

Visão holística do crescimento da base de clientes numa empresa de telecomunicações

Paulo Renato Martins Pereira

Dissertação de Mestrado

Orientador na FEUP: Prof^ª. Maria João Pires



Mestrado Integrado em Engenharia e Gestão Industrial

2019-07-01

Aos meus pais,

Resumo

Este projeto foi desenvolvido no departamento de *Customer Relationship Management* da NOS Comunicações e teve como principal objetivo identificar os principais fatores que influenciam a evolução do número de clientes TV por subscrição.

Nesse sentido, foi construído um modelo preditivo que permite no curto prazo prever o número de movimentos que ocorrem para dentro e fora da empresa.

Assim, por um lado, é realizada a previsão do número de clientes captados com recurso a um método de previsão, o *Holt Winters*. Como exceção, parte das vendas de serviços de fibra é estimada com base na data sua instalação em novas zonas e na estimativa de conversão em vendas brutas.

Por outro lado, é realizada uma previsão do número de clientes que cancelam os seus serviços com recurso a um método causal. Fatores como a tecnologia do serviço do cliente, o canal pela qual entra em retenção e o momento do período de fidelização no qual se encontra influenciam os valores de *churn*.

No caso de o cliente não efetuar o pagamento, o *churn* é influenciado pelo ciclo de faturação ao qual o cliente pertence e a respetiva data de faturação.

Para que um cliente seja captado ou perdido pela empresa, há uma sequência temporal que este tem de percorrer. Assim, para ambas as situações é efetuada uma estimativa do tempo que decorre entre a venda, a entrada em retenção ou o atraso no pagamento até que o cliente é captado ou perdido, bem como as respetivas taxas de conversão.

Os resultados mostram um boa *accuracy* do modelo, sendo que, à medida que o mês decorre existe um ganho de informação real quanto às etapas intermédias dos processos de captação e *churn*, e consequentemente, a precisão dos valores de fecho do mês vai aumentando.

Holistic view of the customer base growth in a telecommunications company

Abstract

This project was developed in the Customer Relationship Management department of NOS Comunicações and the main objective consisted in identifying the main factors that influence the evolution of the number of Pay-TV clients.

In this sense, a predictive model was constructed that allows in the short term to predict the number of movements that occur both inside and outside the company.

Thus, on the one hand, it was performed a forecast of the number of customers captured using a forecasting method, Holt Winters. As an exception, part of fiber services' sales is estimated based on the installation date in new zones and on the estimate of gross sales conversion.

On the other hand, a forecast of the number of customers who cancel their services is executed, applying a causal method. Factors such as the customer service technology, the channel by which the client enters into retention, and the moment of the loyalty period are found to influence churn values.

In case the customer does not make the payment, churn is influenced by the billing cycle to which the customer belongs and the respective billing date.

For a customer to be picked up or lost by the company, there is a time sequence that the customer has to go through. Thus, for both situations, it is performed an estimate of the time elapsed between sale, retention entry or payment delay until the customer is captured or lost, as well as the respective conversion rates.

The results show a good accuracy of the model. As the month elapses there is a gain of real information regarding the intermediate steps of the capture and churn processes, and consequently, the precision of the closing values of the month increases.

Agradecimentos

À NOS pela oportunidade e excelentes condições de trabalho que me proporcionaram.

À equipa de *Customer Relationship Management* na qual estive inserido, em especial, ao André Campos e João Gomes, que me acompanharam neste projeto e que me transmitiram conhecimentos e ensinamentos importantes para o sucesso do mesmo bem como para o meu futuro profissional.

À minha orientadora da FEUP, Professora Maria João Pires, pela disponibilidade e apoio prestado bem como a todos os professores que participaram no meu percurso académico.

A todos os meus amigos e namorada, que me acompanharam ao longo dos últimos anos e que contribuíram para o meu sucesso académico.

De uma forma especial aos meus pais, a quem dedico esta dissertação, pelo apoio incondicional e valores que sempre me transmitiram.

Índice de Conteúdos

1	Introdução	1
1.1	Enquadramento do projeto e motivação	1
1.2	NOS e o Mercado das telecomunicações	1
1.2.1	A NOS	1
1.2.2	Setor das telecomunicações	2
1.3	Objetivos do projeto	3
1.4	Método seguido no projeto	3
1.5	Estrutura da dissertação	4
2	Revisão da Literatura	5
2.1	Vendas no mercado das telecomunicações	5
2.2	Churn e retenção no mercado das telecomunicações	5
2.2.1	Churn	6
2.2.2	Retenção	7
2.3	Previsão de vendas e churn	8
2.4	Métodos de Previsão	9
2.4.1	Holt Winters	10
2.4.2	Análise dos erros de previsão	11
2.4.3	Regressão linear	12
3	Descrição do projeto	14
3.1	Funil de Vendas	14
3.1.1	Vendas Brutas	14
3.1.2	Gross-Adds	17
3.2	Funil de retenção	20
3.2.1	Retenção por Canal	21
3.2.2	Período de fidelização	21
3.2.3	Série Temporal	22
3.3	Churn involuntário	23
3.3.1	Série temporal	24
4	Modelo preditivo da evolução da base de clientes	26
4.1	Previsão dos gross-adds	26
4.1.1	Vendas brutas HFC	27
4.1.2	Vendas brutas DTH	28
4.1.3	Vendas brutas FTTH	28
4.1.4	Feriados e dias especiais	30
4.1.5	Taxas de conversão de vendas brutas em gross-adds	32
4.1.6	Tempo entre venda e gross-add	32
4.2	Previsão do churn	34
4.2.1	Previsão do churn voluntário	34
4.2.2	Previsão do churn involuntário	38
5	Resultados	41
5.1	Previsão dos gross-adds	41
5.2	Previsão do churn voluntário	45
5.3	Previsão do churn involuntário	47
6	Conclusões	49
	Referências	51
ANEXO A:	Entradas em Retenção por dia da semana e canal	53
ANEXO B:	Churn voluntário para cada tecnologia	54
ANEXO C:	Holt Winters	55

