.

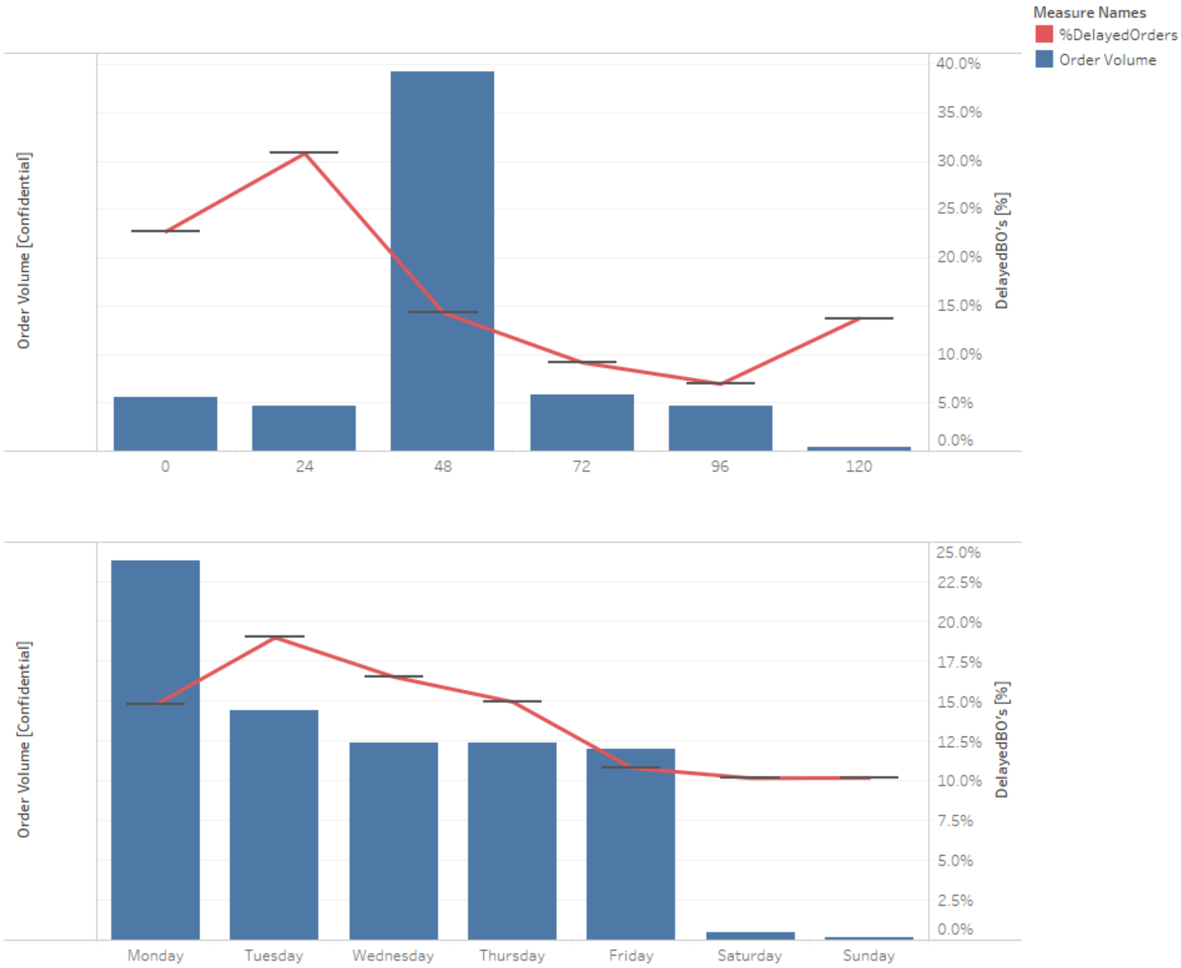

Figure 5.9: Categorical variables analysis

As expected, all variables have a meaningful impact on delayed parcels. Despite that, to produce reliable and powerful outcomes, several pre-processing tasks had to be conducted, which comprise the topic of the next section.

5.2 Data pre-processing

As exposed in section 2.4, data cleaning and understanding are essential to achieve reliable and accurate outcomes. Regarding tracking-related variables some intensive data treatment was done, mainly due to low structured nature of this datapoint. Briefly, each tracking event is composed by three main elements - (i) two-digit code, indicating type of event (such as “DF” stands for Departed Facility or “CR” means Customs Release); (ii) location (country and city level); (iii) date and time (on time zone where event occurred). In addition, other classical pre-processing tasks such as handling missing data or outlier removal were performed.

1. Time zones normalization

Consisted of converting each event timestamp to origin, destination time zones and to the basis of time zone designation (GMT+0) in order to fully capture three perspectives - customer, boutique and a standardized one.

2. Duplicated checkpoint locations removal

While exploring location variable, it was detected that in some facilities, parcels were scanned more than once which occurred due to internal hub movements such as going in and out of a set of building where different processes were conducted. Therefore, these intermediary movements on each hub were removed and the last timestamp stored.

