Returns in the luxury fashion e-commerce: predicting and understanding their impact in the customer experience

José João Guimarães Fernandes

Dissertação de Mestrado

Orientadores na FEUP:

Prof. Jorge Daniel Grenha Luís Teixeira e Prof. Vera Lucia Miguéis Oliveira e Silva

FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Mestrado Integrado em Engenharia e Gestão Industrial

2019-07-01

Abstract

Returns are one of the fundamental components of any e-commerce company. Typically, the costs associated with reverse logistics represent one of the biggest financial expenses for those companies. For customers, the impact of returns is much more significant and goes beyond the return process itself. Returns and return policies affect customers in every phase of the purchase journey, influencing their perceptions and attitudes towards companies.

The project was developed within the biggest online luxury fashion marketplace. In this company, around 20% of all orders are returned. Along the years, the growth of sales combined with a lenient return policy is causing an increase in the company's return rate. Hence, one of its main priorities is to understand and mitigate returns, which was the focus of this dissertation.

The first step was a holistic analysis of the customer experience. To do so, the company processes and the feedback provided by customers were studied. Aside from identifying the main pain points along the journey, it was found that more than 65% of the returns happen due to *Size and Fit* issues. Those insights highlighted the importance of the available product information, for fashion e-commerce websites.

The second phase of the project was the development of an algorithm capable of predicting the probability of an order being returned, based on *a priori* information. Given the large volume of accessible data and the intention of predicting returns, a machine learning approach was followed. The model provided valuable insights and created the opportunity for several initiatives that were identified and discussed.

Finally, after analysing the results of the model and given the customer feedback, the project addressed one of the friction points in the customer journey - product information. The goal was to support the definition of a strategy to implement a content-related initiative, which is expected to improve the information provided to the customers, increasing their trust and reducing the probability of returns. Since one of the customers' main difficulties was identifying their appropriate size, this initiative focused on obtaining specific product measurements for every size available on the website. In the final sections, the estimated impacts and required actions are fully detailed.

Resumo

As devoluções são uma das componentes principais de qualquer empresa de comércio *online*. Tipicamente, os custos associados com a logística inversa representam uma das maiores despesas destas empresas. Por outro lado, para os clientes, o impacto das devoluções é muito mais significativo e vai para além do processo de devolver um produto. As devoluções e respetivas políticas afetam os clientes em todas as fases do processo de compra, influenciando as suas perceções e atitudes relativamente às empresas. Este projeto foi desenvolvido no maior marketplace de moda de luxo *online*. Nesta empresa, cerca de 20% de todas as encomendas acabam devolvidas. Ao longo do tempo, o crescimento das vendas, associado a uma política de devoluções permissiva, está a ser responsável por um aumento da taxa de devolução da empresa. Assim sendo, uma das suas prioridades passa por perceber e mitigar as devoluções, o que constituiu o foco desta dissertação.

O primeiro passo foi uma análise holística da experiência de cliente. Para tal, foram analisados os processos da empresa, bem como o *feedback* dado pelos clientes. Para além da identificação dos principais desafios ao longo de todo o processo, descobriu-se que mais de 65% das devoluções acontecem devido a problemas de tamanho e *fitting* dos artigos. Estas descobertas sublinham a importância da informação acerca dos produtos que é disponibilizada aos clientes em plataformas de venda de moda *online*.

A segunda fase teve por base o desenvolvimento de um algoritmo capaz de prever a probabilidade de uma encomenda ser devolvida, com base em imformação *a priori*. Dado o elevado volume de dados e a intenção de prever as devoluções, optou-se por uma abordagem *machine learning* com o desenvolvimento de um modelo de classificação. O modelo permitiu descobertas interessantes e criou a oportunidade para algumas iniciativas que foram identificadas e discutidas no documento.

Finalmente, depois de analisados os resultados do modelo e com base no *feedback* dos clientes, o projeto abordou um dos problemas identificados na experiência de compra – a informação sobre os produtos. O objetivo foi dar suporte à definição de uma estratégia para implementar uma iniciativa relacionada com o conteúdo do website, que se espera ser capaz de melhorar a informação, aumentando a confiança dos clientes e reduzindo a probabilidade de devoluções. Uma vez que uma das maiores dificuldades da compra passa por escolher o tamanho adequado dos produtos, a iniciativa focou-se em obter medidas específicas para todos os tamanhos disponíveis na plataforma. No final, são explicados os impactos bem como as ações necessárias no futuro.

Acknowledgements

First of all, I would like to express my gratitude to my supervisors at Feup, to Prof. Jorge Teixeira for his constant support and for helping me find the correct path to follow and to Prof. Vera Migueis, for her analytical guidance and encouragement.

To Frederico Alves, my supervisor at Farfetch, for sharing his passion for the project and never-ending ambition. To Inês Pereira, for being a such an inspiration and for all the wise advice.

To Margarida, for always being there, never letting me give up.

And finally, to my parents and brother whose unconditional love and support helped me get through this challenge.

Contents

Acronyms

AOV	Average Order Value
AUC	Area Under the Curve
AWB	Airway bill
BFC	Brand Family Category
CRISP-DM	Cross Industry Standard Process for Data Mining
DACH	Germany, Austria and Switzerland (German-speaking Europe)
IPO	Initial Public Offering
NPS	Net Promoter Score
PDP	Product Display Page
PLP	Product List Page
ROC	Receiver Operating Characteristics
SQL	Structured Query Language
UK	United Kingdom
US	United States
VIP	Very Important Person

List of Figures

List of Tables

1 Introduction

The luxury fashion market is in clear expansion and currently, the share of goods sold online accounts for just 8% of this 254 billion market. However, it is expected to triple by 2025 (Achille, Remy, and Machessou 2018). With such fast growth in terms of sales, fashion e-commerce companies must increasingly deal with the fact that a high percentage of their sales ends up being returned by customers. Hence, mitigating this phenomenon creates an opportunity for companies to increase service levels and raise profitability.

Since the project was developed in a business environment, the document starts with a brief presentation of the company in section 1.1. After that, the motivation for the development of the project is detailed in section 1.2 and its main objectives are highlighted in section 1.3.

Due to the fact that Farfetch deals with large volumes of data, it was decided to tackle the project by following a data mining methodology, i.e., Cross-Industry Standard Process for Data Mining (CRISP-DM) (Chapman et al. 2000), that is explained in detail in section 1.4. Finally, in section 1.5, the overall structure of the document is described.

1.1 Company Description

Farfetch, a luxury fashion online marketplace, was founded in 2008 by José Neves in London and has become the major player of the luxury online market.

For the past years, the company has been experiencing tremendous growth with the emergence of other business units that have transformed the company into a platform business with the goals of being a “commerce enabler and an innovation partner for the future”. The recently created business units are: *Black and White*, responsible to conceive tailor-made e-commerce solutions for luxury brands or boutiques; *Fashion Concierge*, allowing VIP clients to access nearly every possible item, even if not currently available on Farfecth’s platform; *Store of the Future*, aiming to transform the customer shopping experience through online and offline interactions; and *Browns*, an iconic British boutique, acquired by Farfetch in 2015, that works as an offline extension of its business model and as a way of testing possible initiatives. However, the core business of the company is still its marketplace, Farfetch.com, that connects supply and demand for luxury fashion products.

Upon its creation, this unique business model was able to suppress the specific needs of the market. On one side, many of the partner boutiques that are now working with Farfetch did not have the dimension neither the know-how to implement an e-commerce strategy. On the other side, high-end luxury clients were not capable of buying luxury goods anytime and anywhere. Hence, the company was able to establish itself allowing its partners to access a huge customer base, and its clients to have the possibility of having a multi-brand luxury experience.

In addition, what makes the company unique is the fact that it is capable of this without being responsible for the stock available on the website or any transportation systems. Once a customer orders a product, it is packed and prepared by the partner at its stock points, and finally picked and delivered to the customer by a third-party carrier.

According to the founder of the company, despite the introduction of new software, boutiques still work very similarly to how they did in the 90s, and the majority of them is completely isolated from the digital world. In his words, there will be a day in which the distinction between online and offline sales will not matter, after all, it is just sales. The aim of Farfetch is to empower the industry, being disruptive and able to combine both worlds, bringing unique Boutiques together and combining their items into a single platform available worldwide.

Considering the report of Dauriz, Remy, and Sandri (2014), mono-brands still have the lion's share but multi-brand marketplaces, like Farfetch, have the highest potential for growth, as customers increasingly seek multi-brand experiences.

With such high potential, Farfetch has established as the global leader of the online luxury fashion market, serving customers from 190 countries and working with more than 1000 partners, allowing the company to have the widest range in terms of product offering.

The most recent achievements of the company were the New York-based initial public offering (IPO), allowing the company to enter the stock exchange market, the acquisition of Stadium goods, an American streetwear brand, and the partnership with Harrods, the world's most famous department store.

1.2 Motivation for the Project

With the fast development of e-commerce, one of its major challenges is to bring the online experience closer to the physical experience in store. This challenge acquires additional importance for luxury fashion items, where the involvement, both at an emotional and informational level about the products, is much more significant. Whenever there is a mismatch between customer expectations and the products received, customers are forced to return their purchases, causing a pain point in the purchase journey. This journey is part of the customer experience, which comprises every interaction between consumers and companies that takes place before, during and after a purchase.

Returns are one of the fundamental components of any retail and e-commerce operation. Currently, Farfetch has to face the fact that more than 20% of the items bought end up being returned. On top of that, with the goal "Amaze Customers" [1], the company developed a very lenient return policy that implies the acceptance of all returns, completely free of charge on the customer side. Because of that, returns represent one of the biggest financial burdens for Farfetch. Analysing data from 2014 to 2018, it is possible to conclude that, with the growth of sales, the return rate (i.e. the ratio between the number of returned orders and the total number of orders) of the company is also increasing (Figure 1).

[1] One of the Farfetch values. *Be Brilliant, Be Revolutionary, Think Globally, Todos Juntos, Be Human* and *Amaze Customers* .

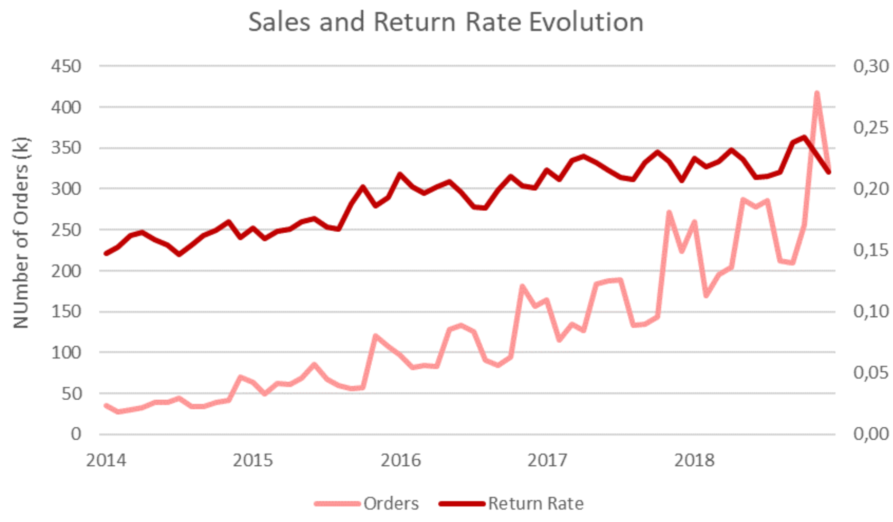


Figure 1 - Farfetch sales and return rate evolution from 2014-2018

Hence, the main objective of this project was to understand returns root causes and develop a strategy to predict and mitigate this phenomenon. This will provide an understanding of customers' expectations and, after that, the development of initiatives that can, not only impact the customer experience but also the profitability of the company.

Additionally, it is important to refer that the effect of returns on customer experience goes beyond the moment where there is a need to return a product. Returns and return policies, impact all phases of the purchase journey and can deeply influence the attitude of customers towards the company.

Given the volume of data available, it was decided to follow a data mining approach in order to develop a predictive model that would be the basis of some initiatives to mitigate returns. Being able to predict and develop incentives to act on this friction moment of the customer journey, will substantially reduce the returns related costs, enhance customer experience and promote higher loyalty and repeated purchase behaviours.

Moreover, when diving into the reasons why Farfetch customers were returning their products, it was found that approximately 65 % of the returns happened due to what is called *Size and Fit* issues (e.g. "product is too big"). This means that product information was not matching customer expectations and, more than not being congruent with the luxurious identity of the company, this gap in terms of information was being responsible for the majority of the returns of customers that were being wrongly informed about the characteristics of the products they were buying.

Potential improvements in terms of product content, not only impact the luxurious perception that the customers have about the company, but also contribute to its profitability due to the way they can affect the conversion and return rates. However, being a marketplace without any access to the items that are sold through the platform creates additional challenges and requires the company to develop a plan, together with its partners, to efficiently gather all the possible product information.

Hence, the second phase of the project focused on providing the adequate support to roll out a strategy to improve the content displayed on Farfetch website, as it proved to be an extremely powerful tool to mitigate returns and provide a better customer experience.

1.3 Project Goals

Given the motivation for the project, its main goals were:

- Understanding how returns impact customer experience;
- Analysing the customer perceptions and the major pain points of their purchase journey with the company;
- Developing a predictive model that is able to predict if certain orders are going to be returned;
- Proposing initiatives to leverage the predictive capacity of the model with the goal of mitigating returns;
- Giving support to an initiative that is capable to tackle one of the major pain points identified by the Farfetch customers.

1.4 Methodology

Given the scope of the project and its goal of explaining returns and developing initiatives that impact customer experience on an online marketplace, it was decided to follow a data mining methodology - Cross Industry Standard Process for Data Mining (CRISP-DM) (Chapman et al. 2000). This framework is widely used in the industry and encompasses a series of steps to be followed when tackling a data mining project. Its main steps can be seen in Figure 2. This choice is supported by the fact that Farfetch collects a huge amount of data related to every order on the platform. Hence, to be able to deal with this volume of data and obtain valuable insights a data mining methodology was deemed as useful.

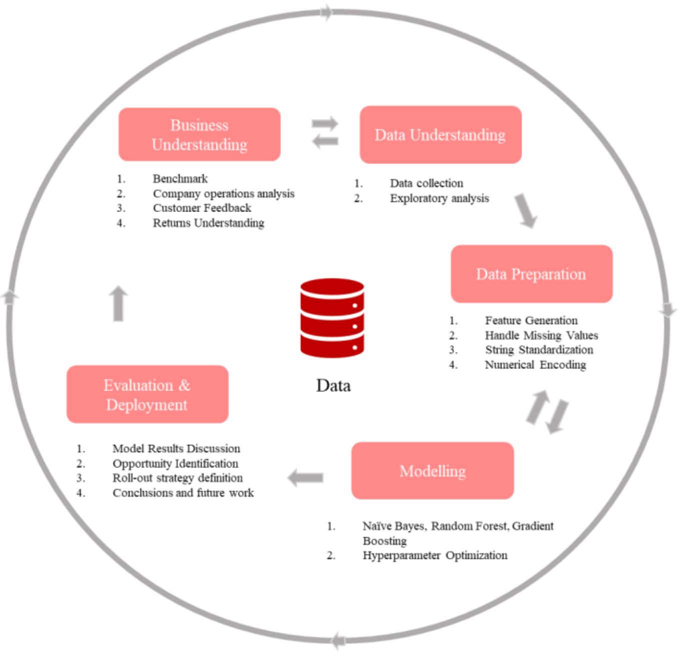


Figure 2 – CRISP-DM framework, adapted from (Chapman et al. 2000)

Following the CRISP-DM framework, the first step involved getting an overview of the Luxury market and its main players, comparing them to Farfetch. After that, it was important to have a clear idea of the company operations and tools to be used in the subsequent phases. This was achieved through several inductions and sharing sessions with internal stakeholders of the company. As an additional step in this phase, it was decided to explore several sources of customer feedback. This has been done following the principles of customer centricity of the company, aiming to understand what customer perceptions and feelings are, concerning their experience with Farfetch.

The next main step was to dive into the data structure of the company and understand, among the several sources, which should be considered to collect all the required information. This has been one of the most time-consuming tasks as there was a need to go through the learning curve of SQL and combine information stored in two main databases. First, the official company system, *SQL server management studio*, and *Google BigQuery*[2], a web-based solution from Google that significantly reduced the processing time.

After collection, the majority of time was spent in data cleaning and preparation processes to ensure that the data was suited to be fed into the models.

During the project, the applied modelling techniques were Naïve Bayes, Random Forest and Gradient Boosting using decision trees as a base predictor. Throughout the analysis, the models are compared and evaluated, with the main criteria being the recall, that will be explained later on. After a first iteration, the hyperparameters of the models were optimized and once again, the results were discussed and compared.

The final two steps of the CRISP-DM methodology are evaluation and deployment. For time reasons, there was no possibility to implement the modelling approach in a real environment. Nevertheless, in the last phase, the results of the models were analysed, discussing their applicability for the company. Moreover, based on the results and the business understanding phase, it was decided to support the planning of the roll-out strategy for a content-related initiative.

In the final stage, all the insights of the project were discussed as well as the potential opportunities for future actions in the company.

1.5 Document Structure

This dissertation is organised in seven chapters whose outline is the following:

- Chapter 1 introduces the project, starting with the company presentation, the main motivation for the project and its main objectives, as well as a high-level description of the methodology used to pursue it;
- Chapter 2 provides the theoretical background that sustains most of the assumptions made during the project. It is organized into three main areas. First, the luxury fashion market and what makes it so unique. Second, an analysis of customer experience and its main phases to frame the return process as one of the most crucial parts of the customer journey. And finally, machine learning techniques, a section that covers the models and techniques that were applied during the project;
- Chapter 3 corresponds to the step of Business Understanding of the CRIPS-DM methodology. Hence, it starts with a high-level overview of the luxury fashion market. Then, all the company operations and ordering processes are explained, both from the company and customer perspective. After that, the feedback of the customer is analysed focusing on the impacts of the main pain points identified. Finally, all the return procedures and their implications are discussed;

[2] Google Big Query is a database cloud solution that functions like a SQL server. However, as it runs on Google servers, it can reduce the downtime due to as serverless infrastructure. Additionally, since it uses automatic scaling and high-performance streaming ingestion to query data, it is able to significantly shorten the querying times.