ANEXO D:	Resultados Regressão - ANOVA	56
ANEXO E:	Previsão de <i>gross-adds</i> cabo - abril.....	58
ANEXO F:	Previsão de <i>gross-adds</i> DTH - abril	59
ANEXO G:	Previsão de <i>gross-adds</i> cabo – maio.....	60
ANEXO H:	Previsão de <i>gross-adds</i> DTH – maio	61
ANEXO I:	Previsão do <i>churn</i> voluntário cabo.....	62
ANEXO J:	Previsão do <i>churn</i> voluntário DTH	63
ANEXO K:	Previsão do <i>churn</i> involuntário cabo	64
ANEXO L:	Previsão do <i>churn</i> involuntário DTH.....	65

Siglas

ANACOM – Autoridade Nacional de Comunicações

CRM – *Customer Relationship Management*

DLP – Data limite de pagamento

DTH – *Direct-to-home satellite broadcasting*

FD – *Final disconnect*

FTTH – *Fiber-to-the-home*

HD – *Hard disconnect*

HFC – *Hybrid fiber-coaxial*

IVR - *Interactive voice response*

MAPE – *Mean absolute percentual error*

SD – *Soft disconnect*

UA – Unidade de Alojamento

Índice de Figuras

Figura 1 - Quotas de mercado por segmento in Relatório e Contas 2018 - NOS.....	2
Figura 2 – Funil de vendas	14
Figura 3 - Série temporal de vendas brutas para um mês ilustrativo.....	15
Figura 4 - Expansão da rede de FTTH in Relatório e Contas 2018 - NOS	16
Figura 5 - Logótipos das empresas detentoras de rede fibra	16
Figura 6 - Série temporal - Campanha comercial.....	17
Figura 7 - Taxas de conversão de venda bruta em <i>gross-add</i> por tecnologia	18
Figura 8 - Tempos entre venda bruta e <i>gross-add</i> por dia da semana e tecnologia.....	18
Figura 9 - Tempos entre venda bruta e <i>gross-add</i> por dia da semana e tipo de FTTH	19
Figura 10 - Tempo entre venda bruta e <i>gross-add</i> por momento do mês.....	19
Figura 11- Entradas em retenção por canal	21
Figura 12 - Taxas de não retidos por escalão do período de fidelização.....	22
Figura 13 - Distribuição diária do <i>churn</i> voluntário mensal	23
Figura 14 - Régua de <i>dunning</i>	24
Figura 15 - Régua de <i>dunning</i> - Acordo de pagamento.....	24
Figura 16 - Distribuição diário do <i>churn involuntário</i> mensal	25
Figura 17 - Esquema da previsão de vendas brutas.....	27
Figura 18 - Vendas brutas HFC – Sem <i>outbound</i>	27
Figura 19 - Vendas brutas HFC – <i>Outbound</i>	27
Figura 20 - Vendas brutas DTH – Sem <i>outbound</i>	28
Figura 21 – Vendas brutas DTH – <i>Outbound</i>	28
Figura 22 - Taxas de captação de FTTH	29
Figura 23 - Vendas brutas FTTH - Células antigas	30
Figura 24 - Vendas brutas DST - Células antigas	30
Figura 25 – Dias especiais.....	30
Figura 26 - Dias especiais por tecnologia.....	31
Figura 27 - Dias especiais - Taxas de redução	31
Figura 28 - Distribuição dos <i>gross-adds</i> por dia da semana e semana do mês	32
Figura 29 - Cálculo do número de <i>gross-adds</i> diários.....	33
Figura 30 - Regressão das idas à retenção - HFC.....	35
Figura 31 - Regressão das idas à retenção - FTTH.....	35
Figura 32 - Regressão das idas à retenção - DTH	35
Figura 33 - Taxas de não retidos por escalão do período de fidelização - <i>Inbound</i>	36
Figura 34 - Estimativa do tempo entre entrada em retenção e <i>churn</i> - HFC.....	37
Figura 35 - Estimativa do tempo entre entrada em retenção e <i>churn</i> - HFC e DTH.....	38

Figura 36 - Previsão do <i>churn</i> involuntário - Etapas intermédias	39
Figura 37 - Estimativa do tempo entre HD e <i>churn</i> - HFC	40
Figura 38 - Cálculo de <i>net-adds</i>	40
Figura 39 - Primeira previsão de vendas brutas	41
Figura 40 - Previsão de vendas brutas mais atualizada	41
Figura 41 - Primeira previsão diária de <i>gross-adds</i>	42
Figura 42 – Primeira previsão acumulada de <i>gross-adds</i>	42
Figura 43 - Previsão diária mais atualizada de <i>gross-adds</i>	43
Figura 44 – Previsão acumulada mais atualizada de <i>gross-adds</i>	43
Figura 45 - Primeira previsão de vendas brutas	44
Figura 46 - Primeira previsão diária de <i>gross-adds</i>	44
Figura 47 - Primeira previsão acumulada de <i>gross-adds</i>	44
Figura 48 - Previsão diária mais atualizada de <i>gross-adds</i>	45
Figura 49 - Previsão acumulada mais atualizada de <i>gross-adds</i>	45
Figura 50 - Primeira previsão diária de <i>churn</i> voluntário	46
Figura 51 - Primeira previsão acumulada de <i>churn</i> voluntário	46
Figura 52 - Primeira previsão diária de <i>churn</i> involuntário	48
Figura 53 - Primeira previsão acumulada de <i>churn</i> involuntário	48
Figura 54 - Entradas em retenção por dia da semana e canal.....	53
Figura 55 - <i>Churn</i> involuntário - HFC.....	54
Figura 56 - <i>Churn</i> involuntário - DTH	54
Figura 57 - <i>Churn</i> involuntário - FTTH	54
Figura 58 – <i>Holt Winters</i> DTH - Sem <i>outbound</i>	55
Figura 59 - Regressão - HFC	56
Figura 60 - Regressão - FTTH.....	56
Figura 61 - Regressão - DTH	57
Figura 62 - Primeira previsão de <i>gross-adds</i> cabo - abril.....	58
Figura 63 - Previsão mais atualizada de <i>gross-adds</i> cabo - abril	58
Figura 64 – Primeira previsão de <i>gross-adds</i> DTH - abril	59
Figura 65 – Previsão mais atualizada de <i>gross-adds</i> DTH - abril.....	59
Figura 66 – Primeira previsão de <i>gross-adds</i> cabo - maio	60
Figura 67 - Previsão mais atualizada de <i>gross-adds</i> cabo - maio.....	60
Figura 68 – Primeira previsão de <i>gross-adds</i> DTH - maio.....	61
Figura 69 - Previsão mais atualizada de <i>gross-adds</i> DTH - maio.....	61
Figura 70 - Primeira previsão de <i>churn</i> voluntário cabo.....	62
Figura 71 - Previsão mais atualizada de <i>churn</i> voluntário cabo.....	62
Figura 72 – Primeira previsão de <i>churn</i> involuntário DTH.....	63