3. Code-based location determination

This step is related to determining precisely when a parcel arrives and departures a certain type of facility. Analyzing event codes, it was possible to distinguish logistics hubs and customs facility. Therefore, other granularity level was created based on the gateway a certain parcel goes through on the destination country.

4. Order volume-based factors aggregation

Location-related variables can take a plethora of values, specially the one at destination city level, since that demand is more widespread than supply around the globe. Such high number would compromise the deployment of more robust and computationally expensive algorithms. With that in mind, factors aggregation depending on the granularity level was computed while aiming to lose as less information as possible and decrease computational requirements. The aim was to aggregate levels until a feasible number of factors that did not compromise future modelling tasks.

- City level - Adopting a pareto approach, variables whose combined order volume represented 80% of whole volume were maintained, while others were categorized as “Other”;
- Country level - Countries whose combined weight represented 1% of whole volume were classified as “Other”;
- Gateway level – Customs facilities whose cumulative weight constituted 5% were also classified as “Other”.

5. Handling missing data

Due to the fact that only hubs prior to destination country arrival were captured, a certain parcel could pass by multiple locations (one up to four) depending on its origin. As a result, whenever a certain hub, in an ascending order, was null, it means that the parcel did not go through such number of hubs before destination country. Such levels, interpreted initially as missing data, were allocated to another category “None” and time between them computed as zero. Any other missing values were residual and respective instances were removed.

6. Outliers detection and removal

Concerning continuous variables, values that were more than three standard deviations away from the mean, were labelled as potential outliers and hence, removed. Despite being a simple method, when it comes to large amounts of data, it is sufficiently adequate (Zumel, 2014).

7. Numeric variables correlation study

In order to understand linear correlation degree of numeric variables, Pearson correlation coefficient for each pair was computed and is displayed, on a matrix format, on figure 5.10. As demonstrated, no relationship presents a coefficient higher than 0.7, widely used threshold for determining a strong correlation. Therefore, variables under scrutiny were maintained in the model.

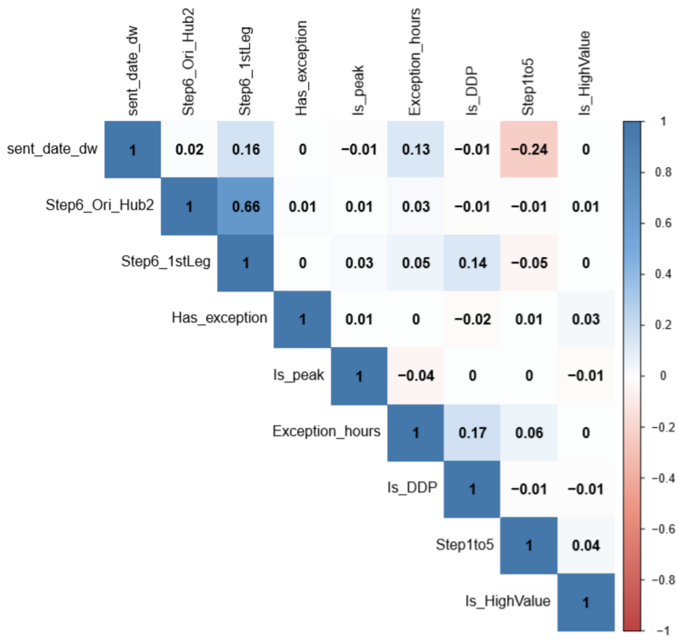

Figure 5.10: Correlation matrix for numeric variables

8. One-hot encoding

One of the predictive algorithms applied required the creation a set of dummy variables that represent the categorical attributes. Therefore, one-hot encoding process was employed to perform the mentioned task.

By exploring and understanding data and its traits, data pre-processing was accomplished on its fullness. After the mentioned steps, dataset is considered prepared for further modelling tasks.

5.3 Modelling

With intrinsic data value exposed as well as data preparation process performed, the following logical step consists on modelling. Select and apply the most suitable algorithms in a coherent manner while assessing performance represent the focus of this section.

5.3.1 Balancing, Splitting Method and Cross Validation

To begin with, it was evident in advance that the dataset would be imbalanced due to that fact that delayed orders represent only a fraction of the entire volume. Therefore, under-sampling method, widely deployed on big data problems, was applied, avoiding overfitting issues. Afterwards, dataset was split into training, validation and testing set. Usually, testing set selected is based on 20% to 40% of the original dataset. Due to the dataset size and computational power limitations, it was decided that 30% would represent the testing set while the remaining 70%, the training one. The mentioned partitioning was conducted randomly resulting in 561.826 data instances for training and 240.784 for testing. A data reduction process had to be performed due to the computational demanding nature of one the algorithm used, revealed on next section. As described, the training set is exploited by the algorithm to capture the inner relationships between variables and constantly validate it in an unbiased manner. To accomplish that, K-fold cross-validation technique was applied avoiding overfitting and hence, creating the foundations for reliable results. As exposed in section 2.4, ten-fold cross-validation is widely used and hence, it was selected. Finally, the testing set was used later to assess the model performance and tune parameters involved for specific algorithms.