- Chapter 4 comprises the steps of Data understanding, Data Preparation and Modelling of the applied methodology. In this chapter, all the procedures are carefully explained starting with the data collection and finishing with the optimisation of the hyperparameters to obtain the best possible modelling results;
- Chapter 5 discusses the results of the applied modelling techniques, focusing on explaining the obtained results and proposing initiatives to leverage the predictive capability of the model. Additionally, given the importance of product information, a strategy to improve this issue of the customer journey is presented, highlighting different approaches, their implications and impacts for Farfetch;
- Chapter 6 presents the main conclusions of the project, its limitations and suggests initiatives to be developed in the future.

2 Literature Review

In this chapter, the goal is to introduce the main concepts that provide a theoretical basis for the project. To do so, it was decided to focus on three main areas: luxury and its online market in section 2.1; customer experience and the broad impact of returns in section 2.2 and a methodological review where the different modelling techniques were explored in section 2.3.

2.1 Luxury Fashion Market

In the beginning, it was important to get to know the essence of luxury and its connection to the online world. This was done as a way of understanding the context of the luxury market, its customers and their high-end demands and, lastly, the benefits and opportunities for luxury fashion players to increase their digital presence.

2.1.1 Luxury Concept

In recent years, with the emergence of e-commerce platforms, luxury goods have become more available to a wide range of customers and, on top of that, the Internet and particularly social media are shifting the consumption model, influencing the way customers perceive luxury.

The term "luxury" is commonly seen as difficult to define. Depending on the area to which it applies and the people that define it, luxury can represent totally different concepts. Kapferer (1997) views luxury as "art applied to functional items" whereas McKinsey Corporation (1990) presents a more economical point of view, defining luxury brands as those that have the highest price/quality ratios in the market for very similar products.

Despite the definition, luxury is a subjective construct and satisfies multiple physical and psychological needs of consumers (Hennigs et al. 2012). Okonkwo (2009) completed this idea stating that it cannot be described as a product, it's part of an identity, a philosophy and a culture, most of the times expressing an indirect social stratification. Vigneron and Johnson (2004) proposed a framework to distinguish luxury from non-luxury brands that comprise five dimensions, based on non-personal perceptions of conspicuousness, uniqueness and quality and personal perceptions of the extended self, and hedonism. Beyond factors such as craftsmanship, sustainability and high-quality materials, the psychological cues are considered to be the most influential ones as wearing a luxury item is said to raise the esteem of its owner (Vigneron and Johnson 2004).

Hennigs et al. (2012) identified four main segments of luxury consumers. Firstly, consumers who strongly desire luxury goods due to their aim of feeling special, and unique. Then the ones that value the social and individual aspects of luxury, associating it with pleasure, and status. Additionally, customers that value the exclusivity higher and want to feel that their goods are only accessible to a few people. Finally, consumers that prefer the superior quality of a luxury product to the aspect of status, mainly enjoying luxury in private.

2.1.2 Online Luxury Market

Luxury Fashion companies traditionally had a very distinct vision of business when compared to the non-luxury ones. As most of the time these companies are ancient and familiar, unlike the majority of businesses, they were not interested in growing as much as possible (Okonkwo 2010). Their focus is on conceiving the best possible products and satisfy their high-end client needs with a focus on the experiences provided (Okonkwo 2010). Also, the luxury market customers tend to be particularly demanding and expect a personalized treatment as they are willing to pay a premium price to have access not only to the uniqueness of the product but also to an exclusive purchasing experience (Deloitte Company 2017).

Especially in the past, some brands felt that growing and particularly entering the e-commerce business would imply losing their exclusivity status (Guercini and Runfola 2015). Even though both areas are constantly pursuing avant-gardism and creativity, despite the proven success of e-commerce in the recent years, the relation between luxury items and e-commerce, at least theoretically, can be a little contradictory (Guercini and Runfola 2015; Okonkwo 2010). The fact that shopping experiences and feelings of exclusivity are not easy to recreate through a web-based platform caused many luxury brands to resist taking advantage of the e-commerce market (Okonkwo 2010).

When buying online, customers lose access to some aspects of the experience such as the feeling of the products, the personalized treatment and the social interactions. Moreover, it is also important for luxury companies to carefully consider other potential disadvantages of their online strategies. Dall'Olmo Riley and Lacroix (2003) pointed, for example, the danger of contributing to a proliferation of counterfeit items due to the fact that the collections are online and easily accessible to everyone, which can damage the brand image and affect the willingness of customers to buy their products.

However, as stated by Dauriz, Remy, and Sandri (2014), the access to a wider range of products, the convenience, the speed of the process, as well as the possibility of shopping almost anywhere and anytime, make online shopping a very powerful tool for the luxury players. Those can also benefit from the possibility of efficiently acquiring new clients and reinforce their positions in emerging markets such as China, India and Russia. (Guercini and Runfola 2015)

Combining both worlds requires companies to understand the fundamental role of customer centricity to create an interactive and engaging website experience, maintaining the personalized feel and emulating the prestigious atmosphere of a luxury store (Bjørn-Andersen and Hansen 2011). It is important that all aspects of a luxury brand website are created to make the customer feel as safe and confident as if buying "in person". This idea acquires additional importance as, for high-end products, the involvement of the consumers is much higher, causing them to search more extensively and, engage in comparisons, implying that companies should focus on providing the best and most accurate information about the products (Angela Hall, Neil Towers 2017).

In conclusion, despite all the potential drawbacks and the fear of losing their exclusivity status, considering the exponential growth that the online luxury market is expected to have (Achille, Remy, and Machessou 2018), the brands that will excel in the market will be those that embrace the Internet as a complement to their offline activities and as a way of efficiently reach international markets, instead of viewing it as a threat (Guercini and Runfola 2015).

2.2 Customer Experience

The section starts by explaining how the customer experience is formed, its main components and the importance for companies to adopt a customer-centric perspective. Afterwards, the customer experience is positioned along the customer journey, where its three main phases are identified: pre-purchase, purchase decision and post-purchase.

2.2.1 Customer Experience Formation

With the emergence of so many service-oriented businesses, the interest in the study of customer experience has increased both from the academic and business perspectives.

Customer experience has been defined in many ways. According to Gentile, Spiller, and Noci (2007), it is a group of interactions between a customer and a product or the company itself. Every customer has a unique experience since these interactions depend on both rational and emotional involvements (Gentile, Spiller, and Noci 2007).

On the other hand, Schwager & Meyer (2007) believe that customer experience includes direct contacts, like customer care throughout the purchase and use of the product or service, but also indirect contacts, in the form of reviews, advertising and recommendations. The customer’s response to these contacts is entirely personal and subjective (Schwager & Meyer 2007) and according to Berry, Wall, and Carbone (2006), sometimes, small details end up having a huge influence on the assessment of the overall perception of the experience.

Therefore, some of these influencing aspects can be controlled by the company, such as price and assortment, but others are out of their control, like expectations and the influence of others (Verhoef et al. 2009). Lemon and Verhoef (2016) reinforce this holistic view of the customer experience, stating that it begins with the search for a certain product or service and it includes the purchase, use and after-sale interactions. It is crucial to manage customer experience following a customer-centric approach (Lemon and Verhoef 2016).

Researchers, such as Zomerdijk and Voss (2010) notice that many service-oriented firms are focusing themselves on customer experience and placing it at the core of their value propositions. Schwager & Meyer (2007) advocate that the quality of the customer experiences should be the major concern of every service-based company. This customer-centric perspective, with the aim of creating distinctive and valuable experiences, can represent an increase in customer satisfaction and loyalty (Klaus and Maklan 2012) boosting firm's’ competitive advantage and revenue (McColl-Kennedy et al. 2015) especially in a moment were the service landscape is developing so fast (Ostrom et al. 2015).

However, from a practical point of view, there is still a lot to do regarding the concern for customer experience. In a recent study performed by Bain & Company’s with data about 362 companies, only 8% of their customers described their experience as “superior” (Schwager & Meyer 2007). Moreover, Angela Hall, Neil Towers (2017) reports that although many firms want a fully personalized shopping experience, very few actually have the necessary means, such as an individualized view of the customers and a complete data view of their target market.

When it comes to the formation of customer experience, Verhoef et al. (2009) have stated influential factors, such as the social environment, retail atmosphere, service interface, price, assortment, the company itself and the past experiences with that company. The authors also add that most of the times, experiences are moderated by contextual factors (e.g. economic climate) and individual factors (e.g. task orientation, customer attitude). Teixeira et al. (2012) considered other factors such as content, convenience (described as the availability and ease of access), speed and feeling of reward.

Those factors should be considered in a holistic view of the construction of these experiences. Lemon and Verhoef (2016) state that customer experience is created throughout the customer journey, defined as the series of contacts established with the firm during the purchase cycle, across multiple touch-points. Those are defined by Schwager & Meyer (2007) as instances of contact between the customer and the firm, or its products/services and Berry, Wall, and Carbone (2006) describe them as sub-experiences that are part of the overall customer experience.

This customer journey perspective requires the companies to clearly identify what are the multiple touch-points that the customers encounter during their journey, and focus on those that represent a higher value to them (Puccinelli et al. 2009). However, Maechler, Neher, and Park (2016) raise an important point, saying that the effort on optimizing each of these touch-points with misaligned initiatives, or poor management attention might cause the risk of losing the perspective of the entire journey, hence missing what the customers actually need and want (Berry, Wall, and Carbone 2006). Focusing on the overall experience provides a superior competitive advantage, enhancing customer satisfaction and increasing the probability of repeated purchases (retention) while reducing end-to-end service costs (Nisar and Prabhakar 2017).

Lipkin (2007) reinforces this idea saying that a holistic perspective does not require companies to monitor every possible factor that influences customer experience. Instead, they should focus on what is important for their customers, keeping in mind that what managers perceive as important might not be what customers value the most. (Andreassen et al. 2016)

2.2.2 Understanding Customer Experience through Customer Journey

Based on the previous section, it is clear that there is a consensus among the literature regarding the fact that customer experience goes beyond the exact moment when a customer encounters a specific product or service (Maechler, Neher, and Park 2016). Therefore, customer experience is the outcome of the entire customer journey.

Lemon and Verhoef (2016) conceptualize this customer journey, dividing it into three main phases.

The first one - pre-purchase - comprises needs' recognition, search and consideration of alternatives. Besides the real need for a specific product, during this stage, there are many factors that can trigger the need to shop. Angela Hall, Neil Towers (2017) state that customers regularly seeks novelties, knowledge and inspiration. Moreover, Sundström, Hjelm-Lidholm, and Radon (2019) add that shopping can be seen as a hunt for joy and satisfaction, or simply a way to counter "boredom".

Then comes purchase, the moment that comprises choosing, ordering and paying for the product or service. This topic will be extensively analysed in the next sections but, to understand what influences the purchase decision, Berry, Wall, and Carbone (2006) identified 3 types of clues, details that are part of the experience (functional, mechanic and humanic), that affect the purchase decision. In addition, besides the factors referred in the previous section Angela Hall, Neil Towers (2017) considers also the influence of the credibility, reliability and trustworthiness of the company as factors that can "make or break" the purchase.

And finally, post-purchase, that contains all the after-sale interactions with the firm or its product, such as the usage/consumption, engagement and support requests. This phase can include actions like returning the products, repurchasing, as well as recommendations or alternative forms of engagement with the company.

This last phase is acquiring increasing importance as some studies such as Court et al. (2009) start including the "loyalty loop" created after purchase as part of the overall customer decision journey. In addition, other researchers such as Verhoef et al. (2009) and Lemon and Verhoef (2016) also state the importance of the temporal factors and found out that previous experiences tend to influence expectations of users, and thus impact the way a customer will perceive the future ones.

2.2.3 The Purchase Decision in the Online Context

Traditionally, the purchase process has been seen by the literature as a funnel where an individual would start with a higher number of brands or products in mind and then, after consideration and marketing exposure, would select one of them and purchase it (Court et al. 2009). However, this concept is no longer able to capture the multiple touch points and the almost infinite product choice and channels available to a well informed and empowered customer that can move through a wide range of journeys (Court et al. 2009; Angela Hall, Neil Towers 2017).

According to Achille, Remy, and Marchessou (2018) the typical luxury customers follow a mixed online/offline journey with 70% of the sales being online influenced. On top of that, Google (2018) states that 85% of the online shoppers start their journey in one device and finish it in another one. These statistics provide a clear idea of how powerful the online presence is and how diversified the "paths-to-purchase" can be.

Dauriz, Remy, and Sandri (2014) reinforce this idea referring that customers are increasingly using product-oriented sources of information where it is possible to compare prices and obtain additional information about the product, which can constitute a decisive factor upon the purchase moment.

Hyeonsoo Kim, Yun Jung Choi (2007) state that, to increase online sales, companies should be able to recreate the in-store environment. To do so the authors identified two types of cues: high task-relevant ones, strongly related to the customers´shopping goals, such as images, product descriptions, price, return policies, etc.; and low task relevant, described as relatively inconsequent website design features, such as colours, icons, layout etc. The results of the study showed that low relevant tasks enable the access to high end ones, influencing the perception that the customer has about the website and its products. Additionally, it revealed that better and more accurate information about a product impacts the attitude towards the brand, the perceived risk, thus increasing the purchase intentions.

Nisar and Prabhakar's (2017) study found a strong correlation between the quality of online shopping experience and customer satisfaction towards the purchase experience. This quality concept was described by the authors as having an appropriate website design, good customer care (such as delivery and refund services), wide assortment and extend product information.

These factors acquire additional importance in a moment where consumers have nearly unlimited purchasing options and channels for product research. With all these possibilities, expectations are also becoming higher, customers expect products to be available everywhere and for its information to be accurate and consistent across all touchpoints (Beach 2018).

In conclusion, brands should be aware of the importance of providing differentiated aesthetic experiences, complemented with meaningful and valuable content, and being able to give customers the information and support they need to make the right decisions (Court et al. 2009).

2.2.4 After Purchase Interactions and the Impact of Returns

According to Court et al. (2009), the post-purchase part of the customer journey is the one responsible for shaping opinions, as the journey is said to be an ongoing cycle where previous experiences might have a great impact on future ones. The same author reinforces the increasing importance of customer-driven touch points such as reviews or word-of-mouth, saying that marketing should prioritize their efforts to meet this trend. However, one of the potential outcomes of a negative post-purchase experience is the necessity of returning the products.

In an online shopping context, during the after-purchase moment, the customer might feel what is described by Heuer, Brettel, and Kemper (2015) as a cognitive dissonance finding out that the item is not that important or that it turns out to have unexpected attributes. According to Heuer, Brettel, and Kemper (2015), gathering valuable product information with the aim of increasing the "tangibility" of the items will reduce customers feeling of distance to the product and potentially eliminate the aforesaid dissonance during the post-purchase evaluations, reducing the risk of returning the product. Puccinelli et al. (2009) add that this is increasingly important to highly involved customers who are more concerned about making the correct decision. Additionally, Nisar and Prabhakar (2017) found out that higher satisfied customers tend to show greater intentions to repeatedly purchase products from an online retailer.

Returns are an undeniable factor of any-commerce company and currently, there is a trend for more flexible return policies as companies believe that it will increase sales more than the number of returns, thus positively impacting profitability (Hjort and Lantz 2016). However, according to Saarijärvi, Sutinen, and Harris (2017) sometimes these lenient policies end up fuelling unnecessary ordering and increasing return rates, which in turn has major implications in managing the complex and costly issue of online returns. The same study also proved that the return policy of the company has a direct influence on how customers search, compare and purchase online.

The importance of returns is related to their ability to reduce the purchase uncertainty and the risk of the transactions, thus increasing the customer's confidence in the company through a series of transactions, which might include returns (Hjort and Lantz 2016; Saarijärvi, Sutinen, and Harris 2017).

The study of Hjort and Lantz (2016) proved that "returners" are the most profitable customers, as they proved to have confidence in the company and generated higher contributions per order. On the other hand, returns policies used for short-term gains, such as those generating more demand or attracting new customers, can have long term negative consequences for retailers. Additionally, the same study, states that high-end consumers, such as luxury customers, typically expect high-quality products when confronted with more lenient return policies.

Summing up, after understanding the specificities of the luxury market and its customers, it was important to have a holistic view of the customer experience that is especially relevant in the luxury market. After exploring the literature, the importance of a customer-centric perspective was notorious. Companies should develop the appropriate mechanisms to understand the main friction points in the customer journey and effectively act on them. In the online context, one of the identified friction points was returns. Its influence on customer experience makes it one of the vital concerns of any e-commerce company.

Furthermore, the analysis of after-purchase interactions revealed that the effects of returns and return policies happen in all phases of the purchase journey. Given their extensive impact, these findings reinforce the importance of providing every possible information to the customers (to ease the purchase decision) and consciously define return policies and procedures. Those that chose to ignore such factors will feel the consequences, not only in customer perceptions but also in terms of profitability.

2.3 Modelling Techniques

Any e-commerce company is able to collect huge volumes of data. To take advantage of this data, transforming it in knowledge that may support customer relationship management, constitutes a relevant issue. Data mining techniques, in particular, classification techniques, are becoming popular tools to assist customer-centric operations. In this context, this section introduces different classification techniques. Naïve Bayes, Decision trees (as those were the base predictors used in gradient boosting), Random Forest and Gradient Boosting algorithms are explained. The hyperparameter optimization procedure is also described in this section.