Figura 73 – Previsão mais atualizada de <i>churn</i> involuntário DTH	63
Figura 74 – Primeira previsão de <i>churn</i> involuntário cabo	64
Figura 75 – Previsão mais atualizada de <i>churn</i> involuntário cabo	64
Figura 76 – Primeira previsão de <i>churn</i> involuntário DTH.....	65
Figura 77 – Previsão mais atualizada de <i>churn</i> involuntário DTH	65

Índice de Tabelas

Tabela 1 – MAPE – Previsão das vendas brutas de abril	42
Tabela 2 - Desvio percentual – Previsão dos <i>gross-adds</i> de abril	43
Tabela 3 - MAPE - Previsão das vendas brutas de maio	44
Tabela 4 - Desvio percentual - Previsão dos <i>gross-adds</i> de maio	45
Tabela 5 – Desvio percentual e MAPE da previsão do número de entradas em retenção	46
Tabela 6 – Desvio percentual das previsões de <i>churn</i> voluntário	47
Tabela 7 - Desvio percentual da previsão de <i>soft disconnects</i>	47
Tabela 8 - Desvio percentual da previsão de <i>hard disconnects</i>	47
Tabela 9 - Desvio percentual da previsão de <i>churn</i>	48

1 Introdução

1.1 Enquadramento do projeto e motivação

Este projeto de dissertação de mestrado teve a duração de 4 meses e decorreu no âmbito do Mestrado Integrado em Engenharia e Gestão Industrial da Faculdade de Engenharia da Universidade do Porto. O projeto foi realizado na empresa NOS Comunicações na equipa de Análise e *Reporting* de Negócio do departamento de CRM e teve como principal objetivo a previsão da evolução do número de clientes no negócio residencial da empresa.

O número de clientes da empresa num dado momento no tempo está diretamente relacionado com o número de clientes captados e o número de clientes perdidos. A previsão deste número assume-se como uma ferramenta de elevada importância na medida em que permite antecipar o ganho ou a perda de clientes num horizonte temporal próximo e, assim, auxiliar no processo de tomada de decisão e planeamento.

1.2 NOS e o Mercado das telecomunicações

1.2.1 A NOS

A NOS é uma empresa de comunicações e entretenimento portuguesa, que nasceu da fusão, em 2013, de duas das maiores empresas de comunicações do país: a ZON Multimédia e a OPTIMUS Telecomunicações. Esta ação estratégica permitiu às duas empresas combinarem forças na oferta de serviços integrados, assumindo uma clara ambição de crescimento num mercado muito exigente e competitivo como é o das telecomunicações. Desta forma, a NOS é hoje uma empresa mais sólida, responsável e orientada para o futuro, que se compromete com a excelência e a satisfação dos seus clientes.

As telecomunicações são o *core business* do grupo, representando, por isso, a sua maior fonte de rendimento. A empresa oferece soluções fixas e móveis de última geração, de televisão, internet e voz para todos os segmentos de mercado - residencial, pessoal, empresarial e *wholesale*, sendo líder de TV por subscrição e em serviços de banda larga móvel de nova geração.

Para além disso, a NOS é, também, a empresa líder na exibição cinematográfica em Portugal, tendo crescido sustentadamente a sua quota de mercado nos últimos anos. Conta com mais de 200 salas de cinema equipadas com tecnologia de ponta e ainda o maior negócio de distribuição de conteúdos audiovisuais em Portugal. Opera neste mercado através da aquisição e gestão de direitos, da distribuição de filmes e séries de produtores independentes para cinema, *home entertainment* e televisão. A empresa fornece também canais e serviços aos diversos operadores, detendo no seu portfólio canais como TVCine, Hollywood, Panda, entre outros.

No final de 2018, a empresa contava com 1.623 milhões de clientes TV por subscrição, com uma quota de mercado de 42%. Detinha, ainda, 25% da quota de mercado de voz móvel e

cerca de 40% de todos os pacotes *multiple-play* (dois ou mais serviços), contando com cerca de 767 milhares de clientes convergentes, sendo este um indicador importante no mercado das telecomunicações pelo facto destes clientes integrarem diferentes serviços.

Em termos financeiros, em 2018, a NOS apresentou 1.576€ milhões em receitas, sendo que 1.506€ milhões provêm do negócio das telecomunicações, o que ilustra a supremacia das telecomunicações no grupo. O resultado líquido consolidado foi de 141,4€ milhões, o que correspondeu a um aumento de 15,8% face ao ano anterior.

1.2.2 Setor das telecomunicações

A NOS insere-se num dos mercados mais competitivos em Portugal e no Mundo, o das telecomunicações. Este setor é um dos mais importantes para a economia nacional, gerando milhares de milhões de euros anualmente. Para além disso, trata-se de um mercado altamente desenvolvido em termos tecnológicos, devido à existência de operadoras que primam pela qualidade dos seus serviços através da constante inovação, apresentando níveis muito elevados de investimento em infraestruturas e penetração de serviços.

Atualmente, no nosso país, os consumidores têm acesso a ofertas de comunicações e entretenimento bastante avançadas, o que conduziu a níveis de penetração de serviços consideravelmente acima de outros países europeus.

Em termos de receitas de mercado, depois de um longo período de declínio anual significativo, que se prolongou até ao ano de 2016, que se deveu essencialmente a níveis de concorrência e descontos em pacotes bastante intensos, o mercado começou a apresentar uma recuperação e estabilização graduais desde então.

Nos dias de hoje, o setor das telecomunicações em Portugal conta com quatro principais operadores: NOS Comunicações, Altice, Vodafone e Nowo.