5.3.2 Model Selection

While exploring algorithms extensively used for similar purposes, three were selected. The aim was to increase performance across more computationally-expensive algorithms and apply different learning principles to fully grasp the dataset potential.

Naïve-Bayes

A simple and computationally-friendly classifier that belongs to the family of "probabilistic classifiers" based on applying Bayes' theorem with strong (naive) independence assumptions between the features. The intrinsic goal is to identify the hypothesis with the maximum posteriori probability.

Logistics Regression

Likewise, based on statistical principles, logistic regression transforms its output using the logistic sigmoid function (also known as logit function) to return a probability value which can then be mapped to two discrete classes.

Extreme Gradient Boosting (XGBoost from now on)

Gradient boosting methods produces a prediction model in the form of an ensemble of prediction models, usually decision trees. It builds the model in a stage-wise fashion like and it generalizes them by allowing optimization of an arbitrary differentiable loss function (Nishida, 2017). Chen and Guestrin (2016) considers that success of XGBoost resides on its scalability in all scenarios. In addition, the author highlights some innovations including, novel tree learning algorithm for handling sparse data or a theoretically justified weighted quantile sketch procedure. Logically, this classifier is expected to outperform the first ones. Gradient boosting methods are currently being used in farfetch ecosystem in similar contexts showing good performance results which validates the application of such algorithm.

With predictive algorithms chosen, development process began by tailoring each algorithm to dataset while controlling corresponding parameters involved. Due to computational limitations, instances used for train XGBoost were reduced to 250.000.

5.3.3 Parameters tuning and model evaluation

After the selection process, it is essential to analyze model parameters used for the underlying learning process. The focus of this task was steered towards the most complex and multi-parameters algorithm, XGBoost. Chen and Guestrin (2016) divides overall parameters in three main categories. Firstly, general parameters are used to guide the overall functioning, while booster ones guide the individual tree (learning entity at each step). Lastly, learning task parameters are responsible for optimization procedures.

Regarding general, its properties are related to the type of model, in this context tree-based (instead of linear ones), running messages and parallel processing which enables a lower debug time. Such parameters alongside with learning task parameters, remain constant across multiple iterations. In what concerns tree-related booster parameters, multiple factors could be studied and fine-tuned. Considering computational limitations and the ones typically tuned across literature, four emerged as key to manage. Firstly, the parameter that controls the maximum number of iterations, which essentially means, the number of trees grown, named **nrounds**. Besides 100, default value, 200 was assigned to be used iteratively. Secondly, **max_depth** parameter defines the maximum depth of trees computed to avoid overfitting, as higher depths allow the model to learn very specific relations. The value was set to fluctuate around its default value (6), 3 and 10, range considered feasible (Jain, 2016). The same range was applied to **min_child_weight parameter** parameter, which determines the minimum sum of instance weight needed in a child. Succinctly, If the tree partition step results in a leaf node with the sum of instance weight less than pre-defined value, then the construction process will end further partitioning. The fourth one considered, denotes the fraction of columns randomly sampled for each tree (named **colsample_bytree**). Since that a higher number of features are considered and one-hot encoding was performed, the recommended and used range lies between 0.3 and 0.7 (Jain, 2016). As a result, in Appendix A, grid search for each model and corresponding metrics are revealed. In short, the optimal parameters that enables the best results are displayed on table 5.2. In addition, ROC curves and feature relative importance are exhibited in appendix B. Analyzing this last datapoint, similar conclusion regarding the most relevant features revealed in section 5.1.2. can be extracted and corroborates the exploratory analysis conducted. Such features comprise not only the most relevant one, “Step1to5”, but also levels “Russia” and “Honk Kong” on destination country, both addressed previously.

Table 5.2: Optimal parameters for both models

Models	nrounds	max_depth	colsample_bytree	min_child_weight
1st	200	10	0.6	3
2nd	200	10	0.7	6

Table 5.3: Performance metrics for first model

Metrics	Naïve-Bayes	Logistic Regression	XGBoost
Accuracy	76.93%	80.63%	82.71%
Precision	77.60%	80.99%	83.10%
Recall	75.71%	80.05%	**83.31%**
Specificity	78.15%	81.21%	82.10%
F-score	76.40%	80.52%	83.20%
AUC-ROC	84.93%	88.66%	91.37%

Table 5.4: Performance metrics for second model

Metrics	Naïve-Bayes	Logistic Regression	XGBoost
Accuracy	76.84%	81.64%	84.01%
Precision	76.23%	81.95%	85.03%
Recall	75.49%	81.12%	**85.54%**
Specificity	78.19%	82.13%	82.47%
F-score	75.86%	81.53%	85.28%
AUC-ROC	84.85%	89.60%	92.42%

Moving towards evaluation and comparison of models, in section 2.4.2., metrics for a model's performance present across literature, were described. Therefore, accuracy, precision, recall, specificity, F-score and AUC-ROC were computed allowing models evaluation. However, is it crucial to understand which metrics are most suitable on this specific context. Since that the end-goal is to create a proactive communication trigger notifying customs in case of a delay, true positive labels are more significant. Contrarily, negative labels, non-actionable on the scope of this project, become less relevant. In light of this, true positive recognition rate (also known as sensitivity or recall) was considered the pivotal metric to assess and compare models' performance. Table 5.3 and 5.4 depict the performance metrics for each of the algorithms and dataset variant exploited (first and second model as explained in section 5.1.1). As expected, best performance is achieved, in both models, using XGBoost for all metrics computed.