2.3.1 Classification Models

In recent years, data mining is gaining increased attention from academics and companies. It is explained by Kotu and Deshpande (2015) as the activity of finding useful patterns in the data, most of the times associated with knowledge discovery, machine learning, and predictive analytics.

Predictive analytics can be described as one of the multiple branches of data mining, where historical data is used to make predictions about future events. Within this vast area, Bishop (2007) subdivides it into three main problems: unsupervised learning, reinforcement learning and supervised learning.

The classification models fall into this last category, described by the same author as problems where data comprises examples of the input vectors along with their corresponding target vectors. The goal of those problems is to build a model that is capable to assign each input record to one (binary classification) or multiple discrete categories (multi-class classification).

2.3.2 Naïve Bayes

The Naïve Bayes algorithm is rooted in the principles of statistics and probabilistic theory (Bishop 2007). It is built under the Bayes' theorem, taking advantage of the probabilistic relationships between features and the class label. Apart from the theorem, its name comes from the naïve (ingenuous) but strong assumption of independence between the input features.

This assumption might not always hold true, but the limitation brought by this assumption is most of the times surpassed by the simplicity and robustness of the algorithm (Bishop 2007). To build the model, Naïve Bayes algorithms calculate the probability for each value of the output variable, for given values of the input features. After that, based on conditional probabilities (Bayes Theorem), for each unlabelled record, the algorithm computes the predictions.

2.3.3 Decision Trees

Decision trees are described by Kotu and Deshpande (2015) as one of the most used and intuitive approaches to classification problems. According to the author, their success is related to their simplified set up in analytical terms combined with an easy interpretation from a business perspective.

As the name indicates, this technique is based on a structure that takes the form of an inverted tree where at each node an attribute is tested, and a binary decision is made. Usually, the response variable has two classes (1 or 0) and the nodes split the data into purer subsets according to the label values of the input attributes. The final output is a tree structure that can be used for the prediction of new unlabelled data.

2.3.4 Random Forest

Random forest can be seen as an evolution of the decision trees algorithm (Kotu and Deshpande 2015). This method is composed of a group of decision trees that make their individual predictions. In the end, the most voted class ends up being the final prediction. The main difference is that, at the moment of splitting the data, on the nodes, the algorithm considers only a random subset of features included in the training set to fit the different decision trees. This is done with the goal of reducing the generalization error, increasing randomness and avoiding poor performances in unseen data.

2.3.5 Gradient Boosting

Gradient boosting is described by Prokhorenkova et al. (2017) as a powerful machine learning technique. The author explains it as the process of "building an ensemble predictor by performing gradient descent in a functional space". The theoretical background of the technique relies on the idea that strong predictors can be achieved by iteratively combining weaker models in a greedy way (Prokhorenkova et al. 2017).

Gradient boosting using decision trees is a technique based on progressively training complex models to minimize the error of predictions. It is described by Ershov (2018) as one of the most efficient ways to develop ensemble models, said to provide state-of-the-art results when applied to structured data. The algorithm is composed of two main steps. One, where gradients of the loss function that is being optimized are computed for each of the inputs. And a second one where, in a greedy manner at each iteration, the tree that improves the loss function the most is selected.

2.3.6 Hyperparameter Tuning

The process of tuning the hyperparameter happens has a way of finding the best "settings" to optimize the performance of the models. However, as the majority of the models act like black-box algorithms, to enhance their performance it is required to test different combinations of input parameters (Bergstra and Yoshua Bengio 2012).

Two of the most widely used methods to perform this tuning are the Grid search and Random search. The first one works by searching exhaustively through a specified subset of hyperparameters that can be seen as displayed in the grid. To obtain the best performance, the model will then be trained with every possible combination of the parameters and come up with the best one. In spite of allowing the model to obtain a good performance, the main drawback of this technique is that it turns out to be very time consuming and computationally expensive.

These constraints can be overcome by performing a random search. In this method, instead of searching the best parameters by looking at every possible combination, the model randomly chooses *n* predetermined combinations and trains the model with them, outputting the best one.

3 Business Understanding

Following the CRISP-DM methodology, the first step is to get an overview of the company and its operations. Hence, in section 3.1, before diving into the complexity of Farfetch, an overview of its competitors is presented to understand what the best practices in the market are, especially regarding return policies and product content.

Additionally, the production process that allows the items to be available on Farfetch.com is introduced in section 3.2. The ordering process is presented in section 3.3, both from the company (section 3.3.1) and the customer perspective (section 3.3.2), exposing the entire purchase journey.

After giving context, the customer perceptions were analysed in section 3.4. This analysis has been performed as a way of identifying the main pain points along with the customer experience and their implications, especially in terms of returns.

Finally, in section 3.5 the return process was analysed in-depth, focusing on its effects and costs for the company.

3.1 Understanding the Market

As a first step of the business understanding process, Farfetch's major competitors were analysed. This analysis was conducted with the aim of identifying the biggest players in the online fashion market are and how they work and position themselves, comparing to Farfetch.

In the online luxury market, there are several players, but Farfetch has been able to leverage its unique business model to obtain a privileged position. As a marketplace, Farfetch does not hold any stock or transportation system, since the partners are the ones responsible for them. This *modus operandi* allows the company to have the most valuable shelf of the market as well as the biggest number of different brands, positioning it as the number one player in the market. In Figure 3 a comparison between Farfetch and its competitor's assortment can be seen.

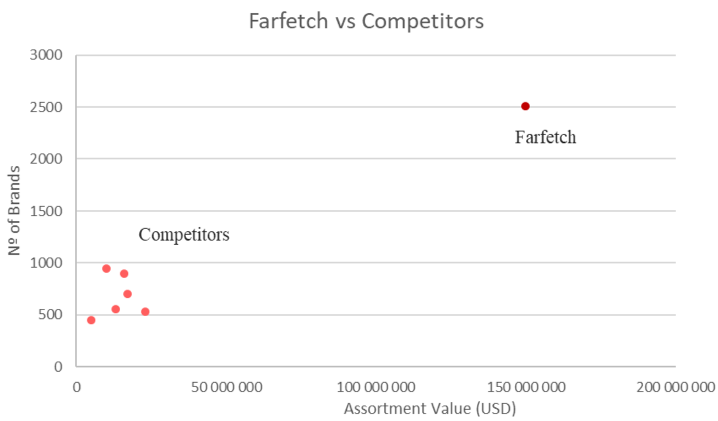


Figure 3 - Farfetch vs Competitors (Nº of brands and Assortment value)

Additionally, some companies were benchmarked. The chosen ones were those identified as competitors by online luxury customers, in a study developed by the company. The results in terms of awareness of the major luxury fashion players are summarized in Figure 4 and were the basis for this analysis.

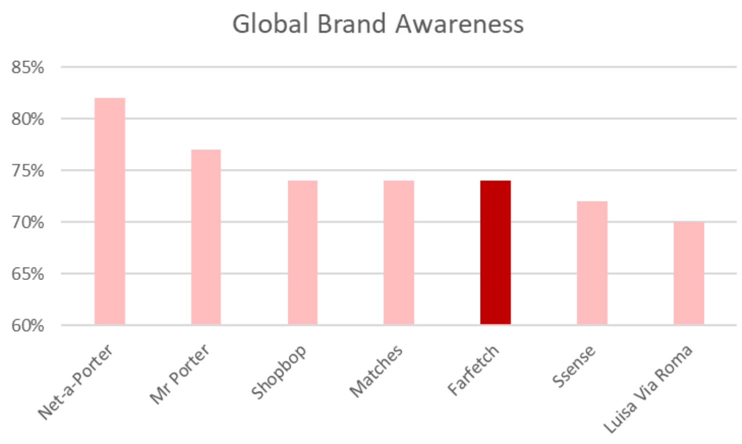


Figure 4 - Farfetch competitors based on brand awareness

Comparing the return policy among these 6 competitors, except for *Shopbob* (that requires the customer to pay for a shipping fee), all of them offer free returns to the customers. Additionally, it was found that *LuisaViaRoma*, similarly to what happens in Farfetch, offers only 14 days for the products to be returned after the delivery moment. The remaining companies allow their customer to return items until 28 days after delivery. This shows the importance that Farfetch competitors are giving to the return policy, as a way of improving customer experience.

Some companies go even further. Despite not being a luxury fashion company, Zappos - an online e-tailer - recently updated its policy to allow customers to return products within one year after the delivery moment.

However, more than providing a flexible return policy, companies should focus on mitigating returns, by having accurate content that can bring the customers' perception of a product closer to reality. Hence, all competitors were analysed regarding the product content that is displayed on their websites.

In this topic, *Net-a-Porter*, *Matches Fashion* and *Mr Porter* clearly stand out. In their websites, product descriptions include, apart from the functional characteristics of the product, inspirational content about the product, brand or the designer. Moreover, these companies also provide measurements for every size of the products, hints about how the product will fit the customer and, for some products, additional content such as videos or runway pictures. However, it should be said that companies benefit from their easy access to their inventory to obtain this information.

After the comparison, it is evident that, despite its dominant position, Farfetch still has much room for improvement. Given the importance that the aforesaid features have in the customer experiences and consequently the profitability of the company, there is the need to invest in better product content and to rethink the return policy as its impacts.

3.2 Farfetch Production Process

Acting as an intermediary between luxury customers and partners from the entire world, Farfetch is not responsible for the available inventory. However, since the company wants to keep a distinctive and coherent style across all its offering, the production process is one of the most complexes in the entire operation.

The production process is shown in Figure 5. It includes every step needed to make a product available online, and it can start more than six months before it is finally displayed.

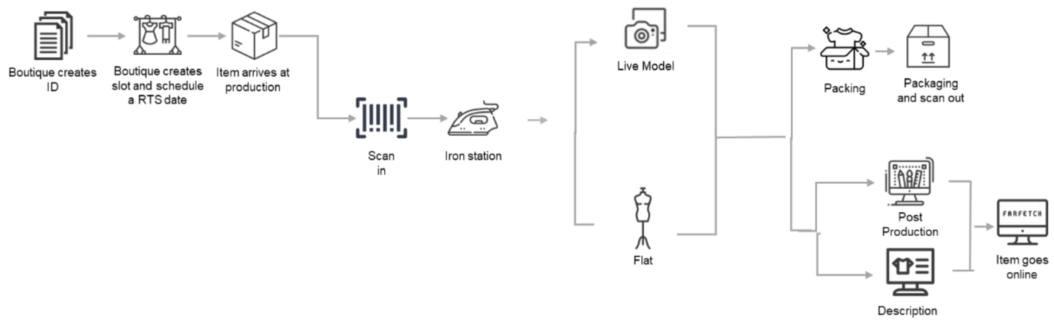


Figure 5 - Farfetch production process

Every time a partner, brand or boutique wants to sell a new item on Farfetch, the first step is the creation of the item on the internal communication platform. Since all fashion items have a unique identification number (ID), partners are encouraged to check if the product already exists in the platform to avoid duplicates. After that, one sample of all the items that are going to be online needs to be delivered to one of the company's production centres (Guimaraes, Sao Paulo, Hong Kong or Los Angeles).

Typically, the items are sent in "slots" that might contain up to 50 different products. Upon arriving at the facilities, the items are scanned in using a bar code. To avoid wasting resources, it is checked if the item was already produced and if there are any quality issues that prevent production. If not, the items are split into different categories according to their characteristics before moving to the next step.

After the adequate preparation, the items are photographed, in the still and live model, and the photographs are then sent to quality control and to post-production, where they are edited and polished to ensure the alinement with Farfetch standards. Once the photos are approved and there is no need for re-shoots, the items are packed and sent back to the partner.

Immediately after taking the photographs, a parallel process is triggered. The photos are sent to the members of the content team that, based on them and sometimes in the information provided by the partner, are responsible to insert all the product information, such as description, functional characteristics, the model measurements and other important details in the database. When both processes are completed, the picture and the information are combined, and the product is ready to go online.

3.3 Farfetch Ordering Process

Once the items are online, customers can interact with them and place orders. In the next sections, the ordering process is explained both from a company and customer perspective.

3.3.1 Internal Perspective

The ordering process starts when a customer places an order on Farfetch website. After choosing the desired product(s) the customer must also define the delivery option and the payment method. When a customer orders several products, although he/she might not notice, the items can come from different boutiques. Hence, every order, that contains all the purchased products, can be split into several boutique orders, according to the availability of stock.

When the orders are placed, they must go through the six different steps of the order cycle until the products are delivered to the customer. All these steps are represented in Figure 6.

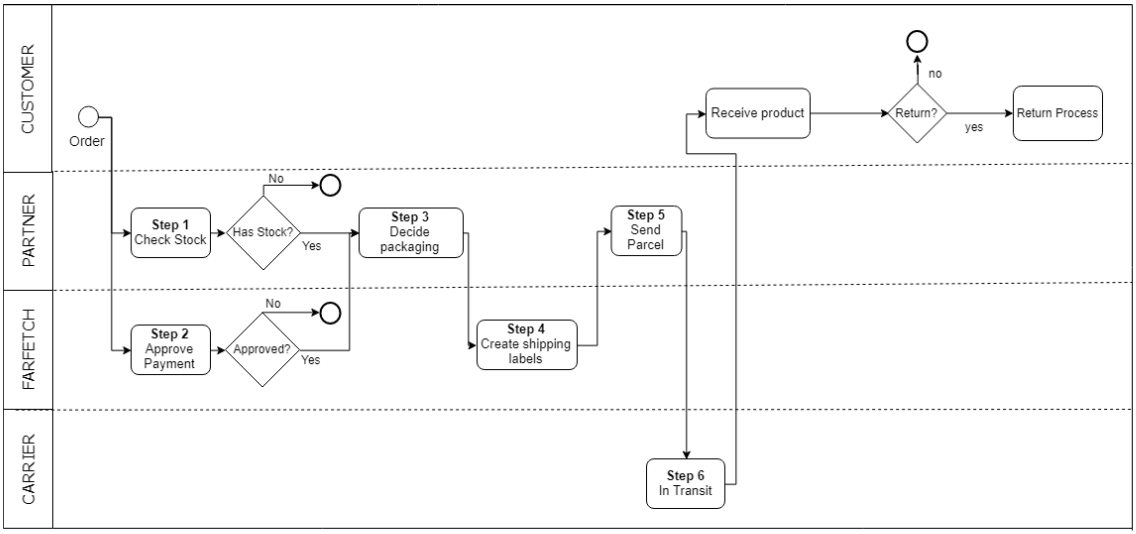


Figure 6 - Farfetch internal ordering process

Step 1 – Check Stock

After receiving an order, the partner must check if it has stock of that specific product. “No Stocks” can happen since offline sales can occur at the same time as the online ones and the system might not be perfectly synchronized. If that happens, Farfetch tries to find the same product in a different store and, if that is not possible, the customer is recommended a similar item or refunded, depending on his/her preferences.

Step 2 – Approve Payment

This step occurs in parallel with the first one. At this point, Farfetch wants to ensure that the placed order is not fraudulent. Most of the approvals are made automatically, but if there are any suspicious signs, the Fraud team is responsible to analyse the details of each case and cancel the order if it seems dubious.

If the order is approved, it proceeds to step 3.

Step 3 – Decide Packaging

In this step, the boutique is responsible for choosing the adequate packaging for the selected items. To do so, partners are encouraged to use a software called STORM where they have access to Farfetch recommendations regarding packaging. The parcel might also include additional personalized details, such as a sticker or handwritten letters, to enhance the customer experience.

Step 4 – Create Shipping Label

The fourth step is related to the creation of the shipping labels, that allow couriers to identify the delivery destination. Most of the times this process is automatically done but, in some cases, the Delivery team needs to intervene to obtain further details or correct misspells in the customer's address.

Step 5 – Send Parcel

Once all the required documents are prepared, the parcel is classified as "ready to send". After this, the couriers are informed to pick up the order and the customer receives an email informing that the items have been shipped.

Step 6 – In Transit

The last step starts at the moment when the courier picks up the parcel from the boutique and comprises all the necessary travelling time until the order is delivered to the customer.

Return Process

Upon receiving its order, if the customer decides that he/she no longer wants to keep the products, a return process is triggered. This process and all the associated costs for the company will be further analysed in the upcoming sections.

3.3.2 Customer Perspective

Once the item is online, it is available to be purchased by Farfetch visitors. However, depending on the motivation of the user, the path taken until successfully buying a product can be very diversified.

According to the company's internal data, it was found that only 25% of the users visit Farfetch with a clear intention of shopping. Other motivations can include inspiration, increasing fashion knowledge or even price comparisons. All these different reasons result in different journeys and interactions with the website.

Customers can access the website directly through Farfetch.com, opening a homepage where inspirational contents such as photos and articles are displayed. After that, if the user wants to refine the search, he/she might use the search bar or the available menus, looking for products of a specific category or brand. By doing that, the visitor will be taken into a product list page (PLP). This page includes a list of products and a menu where the user can filter the search even more or sort the items according to a given attribute (e.g. price).

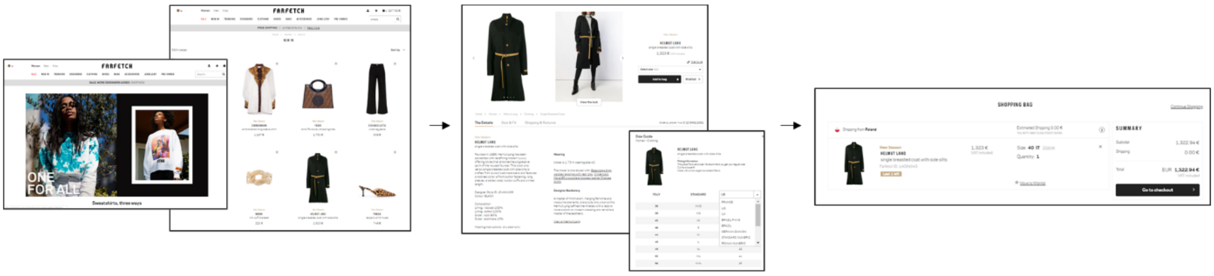
Finally, if the user thinks that a product is worth additional consideration, he/she can click on its picture and will be redirected to the product display page (PDP) where a single product and all its information are displayed.