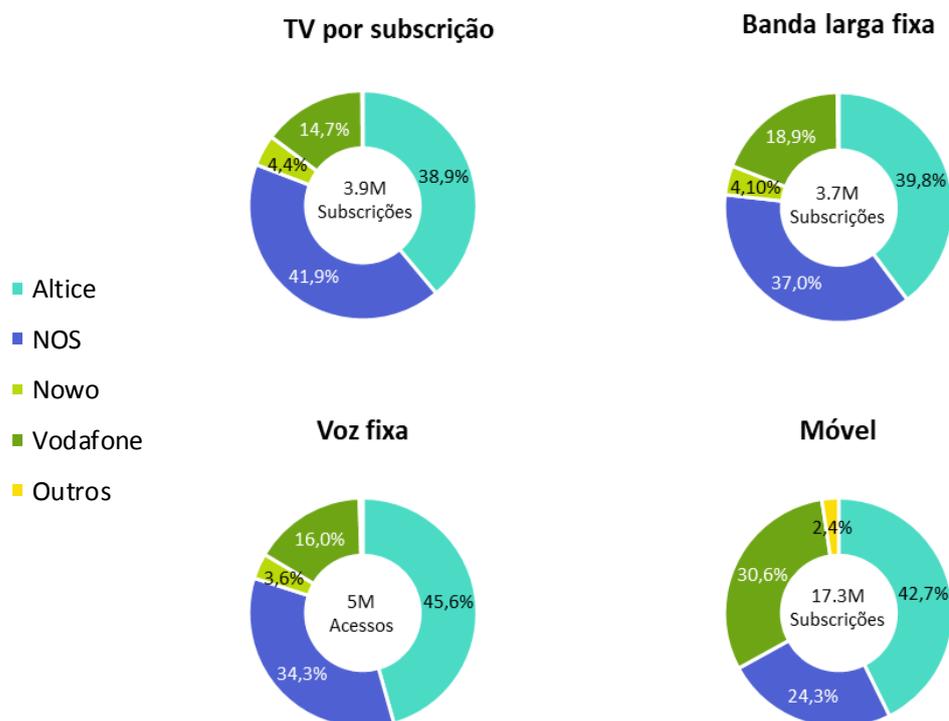


Figura 1 - Quotas de mercado por segmento in Relatório e Contas 2018 - NOS

A Altice Portugal é uma subsidiária do grupo multinacional Altice, desde 2015, e apresenta-se como a líder de mercado nos setores de banda larga fixa, voz fixa e móvel, enquanto que a NOS, o único grupo totalmente nacional, é líder em TV por subscrição.

A Vodafone Portugal é também uma subsidiária, do famoso grupo multinacional britânico, sendo um *player* importante no mercado português, especialmente, no mercado móvel, onde se assume como a segunda líder de mercado.

Por último, a Nowo surgiu em 2016 como resultado do *rebranding* da marca Cabovisão, é a quarta maior operadora em Portugal e tem agitado o mercado desde então.

1.3 Objetivos do projeto

O projeto desenvolvido teve como principal objetivo construir um modelo preditivo da base de clientes particulares de TV por subscrição, de forma a apoiar no processo de tomada de decisão. A finalidade é, assim, que antes do início de cada mês se preveja, com toda a informação recolhida até ao momento, qual será o número de adições/perdas líquidas de clientes para a empresa no fecho do mês. Com o decorrer do mês, as previsões devem ir sendo renovadas, apurando-se com uma maior *accuracy* os resultados finais do fecho.

Com a construção do modelo preditivo, é feito um levantamento dos efeitos que impactam a evolução da base de clientes, percebendo-se as principais variáveis que influenciam os movimentos dos clientes que entram e saem do parque da empresa.

Desta forma, é possível antever as variações na base de clientes da empresa num horizonte próximo e tomar medidas corretivas e preventivas para contrariar potenciais perdas com um grande impacto nos resultados operacionais da empresa.

1.4 Método seguido no projeto

A metodologia seguida neste projeto pode ser dividida em duas fases principais: a construção de um modelo de previsão do número de clientes captados pela empresa, aos quais se dão o nome de *gross-adds* e a construção de um outro modelo que prevê o número de clientes que cancelam os seus serviços – *churn*.

A venda bruta ocorre quando o consumidor pretende aderir aos serviços da empresa e um *gross-add*, quando efetivamente se torna num novo cliente para a operadora e, portanto, apenas uma parte das vendas brutas se torna, efetivamente, um *gross-add*.

Desta forma, na primeira fase do projeto, foi realizada uma previsão do número de vendas brutas com base nos dados históricos e em fatores como a tecnologia do serviço a que está associada a venda (cabo, satélite ou fibra), a instalação de fibra ótica em novas zonas, os canais de vendas, entre outros.

Seguidamente, interessa estimar as taxas de conversão de vendas brutas em *gross-adds*, em que a tecnologia, o dia da semana da venda e a semana do mês são variáveis que apresentam uma influência significativa.

Por último, tendo-se já uma estimativa dos valores de *gross-adds* resultantes das vendas brutas realizadas em cada dia, é necessário prever o dia em que esta conversão se dá por concluída. Após a venda bruta se efetuar, é realizado um agendamento da ativação e apenas quando se conclui a ativação dos serviços, o cliente pode ser considerado um *gross-add*. É este tempo que decorre desde que a venda se efetua até que o cliente se tornar um *gross-add* que necessita de ser previsto.

No final desta fase, é esperado que o modelo preveja o número de adições brutas para a empresa, usando o mês de abril como um teste para avaliar a precisão do modelo e apurarem-se as melhorias necessárias.

Na segunda fase do projeto, foi realizada a previsão do *churn*, que, desde logo, se pode distinguir em *churn* voluntário e *churn* involuntário. Pelo facto de representarem processos completamente distintos e serem influenciados por diferentes variáveis, a previsão do *churn* voluntário e involuntário teve, necessariamente, de seguir diferentes abordagens.