5.4 Impact analysis

Obtaining an accurate and reliable model is unequivocally, leveraged by the actionability of its outcomes. With that in mind, it is crucial to understand the impact of predictive algorithm from a business perspective. Essentially, this new communication trigger could result in customer-support related savings and customer experience improvements which can be measured based on metrics defined on delivery experience framework on section 4.1. The first component could be computed using the variation of tracking related Contacts Per Order (CPO), fundamental metric to assess customer satisfaction, widely used in industry, and cost to serve (labor and overhead costs, among others). Concerning the second member, expected variation of retention rate (RT) was used to assess delivery experience enhancements. Both components are indexed to true positive predictive ability, cases in which the new communication flow has a real impact. Such calculations are summarized on equation 5.1. One assumption consists on the fact that false positive would not impact none of the members neither positively nor negatively. Such impact also depends on the content of the communication flow. Notifying as a delay or a possible delay leads to two different customer interpretations which impact both components under scrutiny.

With both analytical and business perspective covered in detail, results are analyzed on the next chapter.

$$E_{impact} = [(|\Delta CPO_t| \cdot N_{BO} \cdot Cost_{CS}) + (\Delta RT \cdot AOC \cdot N_{users})] \cdot Recall \tag{5.1}$$

E_{impact} = Estimated financial impact
ΔCPO_t = Expected variation on tracking-related Customer Contacts per Order
N_{BO} = Number of boutique orders affected
$Cost_{CS}$ = Overall Customer Support cost per boutique order
ΔRT = Expected variation on Retention Rate
AOC (Average Order Contribution) = Average Order Value (AOV) deducting discounts and other decrements
N_{users} = Number of customers targetted

Besides quantitative outcomes, it is important to reflect about the whole customer experience qualitative dimension revamped, hard to fully measure. Such tailored and meaningful communication flows can truly create long-lasting customer relationships whose value extends beyond short-term financial gains. In addition, this initiative output fits on the overall company customer strategy, by boosting the immersive dimension of Farfetch ecosystem. With both analytical and business perspective covered in detail, results are analyzed on the next chapter.

Chapter 6

Discussion and future work

Across the present dissertation, several findings responsible for steering the workflow were unveiled. As important as results, is how the path to accomplish them was established. Customer experience, namely, delivery experience is a truly multi-dimensional topic where several domains interact determining customer satisfaction. Focusing on the post-order segment of the customer journey, the framework developed helped to capture a holistic overview and understand the impact of all delivery experience elements. Time-related factors revealed, combined with high tracking visibility leveraged by omnichannel-based communication flows and lastly, high control degree over the delivery process were considered the building block of post-order delivery experience. Such domains were materialized in the form of compliance, awareness and empowerment level, respectively. As imperative as revamping experience drivers is ensuring that customer interactions remain confined to Farfetch ecosystem. Not only creating, but also internalizing touchpoints is crucial to accomplish a single point of contact reality, a concept widely exposed across this dissertation. By narrowing the analysis to specific levels, the current situation and improvement points started to emerge. Benchmarking analysis proved a powerful tool not only to check whether external providers value proposition fulfilled the mapped requirements but also to fine-tune them and broaden the scope of alternative elements that could impact delivery experience. Avoiding tunnel vision by understanding how other e-commerce players are reacting to the disruptive area of last-mile delivery surely was helpful. After seizing all improvement opportunities and define vision over time, it was possible to outline workstreams that would deliver individual and complementary contributes. By guiding three workflows in parallel, the project created a sense of awareness inside the organization that facilitated its diffusion across stakeholders.

As a result of a problem-solving approach towards delivery experience enhancements, the need for a more complex analytical tool surfaced. In 2018, 19% of cross-border orders experience a delay and even with recent EDD improvement at checkout, the mentioned value still lies around 10%. Consequently, build the foundations for an innovative proactive communication flow revitalizing the awareness level, was faced as a vital aspect in a quest for the ultimate delivery experience. Therefore, the development of a data-driven predictive algorithm was thoroughly explained. It is important to emphasize the magnitude of pre-processing tasks undertaken. Despite time-consuming, important findings and insights were unveiled which allowed to be more knowledgeable about this specific supply chain segment. In addition, dataset preparation before modelling process is considered consensually across literature, a critical stage of knowledge-discovery. Simultaneously, tracking datapoint exploration proved to be helpful on decision-making processes in other projects. Specific events identification or the fact of being a standardized location-related data source at both country and city levels are evidences of its value-adding potential to the organization.