Additionally, it is also important to highlight that users might not go through this entire journey. Other possibilities for them to encounter Farfetch products include the use of search engines, such as *Google* or *Bing*, online advertisements or email marketing that, depending on the interaction, will take the user to one of the mentioned pages.

In the PDP, the user can find all the information regarding the product, and also shipping and return conditions. Along with the pictures, in the *Details* tab, the description includes features such as price, material composition, washing instructions, information about the designer and about the model that was shot wearing the item. Moreover, on the *Size & Fit* tab, for most of the products, the user can only find information about the measurements of the model and the size of the item he/she was wearing. For *One sizes* (products with single size) and for *Vintage* items, Farfetch also provides specific measurements of the products (e.g. Bust, Sleeves, Shoulders) as those are measured in the production centres.

Before checkout, the customer must choose the size of the product. As different countries have different size scales, there is a *Size Guide* that allows users to compare the designer sizes according to the country's specific scale. Additionally, to help the user with this task, the company is working on a feature called *Fit Predictor*. As the name indicates, it aims to predict the best size for the user and does that based on three questions in which the user must indicate a brand, a category and the size that he/she normally wears with that brand, allowing the algorithm to compute the adequate size of the item that is being considered.

All the navigation process until the purchase moment is summarized in Figure 7.



Inspirational Content & PLP

PDP & Size Selection

Add to Bag & Checkout

Figure 7 - Website navigation and purchase journey

3.4 Understanding the Customer

After getting to know in detail how the company operation works, it was decided to investigate what the perceptions of Farfetch customers are, regarding their experiences with the company. To do so, different sources of data were crossed to get an overview of the main pain points in the customer journey. Additionally, the idea was to understand potential areas where actions might be taken to improve the customer experience while boosting the profitability of the company.

As the focus of the project was to understand returns and what the main factors behind it were, it was required to comprehend how the identified pain points, were impacting customer experience and, consequently, contributing to higher returns.

3.4.1 Customer Surveys

Frequently, the User Research team at Farfetch is focused on developing initiatives to understand company perceptions and get hints of what the feelings of the customers are when moving along the purchase journey with Farfetch. To do so, surveys and interviews are conducted, and the feedback is then gathered and communicated to highlight potential initiatives that can benefit customers.

In a survey conducted in 2017, the company focused on identifying the most important features when shopping luxury fashion items online. The results are summarized in Figure 8.

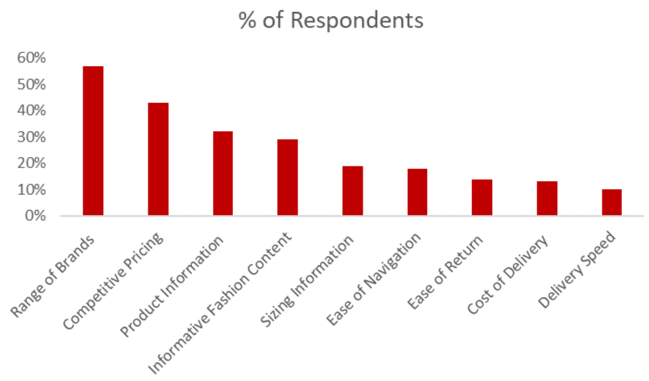


Figure 8 - Most important features of online luxury e-commerce websites

Looking at it, it is clear that luxury fashion customers have high expectations towards the websites where they shop. Furthermore, it is also worth mentioning the presence of 3 features related to product content in the top priorities of customers (Product information, Informative content and Sizing information). After the assortment and price, the information about the available items is crucial in terms of customer experience. This insight was confirmed in a more recent study where the complaints about product information, specifically, sizing information, product description and imagery, were identified as one of the major areas causing friction in the purchase journey.

3.4.2 User Test Sessions

Given the wide range of channels and options for the customers to interact with the Farfetch, the focus on providing a customer-centric experience is seen as a major pillar to support competitiveness.

To increase the knowledge of issues in the customer journey, in August of 2018, the company started to organise “User testing sessions”. Monthly, 18 research participants from the US or UK are asked to visit Farfetch offices and conduct a purchase journey across the multiple platforms (e.g. desktop, mobile web and the iOS app). Each of these sessions is recorded and the comments, issues or insights provided by the customers are written and stored in a library that comprises all the feedback of these sessions.

The so-called “pain points”, problems or expectations stated by users are identified and categorized according to the area where it affects the customer experience and ranked according to the impact on it. Since the start of these experiences, 492 insights were collected and are going to be analysed in this section.

In a superficial analysis of this data, taking advantage of the categorization provided, it is possible to understand that the *Product Content* categories such as *sizing* and *product information* account for 35% of all comments. This first insight confirms the analysis of the previous section, indicating that these topics are of great importance for Farfetch users and thus, represent areas where improvements are relevant and necessary.

To get a deeper understanding of the user feedback, a word cloud was built using the raw data from the collected comments. Word clouds are powerful visualization tools for huge datasets, considered good ways of identifying customer pain points and access feedback about companies or their services (Heimerl et al. 2014). Their way of working is simple: the more frequent a word is on the dataset, the bigger/bolder it will appear on the cloud.

According to the methodology applied by Miguéis and Nóvoa (2017), and using *Rstudio,* after gathering the required data, the first step is to tokenize the text, in other words, every comment is broken into pieces called *tokens*. Additionally, it was necessary to convert all the words to lowercase and after that, remove all the stop-words, classified as the most common words in the English language, such as "the", "is", "at", "which", "on", among others. Moreover, an important step was taken before the construction of the word cloud, the elimination of the word "User", that was directly linked to the way these comments were inserted in the platform and could possibly bias the analysis.

Finally, all the generated terms were put into a *Term document matrix* that characterizes the frequency of the terms occurring in a certain collection of documents. In this type of matrix, rows represent a document, in this case, user comments and columns correspond to terms, every word of each comment. Afterwards, the words were sorted and the word cloud was built (using *word cloud* package) and is represented in Figure 9.

Figure 9 - User test sessions word cloud

Looking at the word cloud, the emphasis of the word "Size", reinforces the idea that product and, particularly, sizing information is crucial when users are navigating across Farfetch website. The explanations for it can be diversified but it should be considered that currently the company still has a defective way of displaying this type of information, which turns out to be impacting customer experience and potentially contributing to the increase of returns.

3.4.3 Customer Contacts

In addition to the surveys and the user test sessions, with the aim of monitoring customer opinions and needs, Farfetch is recently using a platform called *Re:infer* that gathers information about e-mails, call transcriptions and comments made on other forms of interaction such as the Net Promoter Score (NPS) survey that is sent to the customer after every purchase. The sets of text collected are called *verbatims*.

The goal of the platform is to make data about user interactions understandable and actionable. Using machine learning technology, it is possible to categorize each *verbatim* with variable degrees of confidence, also making sentiment analysis, providing a percentage from 0 to 100% in terms of positivity.

The complete *verbatim* data includes 377 945 samples of text, of which only 1% has been reviewed by humans. The rest of the labels have been automatically applied as a result of the platform's algorithm. Along with the text, all the entries have additional attributes such as the country of the user, the store name or the device where the interaction occurred.

When analysing the reasons why customers where contacting the company, it was found that the most representative ones were delivery, packaging and items' quality. However, as those categories had a high positive number of *verbatims* (e.g. 75% of the delivery queries were classified as having a positive sentiment) it was decided not to explore them.

Inspired by the findings of the previous sections, it was decided to take a closer look at the product content *verbatims.* Despite not being one of the main reasons why customers contact the company, as it represents only 2% of all *verbatims*, when looking at the sentiment analysis, 87% of the comments were classified with a negative label, being one of the worst performing categories on this matter.

The data confirms the idea that the available information on the website is, in fact, one of the main pain points of customers, generating a significant amount of negative feedback. Additionally, taking a closer look at what the customers' most common issues are (Figure 10), it was found that around 48% of the comments happened due to Sizing Problems of which only 6,3% were classified as positive.

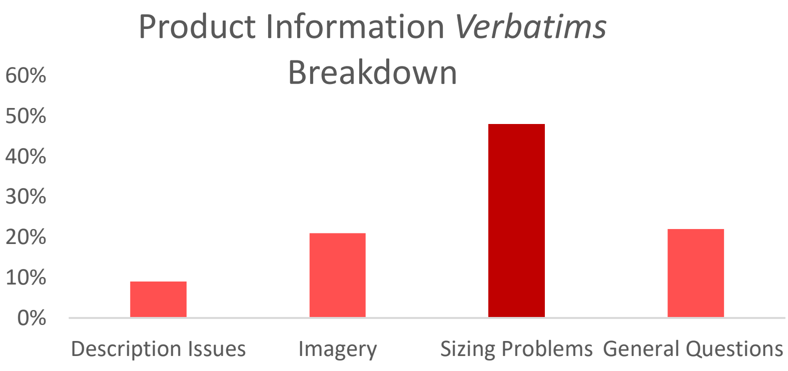


Figure 10 - Product information verbatims breakdown

The observed negativity around product content reinforces its importance. As seen in the explored sources of customer feedback, the information that customers are receiving about Farfetch products is not being sufficient. The high-end needs of luxury customers require the company to seriously tackle this issue, given its impact on customer experience. Moreover, as it will be deeply explored in the next section, this lack of adequate information is also influencing returns and, consequently, raising its associated costs.

3.4.4 Return Reasons Analysis

After analysing customer feedback and understanding the major pain points in their purchase journey with the company, it was decided to analyse what factors were being responsible for the majority of returns. To do that, the return reasons, indicated by the customers, upon a return, were explored.

When customers want to return a product, they must choose within a list of reasons why they want to make that return. After some effort, translating the return reasons stated by customers from several languages, it was possible to group them into 7 main categories that are displayed in Figure 11. From the figure, it is clear that "Size and Fit" issues were the most selected return reason, representing more than 65% of all returns in *Farfetch.com* since 2015.

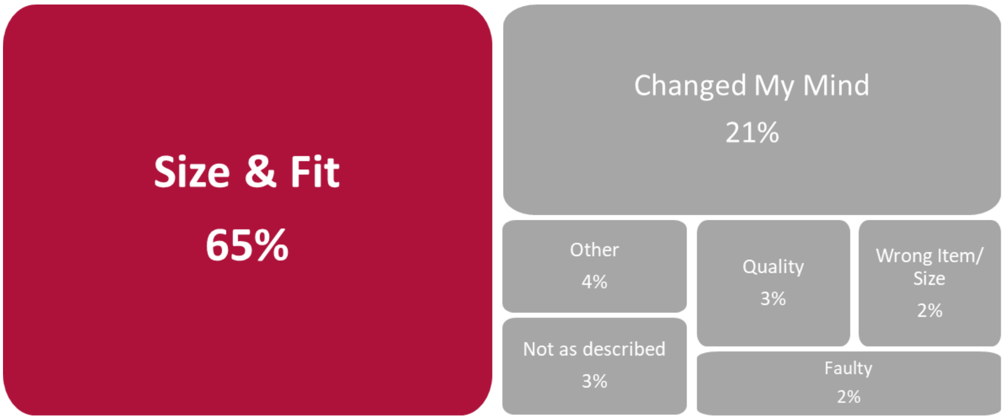


Figure 11 - Main return reasons selected by Farfetch customers

In this prevailing category are included reasons such as "The item is too big" or "Fitting Problems". As the vast majority of the products sold are wearables, the fact that customers do not have access to high-end product content, such as the article measurements or information about the product fit, contributes to this discrepancy.

This creates a massive opportunity, not only for Farfetch but also to other e-commerce companies. To mitigate these returns, it is important to develop initiatives that allow companies to provide better and more accurate information about the products. Those would, not only impact the number of returns but could also enhance customer experience, increasing the probability of loyalty and re-purchase behaviours.

3.5 Returns for Farfetch

As said before, returns are one of the fundamental components of any retail and e-commerce operation. For Farfetch, this is not different as the company must face the fact that more than 20% of the items bought end up being returned.

Once the perceptions of the customers and their impact in terms of returns was understood, the goal of the next sections was to have a clear idea of how the return process of the company work, mainly focusing on its return policy and on its financial implications.

3.5.1 Return Process

Returns represent a high cost for the company but, as an inevitable issue of e-commerce, Farfetch and other e-commerce players started to focus on how to improve their return policies, as a way of boosting their service level and thus gaining the loyalty and trust of the customers, hoping for repeated purchases.

According to Saarijärvi, Sutinen, and Harris (2017), the return policy can have a direct influence on how customers search, compare and purchase online.

As Farfetch views itself as a customer-centric company, it developed a return policy that allows customers to return the items completely free of charge. A pick-up can be scheduled, the products are re-shipped to Farfetch's partners and, after ensuring the quality of the item, the customer is refunded with the entire value of its order.

Every package that is delivered to the customer includes a return Air Waybill (AWB) together with the billing documents. In the past, to return the products, customers had to print this document, but it was changed to provide a seamless experience to those that want to return an item. To initiate a return, the customer must go through the following steps:

1. Access his/her account on Farfetch website an click the "My Orders & Returns" menu;
2. Select the option "By Courier" and chose the items that are going to be returned;
3. Indicate a return reason;
4. Schedule a pick-up time slot;
5. Prepare the parcel for the return, attaching the AWB to the exterior side of the box;

Additionally, Farfetch customers have the possibility of returning an item in-store, to a partner boutique near their address, or to make a return by contacting the customer service team that will help them during all stages of the process.

On the website, customers can find the company's return policy. It implies that returns should be requested no later than 14 days after the delivery date. After being shipped, a returned item needs to be analysed by the partner. If for any reason, the product is not compliant, partners can reject the return, triggering an internal process to handle the situation, especially in terms of cost split.

To avoid these situations the company has developed strict guidelines to ensure the optimal state of the products such as "Items must be returned unworn, undamaged and unused, with all tags attached and the original packaging included".

3.5.2 Cost of Returns

Farfetch's simplified return policy for the customer, together with the fast growth of sales, has made returns one of the major expenses of the company.

Return costs can essentially be divided into four main categories:

- Shipping costs, related to the movement of the products, that are influenced by factors such as the third-party carrier, the provenience and destination and the volume of the item;
- Duties, taxes applied on imports by customs authorities, influenced by the country's regulation, the value and the type of items transported;
- Taxes, additional fees that are charged by some specific countries, also known as value-added taxes;

- Other charges, that might happen as a result of an exception where, for instance, the carrier might ask for additional compensation for deliveries in remote locations.

Given the “full-money-back” return policy of the company, all the aforesaid costs are split in different manners between the partners and Farfetch, depending on the established contract. Additionally, when looking at the impact of returns, besides the monetary issue, the effect that these negative experiences might have in terms of trust and loyalty should be evaluated, as these can foster the risk of losing the customer.

4 Return Prediction Model

As mentioned in the literature review, post-purchase interactions are one of the fundamental components of a customer journey. Besides being a very cost-intensive operation, returns impact, not only influence the perceptions of customer experiences but can also affect the expectations and attitude in future purchases.

Because of that, one of the goals of this dissertation was to build a predictive model that, based on historical data, for each new transition would provide a score regarding the probability of that order being or not returned. This machine learning approach gives the company a tool that can provide insights about the most influential factors concerning returns and creates the opportunity for several initiatives based on the capability of being able to predict the return behaviour of the orders.

The first step was the collection of the data to perform the modelling techniques, that is explained in section 4.1. After that, in section 4.2, the process of feature engineering is explained as the created variables were not available in the data sources. In section 4.3 there was the need to make an initial exploratory analysis of some variables, aiming to have an idea of their impact and potential contribution to higher return rates.

Afterwards, all the required data preparation procedures are explained in section 4.4 and, finally, in section 4.5 the modelling procedures are explained, as well as the cross-validation technique and performance metrics used to compare the models. In this section, there is also the optimization of the hyperparameters of the models. The description of the results and their comparison is conducted in chapter 5.

4.1 Data Collection

Building the dataset required for data mining problems is one of the most important activities in this sort of projects as errors in this stage, might heavily impact all the analysis.

To obtain the final dataset all data had to be queried from the data structure of the company using *Google Big Query,* a web-based SQL management server. Based on business expertise, the main sources of data were identified and merged to obtain the initial dataset.

Additionally, it is important to state that the collected records are at the "product level". As explained in the above sections, a portal order can contain several boutique orders, that might contain several products (Figure 12).

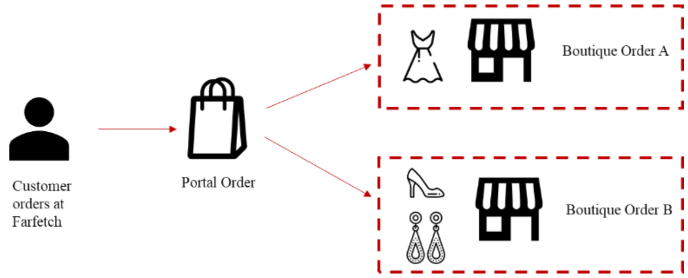


Figure 12 - Portal and Boutique orders

All the analysis was made for orders of specific products as this would allow the company to make the insights of the model much more actionable. Hence, the goal of this phase was to build a dataset where each row would represent a purchased product and each column would contain the attributes of that order and product.

This step was the most time-consuming activity since the database structure was initially hard to tackle and there was a large volume of data to be collected (5 data frames with more than 7 million records). The collected data was then merged and restructured in the computing environment *Jupyter Notebook* using the programming language *Python* and mainly using two libraries: *Pandas* and *Scikit-learn*.

Moreover, an additional step conducted in this phase was the elimination of some features, that had no influence in the return rate. Again, those features were identified based on business understanding and removed to avoid unnecessary noise in the dataset.