O *churn* voluntário (a vontade de cancelar os serviços parte do cliente) é previsto com base no processo de retenção. Em regra, todos os clientes que cancelam os seus serviços têm de passar pela retenção, onde a empresa apresenta propostas de refidelização com o objetivo de manter o cliente. Assim, em primeiro lugar, é feita uma estimativa do número de clientes que passa pela retenção, sendo o momento do período de fidelização dos clientes (que varia de 0 a 24 meses) a principal variável a considerar. Seguidamente, de acordo com o número de idas à retenção, sabe-se que apenas parte se tornará *churn*, como consequência daqueles que não aceitaram qualquer proposta de refidelização. Por essa razão, interessa também estimar as taxas de clientes não retidos que depende de fatores como a tecnologia do serviço, o canal e o dia em que entraram em retenção.

É também necessário fazer uma estimativa dos *timings* do processo, de forma a saber-se, de acordo com as idas à retenção num dado dia, quantas irão dar origem a *churners* e quando.

Quanto ao *churn* involuntário (a empresa procede ao desligamento por falta de pagamento), é previsto com base numa estimativa das taxas de conversão para as principais fases da régua de *dunning* que se inicia na ausência de pagamento por parte do cliente até este ser considerado de *churn*. Nesta fase, há que se considerar os acordos de pagamento realizados durante este processo que alteram os *timings* face ao processo normal.

Concluídas as fases do projeto, foram introduzidos automatismos ao modelo de forma a que com o mínimo esforço humano fossem obtidas as previsões para cada mês sem comprometer a qualidade das mesmas. Posto isto, espera-se que o modelo esteja completamente operacional, sendo capaz de antever os ganhos ou perdas de clientes no curto prazo.

1.5 Estrutura da dissertação

A presente dissertação é constituída por seis capítulos.

No presente capítulo, foi feita uma introdução à empresa e ao projeto realizado, fazendo-se um enquadramento do contexto em que o mesmo se desenrolou, definindo-se os objetivos e a metodologia utilizada na construção do modelo preditivo.

O segundo capítulo apresenta o estado da arte das áreas científicas relevantes para a execução deste projeto, abordando os métodos científicos usados para a construção do modelo de previsão bem como uma revisão dos principais temas do projeto, as vendas e o *churn* no mercado das telecomunicações.

No terceiro capítulo, são apresentados os principais processos que se encontram na base de cada previsão bem como uma breve análise às variáveis macro do projeto, percebendo-se, desde logo, quais serão os principais fatores que têm influência nas mesmas.

No capítulo quatro, é abordado, de forma detalhada, o modelo preditivo, sendo descritas todas as etapas percorridas no desenvolvimento do mesmo.

No quinto capítulo, são apresentados os resultados obtidos com a implementação do modelo para os meses de abril e maio.

As principais conclusões deste projeto de dissertação são sumarizadas no último capítulo, bem como algumas propostas de melhoria para o futuro.

2 Revisão da Literatura

2.1 Vendas no mercado das telecomunicações

“The telecom industry reaches into every corner of our economies, societies, and private lives, and it is one of the greatest drivers of economic growth and human equality the world has ever seen.” (Ekholm, 2018)

O setor das telecomunicações é caracterizado por uma intensa competição, sendo a diferenciação entre empresas e serviços prestados bastante reduzida, o que implica que a concorrência seja praticada por via do preço. Os clientes procuram a operadora que possui a melhor oferta, o que leva os responsáveis de marketing das operadoras a adotar estratégias de angariação, como ofertas de descontos, reduções de preço e promoções agressivas, que depois se traduz numa redução da receita média por consumidor (Gröne *et al*, 2018).

O volume de vendas e respetivas receitas bem como o tipo de serviços vendidos dependem do tipo de consumidor que a empresa está a satisfazer. Morreale (2000) propõe uma divisão dos consumidores de telecomunicações em empresarial e residencial, distinguindo ainda vários grupos dentro dos consumidores empresariais de acordo com a dimensão da empresa. Cada um dos grupos tem uma diferente abordagem em termos de oferta e canais através dos quais a empresa tenta vender os seus serviços, sendo destacados pela autora as vendas diretas, os *call centres* e os agentes.

O processo de vendas numa empresa de telecomunicações exige, assim, o conhecimento das necessidades de cada consumidor. Num mercado tão intenso e competitivo, o objetivo passa por vender o serviço que melhor se adequa às necessidades dos seus diferentes tipos de consumidores, que cada vez mais tem um perfil de maior exigência perante os produtos e serviços contratados (Batorsky, 2000).

Assim, a angariação de novos clientes no mercado das telecomunicações apresenta-se com um grau de dificuldade elevada e acarreta maiores custos do que em qualquer outro mercado.

A previsão de vendas, tal como em qualquer outro setor, é influenciada por uma grande diversidade de fatores e assume-se como um dos aspetos cruciais do planeamento das empresas, sendo um elemento importante para o seu sucesso (Harrison, 1967).

2.2 Churn e retenção no mercado das telecomunicações

“The central purpose of managing customer relationships is for the enterprise to focus on increasing the overall value of its customer base – and customer retention is critical to its success.” (Peppers e Rogers, 2004)

Os clientes são considerados um dos ativos mais importantes das empresas, especialmente em mercados competitivos e dinâmicos, como é o das telecomunicações. Nesta indústria, ao comportamento por parte de um cliente de mudar de um dado operador de telecomunicações

para outro dá-se o nome de *churn* (Strouse, 1999) e este é um dos temas que, nos últimos anos, tem vindo a merecer especial atenção por parte das operadoras.

2.2.1 *Churn*

Segundo Mattison (2005), o *churn* pode dividir-se, essencialmente, em dois grupos: *churn* voluntário e involuntário.

O *churn* involuntário deve-se, essencialmente, à falta de pagamento ou a fraudes cometidas por parte do cliente, sendo, em teoria, o mais fácil de identificar e prever, visto ser a empresa que tem o controlo do mesmo e o poder de decidir cancelar o serviço ao cliente e quando.

Por outro lado, a previsão do *churn* voluntário trata-se de uma tarefa teoricamente mais complexa, uma vez que, ocorre quando o cliente toma a decisão consciente de terminar o serviço com a sua operadora (Hadden *et al*, 2008).

Mattison (2005) aponta como principais razões para o *churn* voluntário: o preço, a qualidade do serviço, cobertura, serviço de apoio ao cliente e imagem da marca. O grau de importância que cada uma destas razões assume varia, inevitavelmente, entre os diferentes mercados e culturas.