Subsequently, the developed classification algorithms showed good performance results, particularly the decision-trees based one, XGBoost. Feeding it with all data until parcel second logistic hub arrival (1^{st} model), it was possible to predict 83.31% delayed shipments (recall). Moving forward on the order journey, on the second data collection trigger (2^{nd} model), the equivalent metric increased to 85.54%. In addition, AUC-ROC values above 90% clearly confirm the quality of the models' predictions. Besides performance, it is crucial to reflect on other evaluation parameters. Being the most complex algorithm, XGBoost presented a running time much higher than logistic regression and naïve-bayes nevertheless, feasible. Despite that, dataset reduction and limiting parameters range for tuning had to be conducted due to the computationally-expensive nature of this gradient booster. Therefore, after this successful proof-of-concept, some improvements could be achieved by accessing additional computational power which would enable to process all data being generated and hone parameters tuning. Since the dataset was extracted from the all year of 2018, it is also crucial to adjust the mentioned algorithm to more recent instances in order to capture recent EDD improvements at checkout and other operational modifications, that directly impact the target variable. Moreover, other variables not available at Fafetch database, at the date of this dissertation, were labelled as potential predictors for modeling purposes. The clearance mode that a given parcel experiences when arrives at customs is an example of that. Furthermore, remote areas identification on supply and demand side could also reveal extremely helpful for determining delayed events. Regarding communication triggers, the algorithm could be easily escalated to more route instances by replicating the learning process resulting in a plethora of opportunities to reach the customer.

Over the short-term, having the algorithm as robust as possible, steps to merge this internal and outsourcing workstreams must be carried out. By taking advantage of near real-time tracking data and communication tools supplied by the external provider as well as, synergies involved, a dynamic automated communication flows should be set up. It is vital to ensure that customer is reached only once across the order journey avoiding any sort of friction. This final stage must be followed by a monitoring process to enable the comparison of real against estimated impact, whose measuring could be obtained considering metrics computed in section 5.4. Concerning other delivery experience enhancements powered by external providers, impacting compliance and awareness, it should be closely followed, ensuring that initial requirements are met. Regarding low-development internal workstream, guide internal initiatives towards envisioned changes in a continuous improvement manner, consists of a vital battlefront to fully revamping delivery experience.

Briefly, the present dissertation was developed adopting a customer-centric approach, proving that data-driven analytical solutions can contribute to drive customer engagement by creating new forms of interacting with customers in a seamless and meaningful way. On e-commerce universe, particularly on the disruptive area of last-mile delivery, such responsive ability is crucial to retain and delight consumers. These competitive advantages will be powered by the predictive model developed once near real-time data is available. Besides that, delivery experience enhancements envisioned are constantly being achieved through the external provider and internal low-development initiatives.

Bibliography

A. Lengnick-Hall, C. (1996). Customer contributions to quality: A different view of the customer-oriented firm. *Academy of Management Review*, 21.

Abu-Mostafa, Y. S., Magdon-Ismail, M., and Lin, H.-T. (2012). *Learning from Data*.

Alabi, K. (2017). Decentralization as a tool for securing ecommerce and business processes. *SSRN Electronic Journal*.

Amed, I., Balchandani, A., Beltrami, M., Berg, A., and Saskia Hedrich, F. R. (2019). The state of fashion 2019. Technical report, McKinsey Company.

Appleseed, J. (2017). Use 'delivery date' not 'shipping speed' – from ux research to implementation roadmap.

Billings, A. and Ho, N. (2017). Supply chains in the age of amazon.

Blair, I. (2018). What is a push notification? and why it matters?

Broillet, A., Dubosson, M., and Varone, S. (2019). Risk awareness from the consumer's side in luxury e-commerce: an empirical study.

Chang, D.-S. and Wang, T.-H. (2012). Consumer preferences for service recovery options after delivery delay when shopping online. *Social Behavior and Personality: an international journal*, 40(6):1033–1043.

Chen, T. and Guestrin, C. (2016). XGBoost. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16*. ACM Press.

Chong, A. Y.-L. (2013). Understanding mobile commerce continuance intentions: An empirical analysis of chinese consumers. *Journal of Computer Information Systems*, 53(4):22–30.

Clark, G. and Johnston, R. (2006). *Service Operations Management: Improving Service Delivery*, volume 17.

D'arpizio, C. and Levato, F. (2018). The future of luxury: a look into tomorrow to understand today. Technical report, Bain Company.

Du, H., Lederer, A., and Wu, J. (2010). *Trust in Electronic Commerce*. IGI Global.

Dvir, D., Raz, T., and Shenhar, A. J. (2003). An empirical analysis of the relationship between project planning and project success. *International Journal of Project Management*, 21(2):89–95.

Falk, M. and Hagsten, E. (2015). E-commerce trends and impacts across europe. *International Journal of Production Economics*, 170:357–369.