4.2 Feature Generation

After collecting all the required data, a very important step, prior to building the model was the identification of potential engineered features that might have some influence in the variable under investigation. As those features couldn't be obtained directly from the database, there was a need to engineer them using SQL queries on the available data.

4.2.1 Dependent Variable

The first and most important variable to be created was the label, the variable that identifies if a product was returned or not. As in this case, the project deals with a binary classification problem, each record in the dataset should be assigned to one of two classes: returned or non-returned. This classification was made by analysing the possible states of an order – *orderslinestatusglobal* - (In Progress, Sent, Cancelled, Returned …). Hence, a binary variable (*returned*) was created assuming the value "1" when the state was "Return" and "0" in any other situation.

4.2.2 Additional Features

As said before, besides the attributes that were available on the dataset, it was thought that some additional information regarding the behaviour of the user and the products purchased could influence the probability of return. To cover this hypothesis, some features were created:

- *User_purchase_count* - The number of products purchased by a specific user before that order;
- *User_return_rate* - The ratio between the number of returns and the number of orders for a specific user*;*

- *Product_purchase_count* - The number of times a product was purchased before that order;
- *Product_return_rate* - The ratio between the number of returns and the number of orders for a specific product*;*
- *Itemcounts* - Number of items contained in a portal order;
- *Same_item_counts* - Number of purchases of the same article in a portal order.

Additionally, to get an understanding of the seasonality effects, there was a need to create variables such as *orderdate_year, orderdate_month, orderdate_yearweek, orderdate_weekday*, to explain the year, month, week and weekday where the order was placed.

4.3 Data Understanding

After obtaining the complete dataset, based on the CRISP-DM framework, it was explored the influence of some potential factors on the return rate. The first set of factors that were analysed are related to the customer. Here, it is important to highlight the distinction between Existing and New customers, where the existing ones are those that already have bought in *Farfetch.com*. Moreover, there is a distinction between New and Old Returners, the last ones being those that have at least returned one product. Besides user information, there is also an exploration of features regarding geography, price and product characteristics

4.3.1 Existing vs New Customers

To understand the influence of the customer purchase behaviour in the return rate, the variable *User_purchase_count* was studied. In this case, if the value of the variable was higher than zero, the customer was classified as an Existing Customer.

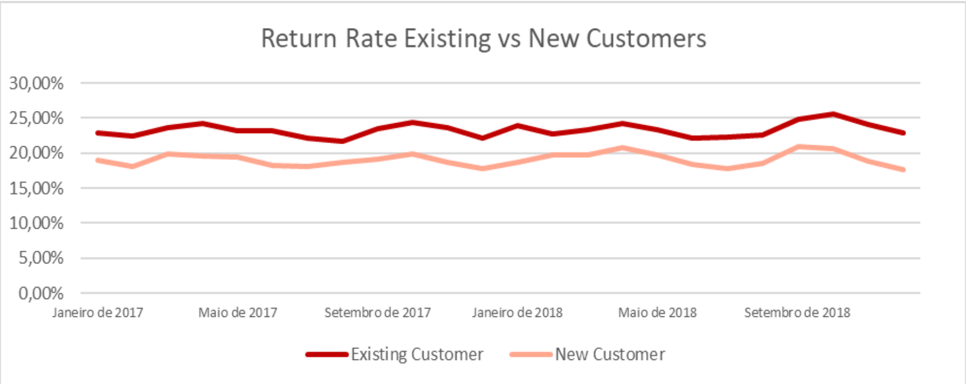


Figure 13 - Existing vs New Customers return rate comparisons

Existing Customers are the ones that generate the highest revenue for Farfetch. However, at the same time, from Figure 13 it is clear that those are also the ones with a higher return rate (23,3%, compared to 19,1% of the New Customers). According to the literature review, customers that had past purchase experiences with companies have an increased trust, being more prompt to return their products in case anything goes wrong with their purchase.

4.3.2 Old Returner vs New Returner

A similar procedure was followed to analyse the impact of previous returns in the probability of a user returning an order. Customers that had a *User_ReturnRate* higher than 0 were classified as Old Returners.

In this case, despite the difference in the number of orders not being very significant, the disparity in terms of return rate is clear when comparing New with Old returners. (Figure 14) From this, it is possible to infer that people that already returned products, value the experience and end up doing it more often than people that have never returned before.

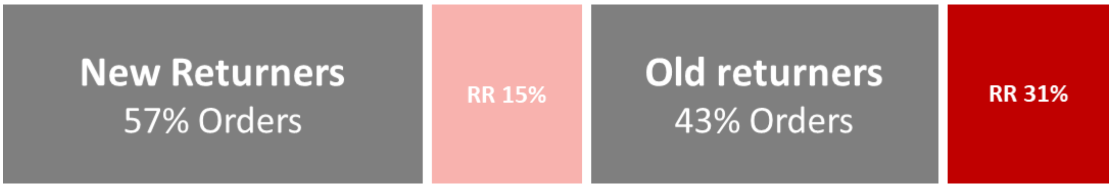


Figure 14 - New vs Old returners comparisons

4.3.3 VIP vs Non-VIP Customers

One of the features included in the dataset is a binary variable called *isvip* that has the value 1 for customers that already have spent more than 10 000 US dollars with Farfetch in the previous year. Those are treated as special customers having access to exclusive campaigns and specialized offerings. It was found that these customers account for around 2% of all Farfetch customers but end up generating almost 23% of the revenues of the company.

Associated with this higher spending, it is also a higher tendency to return the products as it is possible to see in Figure 15, VIP Customers have a return rate that is 1,7 percentual points higher than the regular ones.

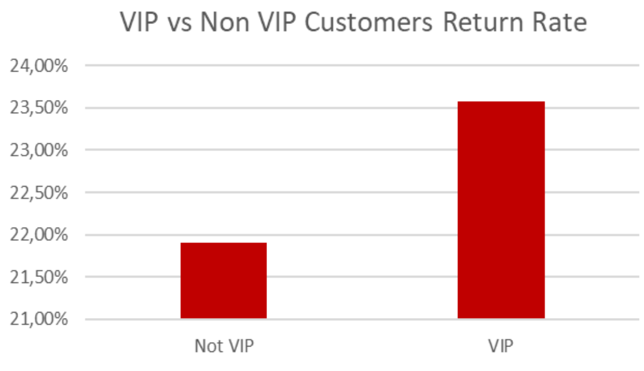


Figure 15 - VIP vs Non-VIP return rate comparison

4.3.4 Region

In order to understand the geographical dimension of returns, the return rate was studied for orders shipped to different regions, considering the variable *Customer_RegionFF*.

The return rate was plotted on a map of the world, as seen in Figure 16. The darker red colours represent regions where the return rate is higher as opposed to the ones with darker grey colours that represent regions with lower return rates.

Analysing the figure three main regions should be highlighted, the USA, UK and particularly the DACH (a region that comprises Germany, Austria, and Switzerland). Based on the empirical experience of teams responsible to handle this matter, this discrepancy is mainly related to cultural reasons. People of the DACH region are more used to return online products, resulting in the highest return rate for the company (31 %).

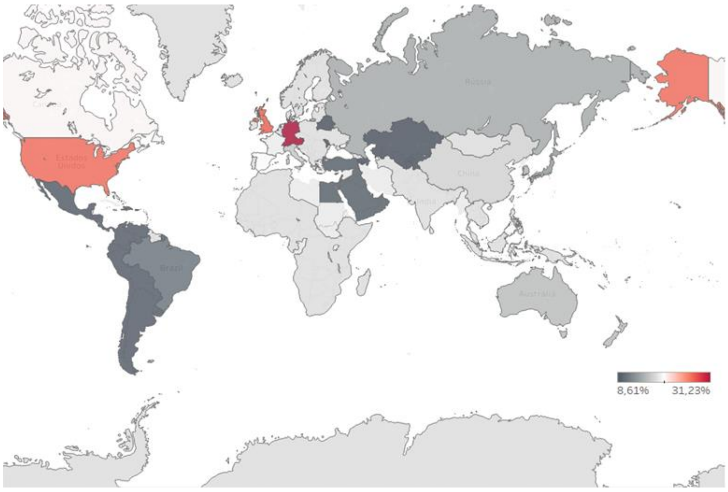


Figure 16 - Region return rate comparisons

4.3.5 Average Order Value (AOV)

Another interesting variable to be explored is the price paid by the customers, to do that it was decided to study the Average order value (AOV). The AOV is a common metric in the business and it is related to the total value that is paid by the user when placing an order. This metric is calculated after summing the *valuewithoutship_user_usd* and the *valueship_user_usd*. Additionally, for plotting purposes, 4 buckets of AOV values were created. Based on the analysis of Figure 17 it is possible to infer that higher value orders are associated with higher return rates. This data seems to indicate that customers might have more conscious evaluations of the products when their price is higher.

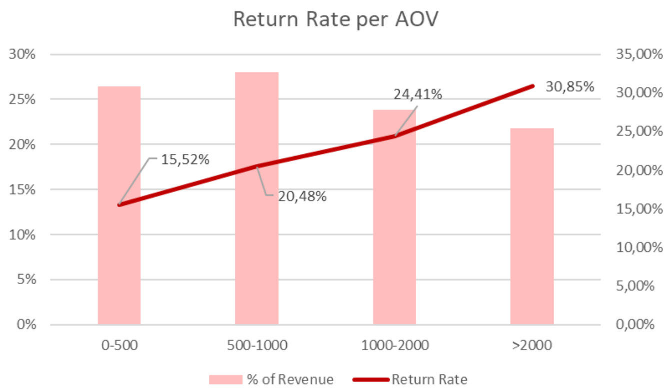


Figure 17 - AOV return rate comparison

4.3.6 Product Gender

To analyse how product attributes might impact returns, one of the characterizes taken into consideration was the product gender. All products have a unique gender that is used to identify the menu where they will appear on the website. If the gender is "unisex" it means that the item can be available on both menus.

As seen in Figure 18, "Women" products are the ones with the highest return rate (25%). Besides being the most purchased products, this might also be related to the type of customer as Women are said to have a higher involvement and be more demanding when it comes to fashion item purchases.

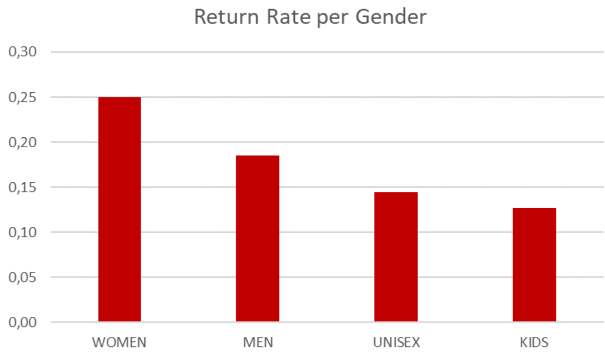


Figure 18 - Product gender return rate comparisons

4.3.7 Product Category

Finally, it was analysed the influence of the feature *category_1stlevel*. From Figure 19 it is clear that clothing (26%) and shoes (22%) are two of the categories where returns are more significant. This acquires additional importance as both categories account for more than 85 % of the company's sales. Hence, being able to predict and act on those will certainly generate a huge impact on Farfetch.

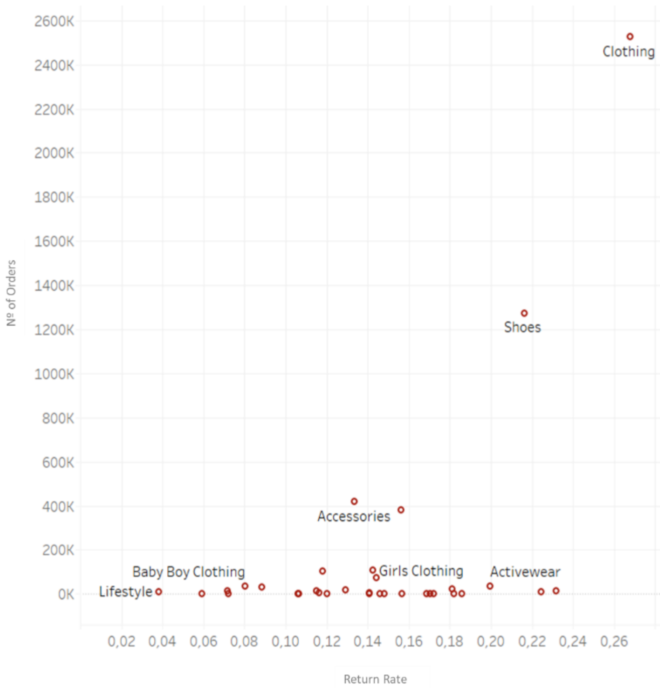


Figure 19 - Product 1st category return rate comparisons

4.4 Data Preparation

Based on the preliminary analysis made in the previous section, the relevance and behaviour of some variables with respect to return patterns could already be analysed.

In the next sections, all the required preparation procedures to apply the modelling techniques will be described.

One of the first processes at this stage was to remove from the dataset features that could not be implemented in practice. A critical example of it is the *Lead Time*, the elapsed time between the moment a customer places an order and the moment when the item is delivered. It is known in the business that this variable deeply impacts customer experience, thus potentially influencing the probability of a return. However, as the goal was to predict the return behaviour at the moment an order is placed, the *Lead Time* had to be excluded from the analysis.

After generating the required features and eliminating potentially noisy features, the final list of variables in the dataset can be seen in Table 1.

Table 1 - List of variables used in the model

Variable	Type	Description
returned	Categorical	Class label (0 or 1)
brand	Categorical	Brand of the product
category_1stlevel	Categorical	Category of the product (e.g. Clothing, Shoes)
category_2ndlevel	Categorical	Sub-category of the products (e.g. Dress, Skirt)
countryoforigin	Categorical	Country where the product was produced
customer_tier	Categorical	Tier of the customer (T1, T2, T3, T4)
designer_code	Categorical	Unique code that identifies specific groups of products
isVIP	Categorical	User VIP status (0 or 1)
itemcounts	Numerical	Number of items included in a portal order
orderdate_month	Categorical	Month of the order
orderdate_weekday	Categorical	Weekday of the order
orderdate_weekyear	Categorical	Week of the year of the order
orderdate_year	Categorical	Year of the order
Product_Purchase_count	Numerical	Number of purchases of a specific product
Product_ReturnRate	Numerical	Return Rate of a specific product
product_gender	Categorical	Gender of the product (e.g. Women, Man)
promocode	Categorical	Promotional discount used
RegionCustomerFF	Categorical	Region to where the order was shipped
sameitemcounts	Numerical	Nº of times the same product was purchased in a portal order
shipper	Categorical	Courier responsible to deliver the order
shiptypeid	Categorical	Delivery Method (e.g. Standard, Express)
sizescalename	Categorical	Name of the scale chosen by the user (e.g. Partner A US, Partner B IT)
sk_size	Categorical	Size of the product
StoreRegion	Categorical	Region of the Store
user_fraud_status	Categorical	Customer Status (e.g. Good Customer, Problematic Customer)
User_Purchase_Count	Numerical	Number of purchases of a specific user
User_ReturnRate	Numerical	Return Rate of a specific user
valueship_user_usd	Numerical	Shipping fee paid by the customer
valuewithoutship_user_usd	Numerical	Price of the product paid by the customer

Moreover, to use machine learning algorithms, it is necessary to perform some data preparation procedures such as removing the missing values and numerical encoding the categorical features. Hence, to ensure that the models were trained and tested using the exact same datasets, the preparation processes were executed beforehand, and the output dataset used for all the modelling procedures.

4.4.1 Handling Missing Values

One of the technical requirements of the modelling procedures, especially when using *Python* is that the models cannot be built using data that contains missing values. The features that had incomplete data can be seen in Table 2.

Table 2 - Missing Values

Feature	Missing Values	% of Missing Values
promocode	4984544	70,0%
customer_tier	806925	11,3%
brand	172913	2,4%
countryoforigin	172311	2,4%
sizescalename	172248	2,4%
designer	172228	2,4%
category_1stlevel	172221	2,4%
category_2ndlevel	172221	2,4%
RegionCustomerFF	122229	1,7%
isVIP	122151	1,7%
User_Purchase_Count	33	0,0%
User_ReturnRate	33	0,0%

Looking at the table, the variable *promocode* immediately stands out. That large number of missing values means that only 30 % of the Farfetch orders in the dataset were made using a promotional code for a discount. To handle the missing data in this feature, an additional category was created (None), that assumed the user had not use any promotional code.

Another special case was the variable *customer_tier*. The customer tier ranges from one to four and is related to the previous spending of the customer with the Farfetch. To handle that, an additional category "T5" was created, to cover situations where customers had no tier associated.

Finally, for the other categories, as the percentage of the missing values was not very significant, those records were replaced with "NULL", that is interpreted by *Python* as an empty attribute, that is disregarded in the modelling process.

4.4.2 String Standardization

The second step required during data preparation was the standardization of the categorical values. This process was done, aiming to homogenize the categorical features that had string values such as *brand, RegionCustomerFF* among others. By executing this, it was possible to eliminate inconsistencies and reduce the number of categories.

The process consisted of lowering the case of every string, remove special characters and replace spaces by the *underscore* symbol.

4.4.3 Categorical Features Encoding

The final step of data preparation was to encode the categorical variables. This was done using the *Target Encoder* library. There are several encoding techniques but, the way this one works is called Target or Mean Encoding. In a nutshell, the technique uses the values of the label class to generate the new encoded features. The categorical features are then replaced by the probability of the target label or by its average (if the label variable is numerical) for each category of the features.

This procedure was chosen due to its simplicity and powerful way of improving the classification models. However, as some features are being defined based on the target variable, there must be additional consciousness to avoid data leakage and overfitting.

The result was a numerical dataset ready to be fed into all the modelling techniques.

4.5 Modelling

Concluded all the preparation process, the final dataset included data of all orders of FF since 2018, containing around seven million records. As expected, this volume of data and the computationally demanding algorithms that were to be applied would result in a prohibitive running time of several days. To overcome this constraint, a temporal data sample containing only data for one month (approx. 460 thousand records) was taken. The resulted dataset was then fed into all the models. The adopted procedure has some limitations that are going to be discussed in the final section of this document.