Ainda de acordo com o mesmo autor, o *churn* voluntário pode ainda subdividir-se em *churn* deliberado e *churn* incidental. O *churn* deliberado ocorre quando o cliente decide mudar para uma operadora concorrente, sendo este o tipo de *churn* que as empresas tentam combater. O *churn* incidental representa, por norma, uma pequena parte do *churn* voluntário e ocorre quando se trata de uma decisão consciente, mas forçada por parte do cliente, devido a algum acontecimento na sua vida, como por exemplo, mudança de residência ou problemas financeiros.

Independentemente do tipo de *churn* que se trata, é certo que ele traz diversas consequências negativas para a empresa. A principal é, naturalmente, a perda direta de receitas. No entanto, esta é apenas uma delas. As empresas podem ser forçadas a reagir ao *churn* com uma redução dos preços de forma a convencer os seus clientes a permanecerem com os seus serviços, o que leva, desta forma, a uma redução nas suas receitas anuais. Para além disso, é natural que também surjam campanhas de reacquirição de clientes, que acarretam mais custos para a empresa, possivelmente recuperados dependendo da taxa de sucesso das mesmas (Mattison, 2005).

A previsão do *churn* que mais interesse tem suscitado e sobre a qual têm surgidos mais estudos é a de tentar prever quais os clientes que irão ser *churn* com base nas suas características e a sua atividade enquanto cliente de forma a sinalizar aqueles que apresentam um maior risco de *churn*. Um dos principais objetivos deste tipo de previsão é auxiliar no planeamento estratégico do processo de retenção. Técnicas de *machine learning* como *decision trees*, *neural networks*, *support vector machines* ou *logistic regressions* são aplicadas nesta previsão (Vafeiadis *et al*, 2015).

Numa perspetiva diferente, Mattison (2005) apresenta a utilidade que um modelo que preveja o número de clientes que irão ser *churn* num determinado espaço de tempo, pode ter na empresa. Este tipo de previsão traz uma série de vantagens a empresa em termos de:

- Gestão de oferta e vendas;
- Planeamento de capacidade;
- Gestão orçamental.

É este tipo de modelo que se pretende construir neste projeto de dissertação.

2.2.2 Retenção

É certo que as vendas assumem um papel crucial em qualquer empresa, no entanto, atendendo à maturidade do mercado das telecomunicações, aos elevados custos de angariar um cliente e ao potencial de crescimento limitado, tem-se assistido a uma mudança estratégica por parte das empresas de telecomunicações, que passa de uma estratégia de angariação de novos clientes para uma estratégia mais defensiva, focada na retenção dos clientes (Keramati, 2011). Esta necessidade leva as empresas a procurarem manter uma relação de fidelidade com os seus clientes, sabendo-se que uma pequena redução na perda de clientes pode representar um aumento significativo das receitas anuais. Assim, de forma a evitar a desistência dos seus serviços, as operadoras apostam os seus recursos no processo de retenção (Jurisic e Azevedo, 2011).

A retenção é definida por Crie (1996) como uma estratégia que visa manter os clientes da empresa em portfolio. Esta é a principal solução encontrada pelas operadoras para combater o *churn*, assumindo uma importância acrescida nas telecomunicações por se tratar de um mercado altamente competitivo e saturado.

Deste modo, as empresas que forem capazes de atrair e reter os clientes ficam melhor posicionadas perante os seus competidores (Jyh-Fu Jeng e Bailey, 2012). Para além disso, um estudo realizado por Livne *et al* (2011) mostra mesmo que a retenção tem um impacto na performance financeira da empresa, havendo uma relação positiva entre a retenção dos clientes e os lucros da mesma no futuro.

Já Rosenberg e Czeipel (2007) resumem as vantagens da retenção a dois conceitos:

- Efetividade. A retenção permite a construção de uma base leal de consumidores e consequentemente, a defesa da quota de mercado da empresa.
- Eficiência. Os custos de aquisição de clientes são considerados bastante superiores aos custos de os reter, levando assim a um aumento dos lucros.

Contudo, os esforços em atrair novos consumidores não podem, de forma alguma, ser postos de parte. É no equilíbrio entre as estratégias que visam atrair novos clientes e as que permitem reter os mesmos que reside a chave do sucesso (Rosenberg e Czeipel, 2007). As principais estratégias têm por base técnicas de *cross-selling* e *up-selling*, sendo por isso oportuno abordar ambos os conceitos.

O *cross-selling* pode ser definido como uma estratégia de venda de novos produtos ou serviços a um cliente que anteriormente já adquiriu outros produtos ou serviços da mesma empresa. Assim, a dependência do consumidor aumenta e por consequência, o *churn* reduz-se. Assume-se, desta forma, como uma ferramenta bastante importante num mercado de telecomunicações caracterizado por uma elevada volatilidade e baixa lealdade dos consumidores (Jaroszewicz, 2008).

O *up-selling*, por sua vez, define-se como uma estratégia de venda na qual a empresa leva o cliente a comprar mais do que uma unidade do mesmo produto ou uma versão de qualidade superior do mesmo (Kamakura, 2008).

O mesmo autor considera que, apesar de tanto o *cross-selling* como o *up-selling* serem técnicas relativamente antigas e já estabelecidas, a forma como são aplicadas alterou-se substancialmente com o aparecimento do conceito de *customer relationship management* e a consequente forma de como as empresas se relacionam com os seus clientes.

As técnicas modernas consistem na determinação dos clientes com maior probabilidade de *churn*, identificando-se aqueles que representam um maior valor para a empresa, sendo depois desenvolvidas as estratégias mais adequadas para o efeito. São usadas ferramentas analíticas no apoio às técnicas de *cross-selling* e *up-selling* que permitem, com base no o

comportamento histórico do cliente e nas suas características identificar o potencial de venda a cada cliente, resultando em abordagens personalizadas e mais persuasivas (Huange *et al*, 2015).