Géron, A. (2017). *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media.

Han, J. (2011). *Data Mining: Concepts and Techniques (The Morgan Kaufmann Series in Data Management Systems)*. Morgan Kaufmann.

Hand, D. J., Mannila, H., and Smyth, P. (2001). *Principles of Data Mining (Adaptive Computation and Machine Learning)*. A Bradford Book.

Jacobs, F. R. and Chase, R. B. (2017). *Operations and Supply Chain Management*. McGraw-Hill Education.

Jain, A. (2016). Complete guide to parameter tuning in xgboost with codes in python. Technical report, Analytics Vidhya.

Joerss, M., Neuhaus, F., and Schröder, J. (2016). How customer demands are reshaping last-mile delivery.

Johnson, R. A. and Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis (6th Edition)*. Pearson.

Kayikci, Y. (2019). E-commerce in logistics and supply chain management. pages 1015–1026.

Kim, E. (2019).

Kim, S. and Park, H. (2013). Effects of various characteristics of social commerce (s-commerce) on consumers' trust and trust performance. *International Journal of Information Management*, 33(2):318–332.

Kini, A. and Choobineh, J. (1998). Trust in electronic commerce: definition and theoretical considerations. 4:51–61 vol.4.

Ko, E., Costello, J. P., and Taylor, C. R. (2019). What is a luxury brand? a new definition and review of the literature. *Journal of Business Research*, 99:405–413.

Kourouthanassis, P. E. and Giaglis, G. M. (2012). Introduction to the special issue mobile commerce: The past, present, and future of mobile commerce research. *International Journal of Electronic Commerce*, 16(4):5–18.

Lambert, D. M., Cooper, M. C., and Pagh, J. D. (1998). Supply chain management: Implementation issues and research opportunities. *The International Journal of Logistics Management*, 9(2):1–20.

Laudon, K. C. and Traver, C. G. (2018). *E-Commerce 2017: Business, Technology, and Society*. Pearson.

Lee, B. (2007). Consumer attitude toward virtual stores and its correlates. *Journal of Retailing and Consumer Services*, 14(3):182–191.

Lee, H. L. (2004). The triple-a supply chain. *Harvard Business Review*.

Lemon, K. N. and Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing*, 80(6):69–96.

Lin, Y., Luo, J., Cai, S., Ma, S., and Rong, K. (2016). Exploring the service quality in the e-commerce context: a triadic view. *Industrial Management & Data Systems*, 116(3):388–415.

Löfgren, M., Witell, L., and Gustafsson, A. (2011). Theory of attractive quality and life cycles of quality attributes. *The TQM Journal*.

McAfee, A. and Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review*.

Meena, S., Wu, S., O'Grady, M., and Kumar, S. (2017). Forrester analytics: Online fashion retail forecast, 2017 to 2022. Technical report, Forrester Analytics.

Mehrjoo, M. and Pasek, Z. J. (2014). Impact of product variety on supply chain in fast fashion apparel industry. *Procedia CIRP*, 17:296–301.

Molla, A. and Licker, P. (2001). E-commerce systems success: An attempt to extend and respecify the delone and maclean model of is success. *J. Electron. Commerce Res.*, 2:131–141.

Nishida, K. (2017). Introduction to extreme gradient boosting in exploratory. Technical report, Medium.

Olivier Abtan, Antonio Achille, J.-M. B. Y. K. V. L. A. M. A. M.-P. S. W. (2014). Shock of the new chic: Dealing with new complexity in the business of luxury. *Boston Consulting Group*.

Ono, A., Nakamura, A., Okuno, A., and Sumikawa, M. (2012). Consumer motivations in browsing online stores with mobile devices. *International Journal of Electronic Commerce*, 16(4):153–178.

Opresnik, D. and Taisch, M. (2015). The value of big data in servitization. *International Journal of Production Economics*, 165:174–184.

Otaris (2018). Data analysis, modeling and reporting.

Ramaswamy, V. (2009a). Leading the transformation to co-creation of value. *Strategy & Leadership*, 37(2):32–37.

Ramaswamy, V. (2009b). Leading the transformation to co-creation of value. *Strategy & Leadership*, 37(2):32–37.

Rangel, A. (2018). An investigation of alignment of time-to-ship offerings and consumer value in ecommerce.

Ruiz, A. (2017). The 80/20 data science dilemma.

Singh, J. and Crisafulli, B. (2016). Managing online service recovery: procedures, justice and customer satisfaction. *Journal of Service Theory and Practice*, 26(6):764–787.

Sola, J. and Sevilla, J. (1997). Importance of input data normalization for the application of neural networks to complex industrial problems. *IEEE Transactions on Nuclear Science*, 44(3):1464–1468.

Statistica (2018). Retail e-commerce sales worldwide from 2014 to 2021.

Stiller, B., Bocek, T., Hecht, F., Machado, G., Racz, P., and Waldburger, M. (2010). Mobile Systems IV. Technical report, University of Zurich, Department of Informatics.