4.5.1 Dataset Balancing

Another step needed prior to applying the modelling techniques was the balancing of the dataset. This operation was required because the number of returned and non-returned observations was significantly different. Of all the dataset records, only 19% corresponded to returned orders.

The strategy to perform the balancing was *undersampling*. In a nutshell, when applying this technique, all the records of the less represented class (in this case ‘returned’) are considered and after that, an equal size random sample of the other class is taken to complete the dataset as seen in Figure 20.

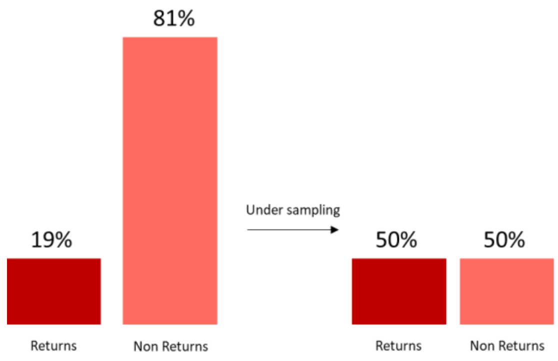


Figure 20 - Dataset balancing operation

In this case, the result of the balancing process was a dataset containing 161 424 records with an equal proportion of returned and non-returned orders.

4.5.2 Performance Measures

Following the CRISP-DM methodology, after modelling, there must be an evaluation step. Hence, before diving into the modelling procedures, it was decided to define and analyse the main performance metrics that would be used to compare the different modelling techniques.

The result of the chosen classification models is a prediction (0 or 1) of the target variable, in this case, *returned*. Given that, for each record of the test set, in addition to the real value of the label, the algorithm will make a prediction and compute a predicted label.

The scope of the project was to predict returns. Because of that, records, where the target variable had the value of 1, were considered positives. In opposition, if the value was 0 it was considered a negative.

Whenever observations are correctly labelled by the model as a return (the positive class), these predictions are called *true positives* meanwhile, observations that are properly classified as non-returns are called *true negatives*. However, there are two more potential outcomes of classification that imply an error. The first one is called in the literature as *type I error* and happens when the model predicts a positive class but in fact, the label of the observation is negative, this being called a *false positive* prediction. The second one is the *type II error* and, being the opposite of the first one, happens when the negative prediction (0) and the positive label don't match. These predictions are called *false negatives*.

In the context of the project, a *type I error* would correspond to an order that is predicted as a return but ends up being kept by the client. On the other hand, a *type II error* would occur if the model predicted a certain order as a non-return when in fact it happens to be returned by the customer. Considering the business perspective, the second one is the one with the most impact for the company and thus will be the main focus of the project. One of the most common ways of visualizing this "metrics" is a confusion matrix (Table 3).

Table 3 - Confusion Matrix

		Real Class	
		0	1
Predicted	0	True Negative	False Negative
	1	False Positive	True Positive

From the above matrix, it is common to calculate several performance metrics that are the basis for most of the model comparisons in binary classification problems. Examples of those are:

$$Accuracy = \frac{True\ Positives\ +\ True\ Negatives}{True\ Positives\ +\ True\ Negatives\ +\ False\ Positives\ +\ False\ Negatives}$$

$$Precision = \frac{True\ Positives}{True\ Positives\ +\ True\ Negatives}$$

$$Recall\ = \frac{True\ Positives}{True\ Positives\ +\ False\ Negatives}$$

$$F1Score\ = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Once again, considering the context of the problem it was decided that, despite all the aforesaid metrics should be taken into consideration, the *recall* should be the main focus when comparing the different models and evaluating their performance. This metric allows the company to understand, of all the returned products, how many did the model predict correctly, having a clear notion of its capabilities and potential gains when applied in the real environment of Farfetch.

ROC curve

Another tool that is commonly used in classification problems is the Receiver operating characteristics (ROC) curve, described by Fawcett (2006) as a simple and powerful technique for visualization, and selection of classifiers based on their performance.

It is essentially a two-dimensional graph in which the true positive rate is plotted on the y-axis and the false positive rate is plotted on the x-axis, representing the trade-offs between benefits (*True Positives*) and costs (*False Positives*) (Fawcett 2006) (Figure 21). To understand it better, two points in the chart are worth mentioning. The first one is the (0,1) point that represents the perfect classification. In opposition, there is the (1,0) point that corresponds to the worst possible classifier. Hence, the closer the classifiers are to the top left corner, the better their performance.

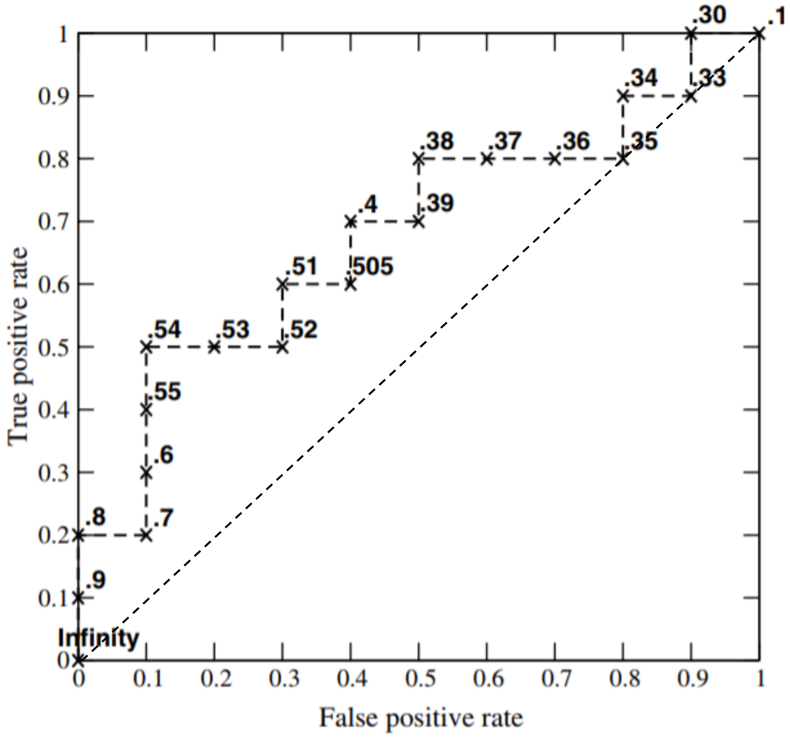


Figure 21 - ROC curve Source : (Fawcett 2006)

The diagonal line in the chart corresponds to the random performance. Fawcett (2006) says that if a classifier was to randomly guess the value of the target variable, it will probably get half of the positives and half of the negatives correct. Any classifier that is below the diagonal line is worse than random predictions and models should be able to exploit information in the data to move for the upper triangular region.

Probabilistic-based algorithms predict numeric values that represent the degree to which an instance is a member of one of the classes. Using a threshold, these predictions can be turned into discrete classifiers. Each of the defined thresholds produces a different point in ROC space. To build the ROC curve, this threshold is varied between -∞ and +∞ and the curve is plotted in the ROC space.

Nevertheless, to compare the performance of different models, there is a need for a single estimator that can measure performance. This is done by calculating the area under the curve (AUC). In simple terms, the AUC is described by (Fawcett 2006) as the probability that the model will rank a random positive instance higher than a random negative one. In other words, the higher the AUC, the better the model can predict positives as positives and negatives as negatives.

4.5.3 K-Fold cross-validation

When applying a machine learning technique, the way data is split between training and test subsets can have a major influence in the modelling results. Although that is much more significant for smaller datasets, this was taken into consideration.

One of the most common approaches is to use cross-validation. When applying this validation technique, instead of using the holdout method, where the dataset is split between a train and a test-set, the performance of the model is evaluated by repeating the training-testing procedure several times. During the project, this technique was applied for both iterations of the modelling phase.

K-fold cross-validation is one of the most common processes of cross-validation, as it is capable to mitigate the high computational effort required by other techniques. In this technique, the training set is divided in k-folds, with each of the folds being left out one time for the training-validation sequence.

In the first iteration of the project, this technique was applied but, in this case, the entire data set was. Hence, at each step, k-1 folds were used to train the model with the default parameters and the left-out fold used for testing purposes. In the end, the performance metrics correspond to the average of the results obtained at each of the k test folds, as represented in Figure 22.

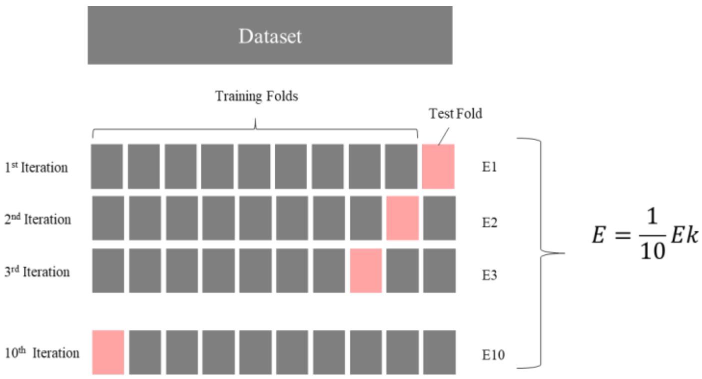


Figure 22 - 10-fold cross-validation, 1st iteration

In the second iteration, the idea was to use cross-validation on the train set, to obtain the best parameters that were later applied to the test set, as it will be further discussed in the next section.

The cross-validation procedure was applied to all the models (Naïve Bayes, Random Forest, and Gradient Boosting) with the k = 10. This choice was made as to when there is not any indication for choosing the number of k and the dataset is bigger enough, ten is a common practice in this sort of problems.

4.5.4 First Modelling Iteration

During the first iteration of the project, three different classification models were applied. As mentioned, to ensure the comparison of the results, the same dataset, containing only numerical data, was fed into all three algorithms. The chosen models were Naïve Bayes, given its simplicity and ability to provide surprisingly good results, Random Forest, an algorithm that is described as easy to use and capable of good results even with the default parameters and Gradient Boosting, an ensemble method that is said to have a state-of-the-art performance. Additionally, the choice of this last algorithm was supported by the fact that it is currently being used in other areas of the company.

All the mentioned models were trained using a 10-fold validation method, following the rationale explained in the previous section.

The average results can be found in Table 4 and on the ROC curves displayed in Figure 23.

Table 4 – 1st Modelling iteration results

Algorithm		10-Fold Cross-Validation				
Name	python library	Accuracy	Precision	Recall	F1	AUC
Naive Bayes	sklearn.naive_bayes	0,6281	0,6680	0,6270	0,6469	0,740
Random Forest	sklearn.ensemble.RandomForestClassifier	0,8139	0,8142	0,8139	0,8141	0,901
Gradient Boosting	catboost.CatBoostClassifier	0,8178	0,8182	0,8178	0,8180	0,908

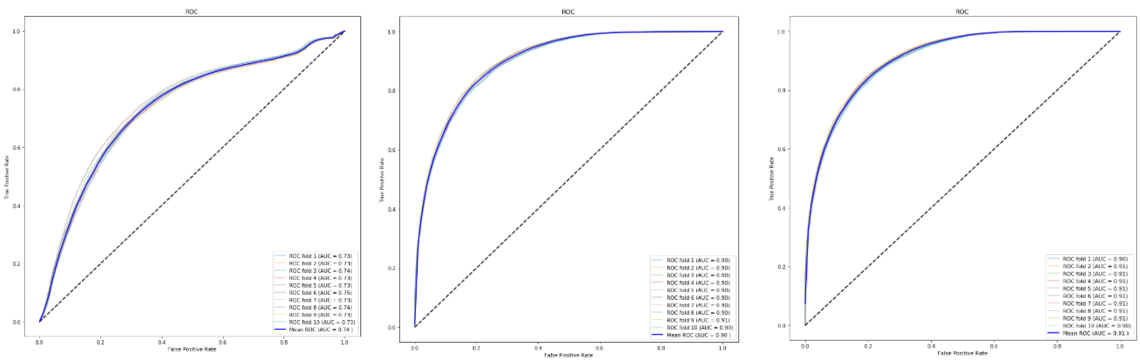


Figure 23 - Naive Bayes, Random Forest, and Gradient Boosting ROC curves

From the table, it is clear that the Naïve Bayes model is the one showing a significantly worse performance. In opposition, Random Forest and Gradient Boosting had similar performances, both showing very interesting results with a slight advantage for the Gradient Boosting in all performance metrics.

Hence, the next step was the optimization of the models. Following the explanation of Probst, Wright, and Boulesteix (2019) the optimization of the hyperparameters of the models, tends to improve their performance. Hence, aiming to obtain the best possible results, it was decided to search for the best hyperparameters of Random Forest and Gradient Boosting, excluding Naïve Bayes from further analysis. This optimization was done by tuning the models in a random fashion way, explained in the next section.

4.5.5 Second Modelling Iteration

During the second iteration, the focus was to optimize the hyperparameters of the model, looking for the best version of the models. As explained before, the main metric chosen to compare the algorithms, was the *recall*. Hence, this is the metric that the models were forced to optimize when looking for the best parameters.

As said in the *k-fold cross validation* section, in this stage, a ten-fold cross-validation method was also followed. The first step was to split the dataset into train and test subsets. This was done using a hold-out method where 25 % of the records were separated and kept aside during the training-validation phase. The rest 75 % of the data, the training subset, was used to find the best hyperparameters and train the model before the test phase (Figure 24)

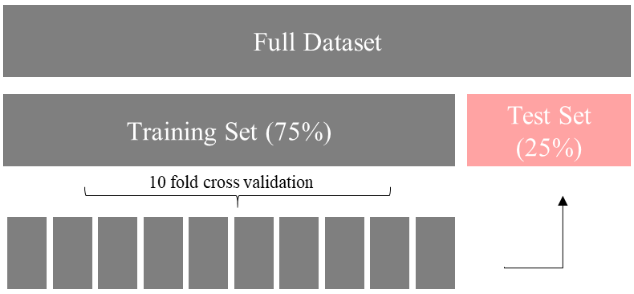


Figure 24 - 10 cross-fold validation, 2nd iteration

Using cross-validation, 10 folds of data where created and, during each iteration, nine folds were used to train the model with certain parameters and the left out one, used for validation purposes. For both algorithms, this procedure was repeated 30 times to obtain the best parameters that optimized the recall.

As mentioned, this optimization was done using a random search method. This means that a range of the parameters was provided and then, the algorithm chose random combinations of parameters to train the model. Another possibility to execute that was to use a grid search technique. In this case, the algorithm would consider and train the model using every possible combination of the hyperparameters. Despite potentially proving better results, this technique would drastically increase the computational effort and the time required to obtain the results. Because of that, it was decided not to use it during the project. The chosen hyperparameters to be optimized are displayed in Table 5.

Table 5 - Hyperparameters optimized

Model	Hyperparameters	
Random Forest	*max_depth:*	Maximum depth of each tree in the forest
	max_features:	Maximum number of features to be considered
	n_estimators:	Number of trees in the forest
	criterion:	Gini index or Entropy
	boostrap:	True or False
Gradient Boosting	*learning_rate:*	Controls the amount of change to be made in each step
	depth:	Depth of the base trees
	bagging_temperature:	Intensity of the Bayesian bootstrap
	l2_leaf_reg:	L2 Regularization coefficient (controls overfitting)

After obtaining the best set of hyperparameters, the model using it as input was trained in the training subset and later used to predict the labels of the test set. The results for Random Forest and Gradient Boosting can be found in Table 6.

Table 6 - 2nd modelling iteration results

Algorithm		10-Fold Cross-Validation					
Name	python library	Hyperparameters	Accuracy	Precision	Recall	F1	AUC
Random Forest	sklearn.ensemble.RandomForestClassifier	max_depth: 12, n_estimators: 408, max_features: 9, criterion: gini_index, bootstrap: True	0,8150	0,8153	0,8150	0,8152	0,907
Gradient Boosting	catboost.CatBoostClassifier	bagging_temperature: 0.5972, depth: 7.0, 12_leaf_reg: 29.8013 , learning_rate: 0.1935	0,8171	0,8174	**0,8191**	0,8182	0,900

From the table is possible to conclude that, the models do not show a significant difference neither from the ones using the default parameters nor from each other. Despite that, when analysing the main metric chosen for comparisons, Gradient Boosting is the one that had the best performance. Additionally, this algorithm makes it possible to check the relative importance of each feature for its predictions. These results are discussed in the next section.

5 Results Discussion and Recommendations

The idea behind this machine learning approach was to build a tool that would allow Farfetch to predict beforehand if a specific order would be returned. On top of that, when analysing the results of the models, the knowledge about which factors influence returns the most can be crucial. Creating awareness in the company for the reasons behind its return rate will certainly trigger a discussion on the obtained results and the development of several initiatives.

On the other hand, having in mind a customer-centric perspective, another pain point in the customer journey – insufficient product information – was tackled during this phase. The importance of providing valuable content was understood during the feedback analysis, supported by the literature review and reinforced by the absence of such features during the development model.

Focusing on the goal "Amaze Customers", the company will be able to provide better customer experiences, by improving its product information and mitigating returns. At the same time, it is increasing its profitability (by reducing the cost of returns) and promoting re-purchase behaviour.

In this chapter, the results and limitations of the model are discussed in section 5.1, as well as the opportunities created by its development. Then, in section 5.2, the product information improvement initiatives are presented, while discussing the required actions and implications for the company.

5.1 Model Results Analysis

After obtaining the results of every model, it was decided to explore the results of the one that showed the best performance, Gradient Boosting. Hence, the focus of this section is to discuss and come up with an explanation for some of the most relevant features in the model. The relative importance of each feature for its predictions can be seen in Figure 25.