2.3 Previsão de vendas e *churn*

No mundo empresarial atual, as técnicas de *data mining* assumem uma considerável importância quando se pretender prever o futuro, apresentando uma acrescida preponderância no mercado das telecomunicações devido à quantidade de informação que as empresas dispõem (Aronsson, 2015). Estas permitem extrair conhecimento a partir de grandes quantidades de dados e assumem-se como uma ferramenta importante na previsão de vendas e do *churn* nas telecomunicações. Segundo o mesmo autor, algumas das técnicas mais usadas são: classificação, *clustering*, regressão, sequências e *forecasting*.

A classificação é uma forma de agrupar a informação, identificando-se a que categoria ou classe uma determinada observação pertence. As observações existentes são examinadas e infere-se uma série de regras a partir delas. Alguns dos modelos derivados das técnicas de classificação usados nas telecomunicações são árvores de decisão, regras de decisão ou *neural networks* (Han *et al*, 2017).

Por outro lado, o *clustering* consiste na divisão dos dados em grupos que apresentam características/padrões em comum, sendo que a cada grupo dá-se o nome de *cluster* (Berkhin, 2006). Ao contrário da classificação, nenhum campo é definido como *target* cabendo ao algoritmo usado definir os diferentes *clusters*. No setor das telecomunicações, esta técnica é geralmente aplicada quando se pretende segmentar os clientes, permitindo por exemplo perceber a composição atual dos clientes e as características daqueles que são *churn* (Ye *et al*, 2013).

Quanto à previsão de *churn*, tanto as técnicas de classificação como as de *clustering* focam-se em separar os clientes de acordo com o risco de *churn*, atribuindo uma determinada probabilidade de *churn* a cada cliente de acordo com as suas características. Este tipo de previsão assume-se especialmente importante no desenvolvimento das políticas de retenção, representando por isso uma peça fundamental do seu planeamento estratégico (Tsai e Lu, 2009; Lemmens e Croux, 2006; Amina *et al*, 2017).

As sequências são eventos conectados ao longo de um certo período de tempo. A importância que estas podem assumir num modelo de previsão prende-se com o facto de, a observações com uma determinada característica, ser atribuída uma dada probabilidade de possuir uma outra característica, permitindo assim, a inferência de atributos (Aronsson, 2015).

Por sua vez, a regressão procura estabelecer relações entre a variável que se pretende prever (variável dependente) e uma ou mais variáveis que expliquem o seu comportamento (variáveis independentes).

Por último, o *forecasting data mining* é o processo que consiste em estimar os valores futuros de uma variável, sendo que o seu objetivo é encontrar regras que permitam prever o que vai acontecer (Aronsson, 2015). Uma das vertentes deste conceito é o *time series data mining* que tem como propósito extrair conhecimento a partir da forma da série temporal. Podem ser usadas técnicas de classificação e *clustering*, em que se atribui um determinado padrão da série temporal a uma categoria mas também métodos de previsão, como o *Holt Winters* e a decomposição clássica (Esling e Agon, 2012).

Deste modo, e atendendo a que o objetivo do projeto é prever o número de novos clientes para a empresa, é necessário efetuar uma previsão do número de vendas realizadas pelo que foram investigados de uma forma mais profunda os métodos de previsão.

Quanto ao *churn*, o objetivo passa por prever o número de clientes que são *churn* no curto prazo e não quais são esses clientes. Assim, o *churn* é previsto com base no processo que cada cliente tem de percorrer até abandonar a empresa, sendo usada a regressão linear como ferramenta para a previsão do número de clientes que entra em retenção em detrimento das técnicas de *clustering* e classificação, que permitem avaliar os clientes com maior propensão a serem *churn*.

2.4 Métodos de Previsão

Os métodos de previsão representam uma parte integral do processo de planeamento e tomada de decisão da gestão e têm como principal objetivo prever a performance futura com base no desempenho registado no passado.

Uma vez que cada área da empresa se relaciona com as restantes, uma boa ou má previsão pode influenciar toda a empresa. Por exemplo, erros na previsão de vendas podem influenciar negativamente a previsão do budget, despesas operacionais, níveis de inventários, *pricing*, entre outros (Makridakis *et al*, 1998).

Segundo os mesmos autores, os métodos de previsão dividem-se, essencialmente, em dois grupos: métodos quantitativos e métodos qualitativos.

Os métodos quantitativos são usados quando existe informação histórica suficiente e em forma numérica e podem-se, ainda, subdividir em métodos explicativos e métodos aplicados à série temporal.

Os métodos explicativos assumem que o comportamento da variável que está a ser prevista é explicado por uma ou mais variáveis independentes. Assume uma relação causa-efeito entre os inputs do sistema e os seus outputs.

Já os métodos aplicados à série temporal não fazem qualquer tentativa de prever os fatores que influenciam o comportamento da variável, mas, baseiam-se na análise da performance histórica com o objetivo de identificar padrões e extrapolá-los para o futuro. Estes métodos são utilizados em detrimento dos anteriores quando existe alguma dificuldade em identificar os fatores explicativos da variável ou em quantificar as variáveis em causa. A segunda razão para usar um método aplicado à série temporal é que, por vezes, o principal objetivo é apenas prever o que vai acontecer e não a razão pela qual vai acontecer.

Por outro lado, quando a quantidade de informação disponível é mais escassa, são aplicados métodos qualitativos. Estes, por norma, baseiam-se no conhecimento e experiência das pessoas que melhor conhecem os processos. Têm como principal objetivo ajudar o decisor com análises subjetivas e muitas das vezes, complementar os métodos de previsão quantitativos, introduzindo uma componente subjetiva que confere mais autenticidade ao modelo.

Todos os métodos quantitativos e uma grande parte dos métodos qualitativos baseiam-se na premissa de continuidade de que os padrões do passado irão continuar a verificar-se no futuro.

Perante uma grande diversidade métodos de previsão, impõe-se a questão de qual o método que se deve aplicar. Chatfield (1978) defende que a escolha do método de previsão mais adequado depende de uma grande variedade de fatores como o objetivo da previsão, o número de observações disponíveis e o número de variáveis a prever.

Neste projeto, foi usado na previsão das vendas brutas, um método de previsão quantitativo aplicado à série temporal, o amortecimento exponencial de *Holt Winters*.

2.4.1 Holt Winters

O *Holt Winters* insere-se na categoria dos métodos de amortecimento exponencial e adequa-se a séries temporais que apresentem sazonalidade e tendência.