Verhoef, P. C., Kannan, P., and Inman, J. J. (2015). From multi-channel retailing to omni-channel retailing. *Journal of Retailing*, 91(2):174–181.

Witell, L., Löfgren, M., and Dahlgaard, J. (2013). Theory of attractive quality and the kano methodology – the past, the present, and the future. *Total Quality Management Business Excellence*, 24.

Xiang, L., Zheng, X., Lee, M. K., and Zhao, D. (2016). Exploring consumers' impulse buying behavior on social commerce platform: The role of parasocial interaction. *International Journal of Information Management*, 36(3):333–347.

Yu, S., Hudders, L., and Cauberghe, V. (2018). Are fashion consumers like schooling fish? the effectiveness of popularity cues in fashion e-commerce. *Journal of Business Research*, 85:105–116.

Zumel, N. (2014). *Practical Data Science with R*. Manning.

Appendix A

Grid search for parameters

Table A.1: Grid search for first model

max_depth	colsample_bytree	min_child_weight	nrounds	Accuracy	AccuracySD
3	0.5	3	100	0.798128521	0.003742072
3	0.5	6	100	0.797728538	0.003114579
3	0.5	10	100	0.797464232	0.004304628
3	0.6	3	100	0.797899954	0.00418413
3	0.6	6	100	0.797849937	0.004531771
3	0.6	10	100	0.798085671	0.003615858
3	0.7	3	100	0.798349951	0.003870877
3	0.7	6	100	0.798171385	0.003781807
3	0.7	10	100	0.798242791	0.004373216
6	0.5	3	100	0.813542808	0.003516857
6	0.5	6	100	0.813614234	0.003442617
6	0.5	10	100	0.812978511	0.003887595
6	0.6	3	100	0.813121396	0.003320049
6	0.6	6	100	0.813007095	0.003616347
6	0.6	10	100	0.812942796	0.003870678
6	0.7	3	100	0.812978522	0.003834018
6	0.7	6	100	0.812792806	0.003417661
6	0.7	10	100	0.812271389	0.003348345
10	0.5	3	100	0.821457113	0.003210418
10	0.5	6	100	0.820321413	0.002813538
10	0.5	10	100	0.820171402	0.003750654
10	0.6	3	100	0.821114264	0.003719426
10	0.6	6	100	0.820949973	0.003032052
10	0.6	10	100	0.820664265	0.003099159
10	0.7	3	100	0.821699971	0.003089911
10	0.7	6	100	0.821335664	0.003391586
10	0.7	10	100	0.820185663	0.003712158
3	0.5	3	200	0.805392810	0.004161069
3	0.5	6	200	0.805328538	0.003470518
3	0.5	10	200	0.805464255	0.004192341
3	0.6	3	200	0.805099948	0.004210151
3	0.6	6	200	0.805971381	0.004487967
3	0.6	10	200	0.805299956	0.003999659
3	0.7	3	200	0.804807105	0.003741279
3	0.7	6	200	0.805307091	0.004013291
3	0.7	10	200	0.805564233	0.003901238
6	0.5	3	200	0.818692819	0.003170012
6	0.5	6	200	0.818778550	0.003551501
6	0.5	10	200	0.818142828	0.003440959
6	0.6	3	200	0.818442816	0.002963023
6	0.6	6	200	0.818492830	0.003368979
6	0.6	10	200	0.818357102	0.003718313
6	0.7	3	200	0.819014259	0.003823388
6	0.7	6	200	0.818714250	0.003501314
6	0.7	10	200	0.817985666	0.003303733
10	0.5	3	200	0.823257108	0.003140175
10	0.5	6	200	0.823007113	0.00307073
10	0.5	10	200	0.822778529	0.003335753
10	0.6	3	200	0.823699972	0.003626148
10	0.6	6	200	0.823371413	0.003060443
10	0.6	10	200	0.822914256	0.003077802
10	0.7	3	200	0.823514254	0.002650174
10	0.7	6	200	0.822992819	0.00320244
10	0.7	10	200	0.822764238	0.003109409