Surprisingly, *designer_code* was the most valued featured by the model. This code is a unique code given by the brand to identify specific products. As it is created by the brand, the designer code is common to all the fashion players that have those products. Because it captures both colour and sizing information (e.g. the same dress in a different colour and a different size would have the same *designer_code*) it brings additional information that the model would not be able to capture in features such as *productid.* To understand the impact that such variable was having in the model performance, a variant of the models was tested, excluding this feature from the analysis. However, the best scores were always obtained when the models were trained using the *designer_code* feature.

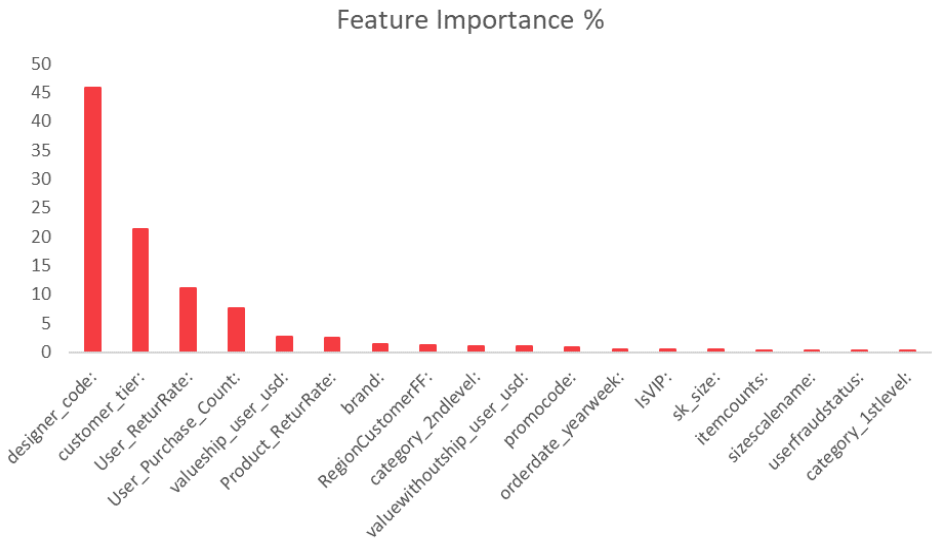


Figure 25 - Gradient Boosting features relative importance

The tier of the customer also proved to be a very influent attribute towards predictions. In Farfetch there are 4 tiers of customers defined according to the value that the users spent in the previous year. The presence of this feature confirms the idea that customers that spent more money with the company are more prompt to return their purchases. This can happen because, in one way, they are already experienced in the process, and in another way, these customers tend to have higher involvement with the products, thus being more demanding and consequently, returning more.

Additionally, the features *User_ReturnRate* and *User_Purchase_Count* also showed a relevant influence when predicting a return. Through these insights, it is possible to infer that orders that come from users that already bought in Farfetch and have returned several products, tend to have a higher probability of being returned. Raise awareness to this matter can be of major importance for the company, giving it the possibility of flagging certain customers and develop actions to improve their experience and mitigate unnecessary returns.

The *valueship_user_usd* also exhibited high importance. Essentially, this feature captures the shipping costs, the fee that is paid by the customer to cover the delivery of the order. Depending on the expected lead time, on the partner or the route, it can have different values. Upon a return, the customer does not have to pay any of this fee, as Farfetch, sometimes in collaboration with the partner, pays for these expenses.

Features directly related to the product such as *Product_ReturnRate*, *Brand*, and *category_2ndlevel* showed a similar impact on the predictions. That can be seen as the influence of the specific product characteristics in the return rate. As an example, it was found that a very specific "hot-selling" pair of trainers was having a higher than normal number of returns. To analyse this situation, Farfetch had to speak to the brand and eventually understood that those shoes had been designed for an oversized fit. However, as the customer was not being informed of this, it was causing many unnecessary returns. Being able to capture and take actions on such events can have a great impact, not only by mitigating returns but also by providing a better customer experience.

After the exploratory analysis, given the verified differences in terms of return rate across the different regions, it was expected that this feature (*RegionCustomerFF*) had a higher relevance when predicting a return. Despite not being one of the most significant ones it is still relevant to know where the orders are going to be shipped, as it can provide clues about how probable it is for the item to be returned.

It is also worth mentioning the variables *valuewithoutship_user_usd* (the price of the product) and the *promocode*, a variable that captures the use or not of a specific discount code. In the first one, based on the sense that higher priced products are returned more, the business expectations were that this feature could be one of the most important. However, the model proved that the influence of price was lower when compared to other variables. Furthermore, the case of *promocode* will be further discussed as a limitation of the project. As the promotional activity of the company is variable in time, the temporal sample might have prevented the model to capture the accurate influence of this variable.

5.1.1 Limitations of the Model

As in any other project, the developed modelling approach has a number of limitations that are important to be discussed. Those are not only relevant to understand its implications to the project but also to use them as clues and guidelines for what should be explored in future projects.

One of the main limitations of the model was the fact that all the modelling procedures considered only a sample of data of 1 month. This was done due to time constraints and as a way of reducing the computational effort. Using this sampled dataset, the model was not able to capture the seasonality effect. During a year, there are times, for example, close to New Year's Eve, where the number of impulsive purchases can be greater, or for example during Christmas, where many undesired gifts can be bought. Both situations result in a higher number of returns that the followed modelling procedure was not able to capture. The same happens with promotions and marketing campaigns. As these actions are not constant during the whole year, a temporal sample of data might be biased towards the presence or absence of such campaigns.
Another limitation has to do with the involved features. Despite the extensive search for variables that could contribute to returns, it is possible that some influencing features have been left out from the analysis. In future iterations, this search for additional features should continue and their inclusion in the model validated.

Moreover, it should be stated that at this point, the company has no visibility regarding the products that have superior content such as fitting advice or measurements per size. In the future, the monitoring of such cues should be guaranteed, and those features included in the model given their identified impact on customer experience.

Finally, technical limitations. In a deeper analysis, other modelling techniques should be explored as well as different validation methods. Additionally, the choice of the parameter optimization method was done based on a trade-off between performance and time. Nevertheless, in future works, other methods and parameters should be considered.

5.1.2 Future Actions for the Company

After discussing the results and limitations of the model, it was important to understand its practical applicability for Farfetch. This phase was done by addressing several internal stakeholders and identifying areas of opportunity that could leverage the prediction capacity of the model. Hence, the critical initiatives that have been identified were:

1. <u>Define localized return policies.</u> Nowadays, Farfetch's return policy is the same in any case. The company can use the insights of the model to understand if some adjustments, for instance for specific regions or product categories would be beneficial;

2. Understand the feasibility of selling a product. In some situations, selling specific items can end up reducing the profitability of the company. One of the influencing factors is related to the number of returns. Hence, having the possibility to predict the return rates for a given item will enable Farfetch to understand those situations and develop tactics to minimize their impact;
3. Develop a more accurate financial forecasting. As said in section 3, return costs represent one of the biggest financial expenses for the company. Hence, being able to estimate returns associated costs and embrace the insights of the model in the financial forecasting will certainly improve the accuracy of the predictions. This topic acquires additional importance as Farfetch has recently become a public company.
4. Create the possibility of warning customers. Based on predefined thresholds, whenever the model identifies a significant probability for the product to be returned, Farfetch could develop mechanisms to alert the customer. This could require the company to suggest alternatives or perhaps offer compensations to prevent the customer from placing specific orders. Nevertheless, these actions could imply a trade-off between profitability and the experience provided to the customer.
5. Increase awareness for fraudulent customers behaviours. Given Farfetch's lenient return policy, some customers may incur in abusive behaviour. To control it, the company developed a flagging system that identifies problematic customers. Currently, this process is manually done by the teams. Using the prediction capability of the model will enable the company to effectively identify those situations and trigger the appropriate actions.
6. Provide insights as a service. Most of Farfetch partners, besides selling their products in the platform, also possess their own e-commerce websites. Hence, given the potential benefits of the model, the company could explore the possibility of providing its capabilities and insights as a service "offered" to brands and boutiques.
7. Work as an input to the order allocation algorithm. As explained in section 1, the same item can be available in different Farfetch partners. Whenever there is a need to define which partner is going to fulfil an order, based on several inputs, such as price and lead time, the order allocation algorithm is responsible to compute the most appropriate partner. Hence, if, based on the predictive model results, it is possible to identify a significant probability of return, that should be taken into consideration. In those cases, as different boutiques have different returning costs, between two similar partners, the algorithm should attribute the order to the partner with the lowest return's cost.

Additionally, it is important to state that the model may suit the immediate needs of the company as currently there is no other way of predicting returns but, as the industry and the market change so fast, there will certainly be the need to adapt it, discover new features and more areas of application.

5.2 Product Information Improvement

After the machine learning approach, some interesting conclusions could be drawn. From the analysis, it can be seen that, despite the high influence of customer behaviour, the characteristics of the products play a vital role when determining if a product is going to be returned or not. Farfetch is working on a number of initiatives to provide a premium product content that meets its high-end customer demands. However, as those initiatives are only in their early stages, data about them cannot yet be accessed, which prevented the inclusion of those features in the model.

Based on that, after crossing this information with the insights on the customer side and analysing what the biggest players in the online fashion market are doing, Farfetch decided to launch four initiatives with a clear intention of improving the customer experience and mitigating returns. Those will allow the company to provide superior value to the customer, as it eases product consideration, potentially increasing the conversion rates.

Moreover, as explained in section 2, the impact of better product content will also have consequences in the post-purchase phase. The necessity of a return is a friction moment in the purchase journey. Hence, guaranteeing that customers can access more and better information about products is expected to reduce the probability of a customer returning a product.

The ongoing initiatives are:

- Enhanced Product Descriptions: improving the description of the products by increasing the details and inspirational features aiming for increased purchase intentions;
- Fitting Advice: hiring and training fitting specialists that would be responsible to describe how clothes are expected to fit the customers and recommend size adjustments (ex: This piece fits small. For a looser fit order one size up);
- Size Scales Accuracy: providing accurate "brand specific" size scales for every possible item;
- Measurements per size: developing a strategy to obtain measurements per every size of the products available in Farfetch.

Considering the limited time of the project, an "impact-difficulty of implementation" matrix (Figure 26) was built and, based on this framework, together with business priorities, it was decided to support the roll-out of the Measurements per Size initiative as it was the one with the biggest impact, considering the time and effort to be implemented.

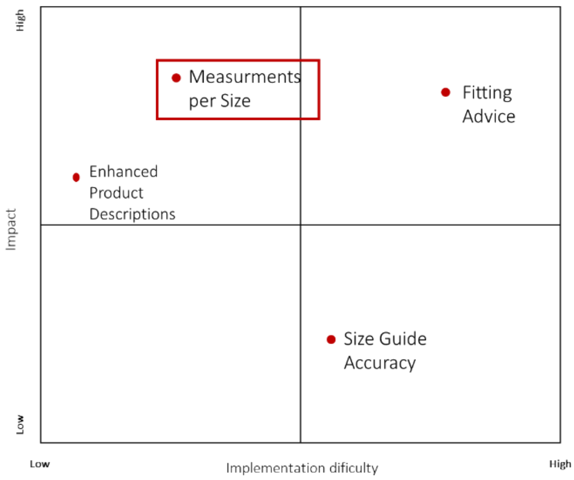


Figure 26 - Impact vs Implementation difficulty matrix

The scope of the problem and the methodology that was followed will be explained in the subsequent sections.

5.2.1 Measurements per Size

As said before, providing better content to online luxury customers can be beneficial in many ways. Apart from providing a better purchase experience, the possibility of mitigating returns will be able to, not only eliminate an unnecessary pain point in the customer journey, but also reduce the costs associated with this process, boosting the profitability of the company.

Hence, the idea behind this section is to explain the motivation for the “Measurements per Size” initiative and the developed plan to obtain this information.

5.2.2 AB Test

To investigate the value of providing measurements for all sizes of the items, before launching this initiative, the company developed an AB test.

In this context, an AB test is defined in the literature by (Beasley 2013) as an experiment where two different competing designs or features of a webpage are tested simultaneously based on a predefined metric for evaluating the test success. One of the most common metrics is conversion rate, that can be defined as the number of users that execute an action (e.g. a purchase), compared to the number of visitors. Once the metric is defined, the test randomly assigns a percentage of the users to each page and compares the results (Beasley 2013).

The aim of this AB test was to understand if having measurements per size would be beneficial for customers and understand how this would translate in terms of conversion and return rates.

Hence, when entering the PDP of one of the products in the test, 50% of the users would see all the measurements (waist, shoulder, bust…) per each size available while the rest would see only the generic product description.

Given the inability of accessing all the products of its catalogue, currently, Farfetch is not being capable of measuring every item. Hence, to gather the necessary information for the test, two main sources were used: *Browns* and partner brands. As explained in section 1.1, *Browns* can be seen as the offline extension of Farfetch.com and, having access to its inventory, *Browns* measurements per size information was manually collected at their warehouse and uploaded to Farfetch website. The second resource was possible due to good relationships with some partner brands that already possessed this information.

The focus of the test was to analyse the impact of the initiative on clothing and shoes items, as those families account for around 80% of Farfetch total revenues.

As expected, the results of the AB test were optimistic, showing positive uplifts of 2,6 % for clothing and 1% for Shoes in terms of conversion rate. Additionally, when analysing the results, it was possible to understand in detail the effects of the initiative according to the different regions, customer types, brands, product categories, among others. Analysing these variables, with this level of granularity, allows the company to identify what the key actions are and what products should be targeted.

5.2.3 Roll-out Strategies

Based on customer feedback and the AB test results, the importance of having measurements per size is undeniable. It proved to be capable of increasing conversion and eliminating one of the major pain points of the customer - sizing information. However, as a marketplace, to obtain all this information about the products that are part of its portfolio, Farfetch needs to be capable of identifying the best strategies and recognising the crucial partners who can efficiently provide this information.

Hence, three main strategies have been identified: leveraging *Browns*, using partner brands’ information and measuring the items at partners’ boutiques. Due to time constraints, these strategies have not yet been fully implemented but, are determinant to accomplish the customer experience improvements that Farfetch aims to obtain.

The first strategy consists in developing a process to ensure that, for every item that is measured in Browns warehouse, its measurements information is automatically uploaded into Farfetch.com. The second one is to collaborate with Farfetch partner brands. That will require communicating the AB test results, to reinforce its positive impacts, and the development of the adequate mechanisms to ease the upload of the information.

Finally, the last strategy would be to manually collect this information at the boutique partners. This could be done either by asking boutiques to measure the clothing and shoes (for products available on their inventory) or by having Farfetch allocating people that would be responsible for visiting some of the top boutiques and measuring the products.

However, there are several constraints regarding the implementation of such a strategy. On one hand, boutiques do not want their operation to be disrupted. Hence, having external people inside their facilities, with a risk of damaging their high-value products, creates some reluctance. On the other hand, for Farfetch to ask the boutiques to measure the products themselves would require the implementation of an incentive plan and would reduce its control over the collected information. This lack of control can generate errors that would result in a negative experience for the customers.

Nevertheless, despite the potential challenges of the third strategy, the company believes in its value, so it was decided that a complementary work should be developed to consolidate the approach. As said at the beginning of the section, the focus of the initiative was to measure only clothing and shoes products. To support the roll-out of the partner boutique strategy, a prioritized list of products to measure was elaborated. This prioritization would not only make the process more efficient but could also be beneficial for the partners that can easily identify what products to measure.

Prioritization Strategy

With the aim of identifying the prioritized products and taking into account the vast number of brands available in *Farfetch.com*, it was decided to proceed the analysis including only the brands that represent 80% of Farfetch total Revenues in 2018 (Figure 27). Additionally, it is important to stress the fact that all the products that are available on *Browns* have been excluded from the analysis, assuming that its measurements would be obtained as a result of a different process.

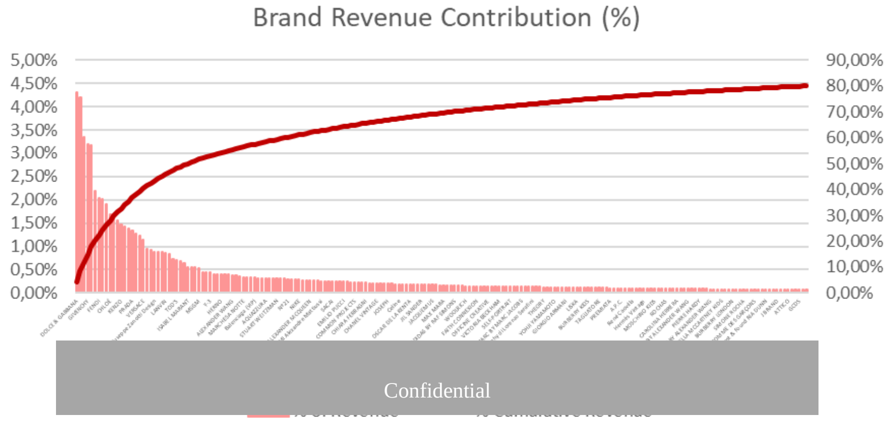


Figure 27 - Brand contribution to Farfetch revenue

After identifying the candidate brands, there was a need to understand what categories (dresses, skirts, trainers, etc) were more relevant to roll out this initiative. After an exploratory analysis, it was found that, for clothing, the ten most sold categories covered more than 90% of the revenue of that family, and the same happened for the top six categories of shoes, so those were the chosen categories were the ones represented in Table 7.

Table 7 - Priority categories to obtain measurements information

Categories	
Clothing	Shoes
Jackets	Trainers
Dresses	Boots
Tops	Sandals
Sweaters & Knitwear	Pumps
Coats	Loafers
Trousers	Mules
T-Shirts & Vests	
Knitwear	
Denim	
Skirts	

In the next phase, for each category, the brands were ranked according to the returned revenue *(Revenue * Return Rate)*. This metric was defined, by business experts, as a good indicator for prioritization as it would allow to capture the effect of both factors, the revenue generated by the brands and their return rates. Afterwards, to make these insights actionable, in a second iteration, there was a need to shorten the number of brands in consideration. To do that, a very practical criterion was used and were only considered, brands with a returned revenue above the average of the category. The obtained output was, for each family and category, a list of the top brands that should constitute a priority regarding the measurements process. In the business, these sets of products with the same brand, family and category are called Brand-Family-Categories (BFCs).

The final step was to combine all the BFCs in a single priority list. In Figure 28 it is possible to have a better understanding of the distribution of the BFCs in terms of return rate and revenue for the company.

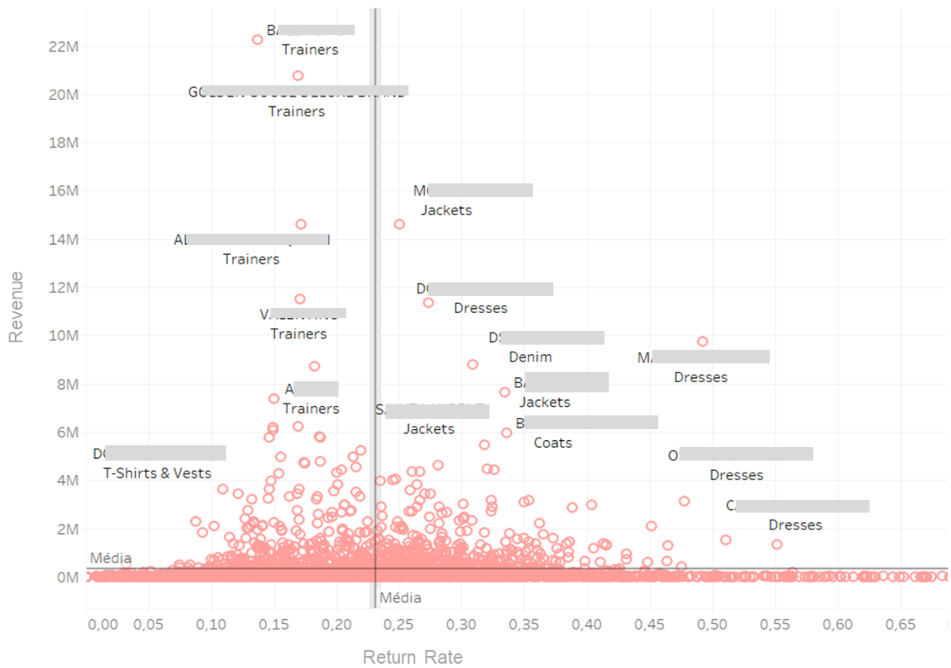


Figure 28 - Return Rate and Revenue distribution of the priority BFCs

The aforesaid procedures were followed as a way of giving the company a double perspective on the problem. On one hand, a ranking of the top priority brands per each family and category of products and on the other hand, a global ranking with the top BFC's that should be priority measured. This way, Farfetch can use this data to communicate with the partners and provide clear guidance on how to efficiently tackle the measurements collection.

As a next step, it would be important to understand exactly how the stock is distributed among the partners. By knowing that, the company will be able to make targeted recommendations and define the appropriate incentive plans to roll-out the initiative. In addition, there will be a need to communicate with the boutiques and understand with them what the best tactics are to acquire this information. Finally, to ensure the implementation is properly done, it is also important to identify the adequate performance metrics and control mechanisms aiming to monitor the process and minimize every possible error.

5.2.4 Implementation and Impact

Before diving into the implementation and its impacts, the challenges associated with this initiative should be reinforced. Working with more than 2500 brands and 600 boutiques, that produces a catalogue of more than three hundred thousand products, creates the need for a strictly planned operation to obtain additional product information. Given that, the importance of this study relies on the identification of the crucial partners and the most efficient manner to tackle the problem.

Nevertheless, all the mentioned strategies are still in a refinement process requiring substantial work, before its implementation. However, based on the AB test results and after identifying product content as one of the main pain points for the customer, providing measurements information will certainly represent improvements for the company. On one hand, it increases the level of information customers have about the products, leading to higher trust and increased purchase intentions. On the other hand, access to more accurate product information is expected to decrease the number of returns, especially in what concerns *Size and Fit* issues.

Besides improving the experience provided to Farfetch customers, the Measurements per Size initiative is expected to allow the company to cover around 25% of its catalogue with better product content, consequently impacting a significant percentage of revenues. Furthermore, the fact that it will enable the company to affect more than 65% of returns (that happen due to fitting issues), is expected to generate a significant impact in terms of profitability by reducing its associated costs.

On top of that, the implementation of this and the other mentioned initiatives to improve product content will trigger the opportunity for their data to be captured and available in the data structure of the company. Such improvements will create the possibility to update the features that are currently included in the developed predictive model, allowing it to learn from this information and improve its return prediction capability.

6 Conclusions and Future Work

The project addressed the issue of returns from a luxury fashion online marketplace perspective. To do that, it was important to position returns along with the customer experience, understanding that their effect goes beyond the returning moment and impacts all phases of the purchase journey.

The main goals of the project were to understand returns and their underlying causes, building a tool that could predict the probability of a given order being returned. During the project, it was required to understand the luxury fashion market and its specificities. Afterwards, the feedback of customers was explored, which provided insights on some of the biggest pain points in the customer journey and how those were influencing returns. Something that was clear from this analysis was the impact of product content on customer perceptions. Then, all the company operations were analysed in detail to understand potential areas of impact for the project. This included the exploration of the ordering process from an internal and a customer perspective as well as everything related to the return procedures and their implications.

After contextualizing, different modelling methods were applied to find the solution that could fit the company’s best interests. The output was a model, capable of predicting returns based on several input features related to the order of a product. On top of this, the model provided insights that created opportunities for Farfetch to leverage its predictive capability.

Finally, given the identified importance of product content, it was decided to support the definition of a roll-out strategy to implement a measurements per size initiative. This initiative will enable the company to provide superior product content to meet the high-end needs of its customers while mitigating the growth of returns.

During the development of the project, there was a possibility of creating awareness for the topics of returns and customer experience. Given that, several opportunities in these areas emerged, triggering new projects and initiatives that should be further explored.

Additionally, there will be a need to support all the mentioned content strategies and contribute to their implementation as a way of improving customer experience and reducing returns. Furthermore, it will be crucial for the company to guarantee that all the relevant data related to it is properly stored in the data structure. Because new features and attributes will be created, these requirements must be passed to the responsible teams that need to define the appropriate location and develop mechanisms for its automatic upload. Having access to the data, will not only enable process monitorization but will also contribute to the inclusion of new features in the returns prediction model. This will improve its predictive capability and enhance its positive impacts on Farfetch.

Since the issue of returns affects most e-commerce players, the applicability of the predictive model is not exclusive to Farfetch and could be extended to other companies and other contexts where returns have a significant impact on business and customer experience.

References

Achille, Antonio, Nathalie Remy, and Sophie Machessou. 2018. "Apparel, Fashion & Luxury Group - The Age of Digital Darwinism." https://www.mckinsey.com/~/media/McKinsey/Industries/Retail/Our Insights/Luxury in the age of digital Darwinism/The-age-of-digital-Darwinism.ashx.

Andreassen, Tor Wallin, Per Kristensson, Line Lervik-Olsen, A Parasuraman, Janet R McColl-Kennedy, Bo Edvardsson, and Maria Colurcio. 2016. "Linking Service Design to Value Creation and Service Research." *Journal of Service Management* 27 (1): 21–29. https://doi.org/10.1108/JOSM-04-2015-0123.

Angela Hall, Neil Towers, Duncan R. Shaw. 2017. "Understanding How Millennial Shoppers Decide What to Buy." *Journal of Management Development* 23 (7): 635–48. https://doi.org/10.1108/MBE-09-2016-0047.

Beach, Gina. 2018. "How Product Content Syndication Software Helps Brands and Manufacturers Realize More Revenue." 2018. https://productsup.io/blog/product-content-syndication-revenue/.

Beasley, Michael. 2013. "In Practical Web Analytics for User Experience." In *Practical Web Analytics for User Experience*, i–iii. Ypsilanti, Michigan, USA: Morgan Kaufmann. https://doi.org/10.1016/b978-0-12-404619-1.00017-4.

Bergstra, James, and Umontrealca Yoshua Bengio. 2012. "Random Search for HyperParameter Optimization." *Journal of Machine Learning Research* 13: 281–305.

Berry, Leonard L., Eileen A. Wall, and Lewis P. Carbone. 2006. "Managing Service Experience Clues." *Academy of Management Perspectives* 20: 43–57. https://doi.org/10.5465/AMP.2006.20591004.

Bishop, Christopher M. 2007. *Machine Learning and Pattern Recognition*. Springer.

Bjørn-Andersen, Niels, and Rina Hansen. 2011. "The Adoption of Web 2.0 by Luxury Fashion Brands." http://aisel.aisnet.org/confirm2011.

Chapman, Pete, Julian Clinton, Randy Kerber, Thomas Khabaza, Thomas Reinartz, Colin Shearer, and Rüdiger Wirth. 2000. "CRISP-DM 1.0 Step-by-Step." *ASHA Presentation*, 73. https://doi.org/10.1109/ICETET.2008.239.

Court, David, Dave Elzinga, Susan Mulder, and Ole Jørgen Vetvik. 2009. "The Consumer Decision Journey." https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/the-consumer-decision-journey.

Dall'Olmo Riley, Francesca, and Caroline Lacroix. 2003. "Luxury Branding on the Internet: Lost Opportunity or Impossibility?" *Marketing Intelligence & Planning* 21 (2): 96–104. https://doi.org/10.1108/02634500310465407.

Dauriz, Linda, Nathalie Remy, and Nicola Sandri. 2014. "Luxury Shopping in the Digital Age."

Deloitte Company. 2017. “Global Powers of Luxury Goods.” https://www2.deloitte.com/cn/en/pages/about-deloitte/articles/pr-consumers-in-emerging-markets-continue-to-drive-luxury-market-growth.html.

Ershov, Vasiliy. 2018. “CatBoost Enables Fast Gradient Boosting on Decision Trees Using GPUs.” 2018. https://devblogs.nvidia.com/catboost-fast-gradient-boosting-decision-trees/.

Fawcett, Tom. 2006. “An Introduction to ROC Analysis.” *Pattern Recognition Letters* 27 (8): 861–74. https://doi.org/10.1016/j.patrec.2005.10.010.

Gentile, Chiara, Nicola Spiller, and Giuliano Noci. 2007. “How to Sustain the Customer Experience:” *European Management Journal* 25 (5): 395–410. https://doi.org/10.1016/j.emj.2007.08.005.

Google. 2012. “Customer Insights.” *Marketing Review St. Gallen* 29 (2): 61. https://doi.org/http://dx.doi.org/10.1007/s11621-012-0126-5.

Guercini, Simone, and Andrea Runfola. 2015. “Internationalization through E-Commerce. the Case of Multibrand Luxury Retailers in the Fashion Industry.” In *Advances in International Marketing,* 26:15–31. https://doi.org/10.1108/S1474-797920150000026002.

Heimerl, Florian, Steffen Lohmann, Simon Lange, and Thomas Ertl. 2014. “Word Cloud Explorer: Text Analytics Based on Word Clouds.” *Proceedings of the Annual Hawaii International Conference on System Sciences,* 1833–42. https://doi.org/10.1109/HICSS.2014.231.

Hennigs, Nadine, Klaus Peter Wiedmann, Christiane Klarmann, Suzane Strehlau, Bruno Godey, Daniele Pederzoli, Agnes Neulinger, et al. 2012. “What Is the Value of Luxury? A Cross-Cultural Consumer Perspective.” *Psychology and Marketing* 29 (12): 1018–34. https://doi.org/10.1002/mar.20583.

Heuer, David, Malte Brettel, and Jan Kemper. 2015. “Brand Competition in Fashion E-Commerce.” *Electronic Commerce Research and Applications*. https://doi.org/10.1016/j.elerap.2015.07.007.

Hjort, Klas, and Björn Lantz. 2016. “The Impact of Returns Policies on Profitability: A Fashion e-Commerce Case.” *Journal of Business Research* 69 (11): 4980–85. https://doi.org/10.1016/j.jbusres.2016.04.064.

Hyeonsoo Kim, Yun Jung Choi, Yuri Lee. 2007. “Web Atmospheric Qualities in Luxury Fashion Brand Web Sites.” *Journal of Management Development* 23 (7): 635–48. https://doi.org/10.1108/MBE-09-2016-0047.

Kapferer, Jean-Noël. 1997. “Managing Luxury Brands.” *Journal of Brand Management*. https://doi.org/10.1057/bm.1997.4.

Klaus, Philipp, and Stan Maklan. 2012. “EXQ: A Multiple-Item Scale for Assessing Service Experience.” *Journal of Service Management* 23 (1): 5–33. https://doi.org/10.1108/09564231211208952.

Kotu, Vijay, and Bala Deshpande. 2015. *Predictive Analytics and Data Mining*. Morgan Kaufmann.

Lemon, Katherine N., and Peter C. Verhoef. 2016. “Understanding Customer Experience Throughout the Customer Journey.” *Journal of Marketing* 80 (6): 69–96. https://doi.org/10.1509/jm.15.0420.

Lipkin, Michaela. 2007. “Customer Experience Formation in Today’s Service Landscape.” *Journal of Management Development* 23 (7): 635–48. https://doi.org/10.1108/MBE-09-

2016-0047.

Maechler, Nicolas, Kevin Neher, and Robert Park. 2016. "From Touchpoints to Journeys: Seeing the World as Customers Do." *McKinsey Digital*, no. March: 1–14. https://doi.org/http://dx.doi.org/10.1103/PhysRevLett.28.1516.

McColl-Kennedy, Janet R., Anders Gustafsson, Elina Jaakkola, Phil Klaus, Zoe Jane Radnor, Helen Perks, and Margareta Friman. 2015. "Fresh Perspectives on Customer Experience." *Journal of Services Marketing* 29 (6–7): 430–35. https://doi.org/10.1108/JSM-01-2015-0054.

McKinsey. 1990. "The Luxury Industry: An Asset for France." *McKinsey*.

Miguéis, Vera L., and Henriqueta Nóvoa. 2017. "Exploring Online Travel Reviews Using Data Analytics: An Exploratory Study." *Service Science* 9 (4): 315–23. https://doi.org/10.1287/serv.2017.0189.

Nisar, Tahir M., and Guru Prabhakar. 2017. "What Factors Determine E-Satisfaction and Consumer Spending in e-Commerce Retailing?" *Journal of Retailing and Consumer Services*. https://doi.org/10.1016/j.jretconser.2017.07.010.

Okonkwo, Uché. 2009. "The Luxury Brand Strategy Challenge." *Journal of Brand Management* 16 (5–6): 287–89. https://doi.org/10.1057/bm.2008.53.

———. 2010. *Luxury Online: Styles, Systems, Strategies*. *Luxury Online: Styles, Systems, Strategies*. New York: Palgrave Macmillan. https://doi.org/10.1057/9780230248335.

Ostrom, A. L., A. Parasuraman, D. E. Bowen, L. Patrício, and C. A. Voss. 2015. "Service Research Priorities in a Rapidly Changing Context." *Journal of Service Research*. https://doi.org/10.1177/1094670515576315.

Probst, Philipp, Marvin N. Wright, and Anne Laure Boulesteix. 2019. "Hyperparameters and Tuning Strategies for Random Forest." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 9 (3). https://doi.org/10.1002/widm.1301.

Prokhorenkova, Liudmila, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, and Andrey Gulin. 2017. "CatBoost: Unbiased Boosting with Categorical Features," no. Section 4: 1–23. http://arxiv.org/abs/1706.09516.

Puccinelli, Nancy M., Ronald C. Goodstein, Dhruv Grewal, Robert Price, Priya Raghubir, and David Stewart. 2009. "Customer Experience Management in Retailing: Understanding the Buying Process." *Journal of Retailing* 85 (1): 15–30. https://doi.org/10.1016/j.jretai.2008.11.003.

Saarijärvi, Hannu, Ulla Maija Sutinen, and Lloyd C. Harris. 2017. "Uncovering Consumers' Returning Behaviour: A Study of Fashion e-Commerce." *International Review of Retail, Distribution and Consumer Research* 27 (3): 284–99. https://doi.org/10.1080/09593969.2017.1314863.

Schwager & Meyer, Andre & Christopher. 2007. "Understanding Customer Experience." *Harvard Business Review*, 117–26.

Sundström, Malin, Sara Hjelm-Lidholm, and Anita Radon. 2019. "Clicking the Boredom Away – Exploring Impulse Fashion Buying Behavior Online." *Journal of Retailing and Consumer Services* 47 (March 2018): 150–56. https://doi.org/10.1016/j.jretconser.2018.11.006.

Teixeira, Jorge, Lia Patrício, Nuno J. Nunes, Leonel Nóbrega, Raymond P. Fisk, and Larry Constantine. 2012. "Customer Experience Modeling: From Customer Experience to Service Design." Edited by Rohit Verma. *Journal of Service Management* 23 (3): 362–76. https://doi.org/10.1108/09564231211248453.

Verhoef, Peter C., Katherine N. Lemon, A. Parasuraman, Anne Roggeveen, Michael Tsiros, and Leonard A. Schlesinger. 2009. “Customer Experience Creation: Determinants, Dynamics and Management Strategies.” *Journal of Retailing* 85 (1): 31–41. https://doi.org/10.1016/j.jretai.2008.11.001.

Vigneron, Franck, and Lester W Johnson. 2004. “Measuring Perceptions of Brand Luxury.” *Journal of Brand Management* 11 (6): 484–506. https://doi.org/10.1057/palgrave.bm.2540194.

Zomerdijk, Leonieke G., and Christopher A. Voss. 2010. “Service Design for Experience-Centric Services.” *Journal of Service Research* 13 (1): 67–82. https://doi.org/10.1177/1094670509351960.