Table A.2: Grid search for second model

max_depth	colsample_bytree	min_child_weight	nrounds	Accuracy	AccuracySD
3	0.5	3	100	0.813398603	0.00401815
3	0.5	6	100	0.81394146	0.004129615
3	0.5	10	100	0.813320038	0.003681715
3	0.6	3	100	0.813777178	0.004361659
3	0.6	6	100	0.813934315	0.004160565
3	0.6	10	100	0.813627179	0.00422549
3	0.7	3	100	0.814198617	0.003737477
3	0.7	6	100	0.814041487	0.003847524
3	0.7	10	100	0.81408431	0.004474444
6	0.5	3	100	0.831148689	0.004199668
6	0.5	6	100	0.830662953	0.004310077
6	0.5	10	100	0.830827259	0.004411222
6	0.6	3	100	0.831405792	0.00440463
6	0.6	6	100	0.831084394	0.004162649
6	0.6	10	100	0.830884425	0.003861126
6	0.7	3	100	0.831548672	0.004443372
6	0.7	6	100	0.831377284	0.004163937
6	0.7	10	100	0.83150585	0.003956797
10	0.5	3	100	0.837234422	0.004569314
10	0.5	6	100	0.836863006	0.003868984
10	0.5	10	100	0.836577269	0.004359408
10	0.6	3	100	0.837684465	0.004127286
10	0.6	6	100	0.836991582	0.003939721
10	0.6	10	100	0.837477289	0.004028999
10	0.7	3	100	0.837091582	0.003888437
10	0.7	6	100	0.837905874	0.003928745
10	0.7	10	100	0.837163025	0.003404162
3	0.5	3	200	0.821548649	0.003892369
3	0.5	6	200	0.821377215	0.004168288
3	0.5	10	200	0.821470081	0.004091017
3	0.6	3	200	0.821427212	0.004617381
3	0.6	6	200	0.82149864	0.004243909
3	0.6	10	200	0.821448657	0.004035676
3	0.7	3	200	0.821634365	0.004190788
3	0.7	6	200	0.821605798	0.004120716
3	0.7	10	200	0.821520043	0.004516306
6	0.5	3	200	0.835391565	0.004359274
6	0.5	6	200	0.835405867	0.004368262
6	0.5	10	200	0.83552729	0.004398036
6	0.6	3	200	0.835627284	0.004237121
6	0.6	6	200	0.835591567	0.00450523
6	0.6	10	200	0.835105863	0.004308139
6	0.7	3	200	0.835684424	0.004791443
6	0.7	6	200	0.835541574	0.004265103
6	0.7	10	200	0.835577276	0.004489694
10	0.5	3	200	0.838584421	0.004954541
10	0.5	6	200	0.838463035	0.003711646
10	0.5	10	200	0.838141571	0.004303372
10	0.6	3	200	0.838741617	0.004118882
10	0.6	6	200	0.83839159	0.004130252
10	0.6	10	200	0.838705862	0.004587247
10	0.7	3	200	0.838255869	0.00378679
10	0.7	6	200	0.838984436	0.004089124
10	0.7	10	200	0.838255876	0.004100488

Appendix B

ROC Curve and Relative Feature Importance

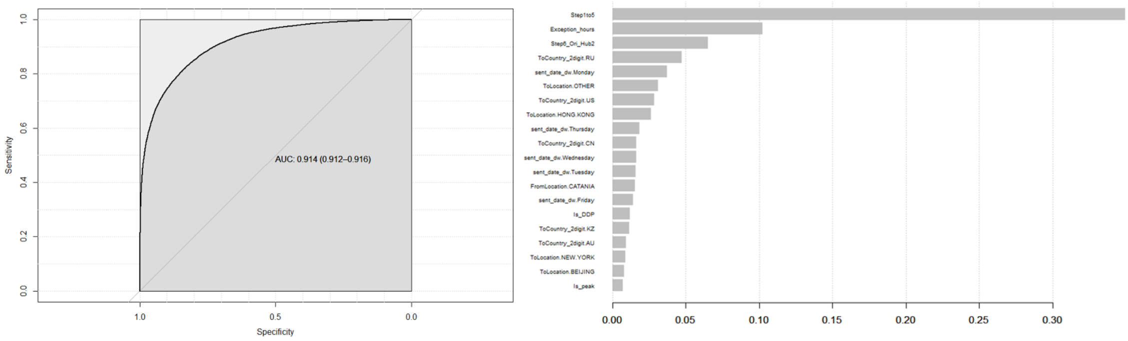

Figure B.1: ROC curve and overview over feature importance for first model

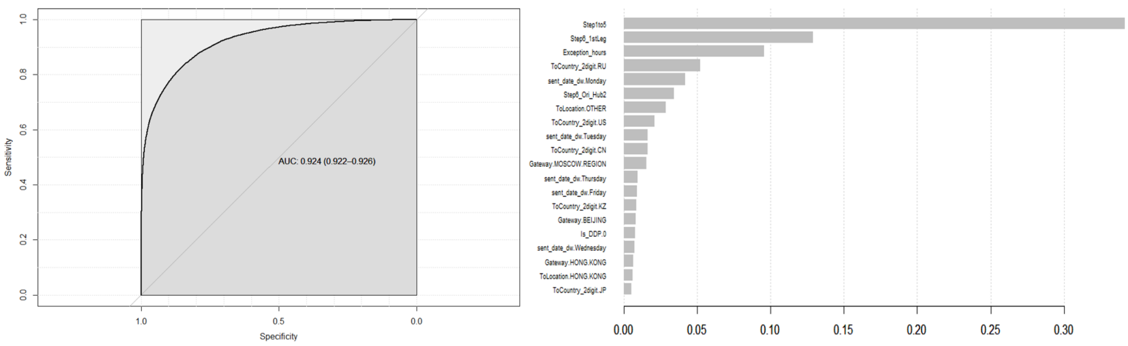

Figure B.2: ROC curve and overview over feature importance for second model