Segundo Makridakis, Wheelright e Hyndman (1998), os métodos de amortecimento exponencial caracterizam-se por atribuírem uma série de pesos aos dados históricos, que, descem de forma exponencial desde os dados mais recentes aos dados mais antigos, implicando que, os dados com menor antiguidade tenham um maior peso na previsão do que as observações mais antigas. Desta forma, é necessário definir três parâmetros que definam a distribuição dos pesos de acordo com a antiguidade dos dados: α , β e γ .

O *Holt Winters*, tal como todos os métodos quantitativos aplicados a séries temporais, assume que os valores das observações históricas são explicados por uma combinação entres os padrões que se registam no passado e o erro aleatório.

Há três diferentes padrões que são considerados neste método. Um deles é o padrão horizontal, ao qual também se pode chamar de nível, sendo que este verifica-se quando os dados flutuam à volta de uma média constante. A tendência é outro padrão, que se regista quando existe um crescimento ou decréscimo da série no longo prazo. Para além destes, tem-se ainda o padrão sazonal, que está presente quando a série temporal regista variações periódicas regulares como por exemplo trimestre, mês ou dia da semana.

Há dois modelos *Holt Winters* que devem ser utilizados de acordo com o tipo de sazonalidade dos dados:

- Sazonalidade multiplicativa, na qual se identifica um aumento ou diminuição da amplitude da componente sazonal de acordo com a tendência da série.
- Sazonalidade aditiva, que se caracteriza por variações sazonais de magnitude constante.

Holt Winters Multiplicativo

Caso o método a aplicar seja o multiplicativo, as componentes do modelo são calculadas por:

- Nível: $N_t = \alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)(N_{t-1} + b_{t-1})$
- Tendência: $b_t = \beta (N_t - N_{t-1}) + (1 - \beta)b_{t-1}$
- Sazonalidade: $S_t = \gamma \frac{Y_t}{N_t} + (1 - \gamma)S_{t-s}$, onde s é o período de sazonalidade.

Para a inicialização do modelo, é necessário atribuir valores iniciais às três componentes, sendo que Makridakis, Wheelright e Hyndman (1998) sugerem:

$$N_s = \frac{1}{s} \sum_{t=1}^s Y_t$$

$$b_s = 0$$

$$S_i = \frac{Y_i}{N_s} \quad (i = 1, \dots, s)$$

A previsão de cada observação futura é dada por:

$\hat{Y}_{t+m} = (N_t + b_t m) S_{t-s+m}$, onde m é o número de unidades de tempo á frente do momento que se pretende prever.

Holt Winters Aditivo

As componentes do modelo aditivo são calculadas por:

- Nível: $N_t = \alpha(Y_t - S_{t-s} + (1 - \alpha)(N_{t-1} + b_{t-1}))$
- Tendência: $b_t = \beta(N_t - N_{t-1}) + (1 - \beta)b_{t-1}$
- Sazonalidade: $S_t = \gamma(Y_t - N_t) + (1 - \gamma)S_{t-s}$

De igual forma, os autores sugerem que estas componentes assumem os seguintes valores para o modelo aditivo:

$$N_s = \frac{1}{s} \sum_{t=1}^s Y_t$$

$$b_s = 0$$

$$S_i = Y_i - N_s \quad (i = 1, \dots, s)$$

Por sua vez, a previsão é dada por:

$$\hat{Y}_{t+m} = N_t + b_t m + S_{t-s+m}$$

O amortecimento exponencial de *Holt Winters* é um dos métodos mais usados devido à sua robustez, facilidade de uso e à boa *accuracy* que caracterizam as suas previsões, o que justifica a sua utilização neste modelo. No entanto, tal como qualquer outro método, apresenta aspetos negativos que devem ser tidos em conta na sua aplicação (Chatfield e Mohammad, 1988).

Estes autores alertam, essencialmente, para três questões:

1. A normalização da componente sazonal, que defendem que esta deve ser feita espontaneamente se o padrão sazonal apresentar variações consideravelmente ao longo do tempo;
2. A divergência de opiniões quanto à escolha dos valores iniciais para as três componentes, sendo a que é proposta por Makridakis, Wheelwright e Hyndman uma das mais aceites;
3. A escolha dos parâmetros de amortecimento, que propõem ser feita pela minimização do erro quadrático médio, alertando também para a escolha dos valores iniciais que pode inviabilizar a otimização dos valores parâmetros.

2.4.2 Análise dos erros de previsão

Segundo Makridakis, Wheelwright e Hyndman (1998), a precisão/*accuracy* de um método refere-se a quão bem o modelo de previsão consegue reproduzir os dados já conhecidos.

De forma a avaliar a precisão do método de previsão ou a efetuar comparações entre diferentes métodos, torna-se necessário fazer uma análise dos erros de previsão, para as quais podem ser utilizadas diferentes métricas.

Uma delas é o erro médio absoluto, que mede o desvio da previsão em relação aos valores reais, em termos absolutos, sendo calculado por:

$$\text{Erro Médio Absoluto} = \frac{1}{m} \sum_{t=1}^m |Y_t - \hat{Y}_t|, \text{ onde } m \text{ é o número de observações usadas.}$$

O erro quadrático médio é também usado como uma métrica que avalia os erros de previsão. Este é determinado pela soma dos erros de previsão ao quadrado e dividindo a mesma soma pelo número de observações usados no cálculo, sendo expresso pela seguinte equação:

$$\text{Erro Quadrático Médio} = \frac{1}{m} \sum_{t=1}^m (Y_t - \hat{Y}_t)^2$$

Tanto os valores do erro médio absoluto como do erro quadrático médio dependem das dimensões do item que está a ser previsto. Nesse sentido, a utilização do erro percentual absoluto médio permite resolver essa questão. Segundo um estudo realizado por Mentzer e Kahn (1995), esta era a métrica mais usada entre os gestores responsáveis por aplicar métodos de previsão. A importância da métrica mantém-se intacta nos dias de hoje. O cálculo é realizado pela média da diferença absoluta entre os valores previstos e reais, expressa em percentagem:

$$\text{Erro Percentual Médio Absoluto} = \frac{1}{m} \sum_{t=1}^m \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \times 100$$

O índice de *Theil* é também útil na medida em que se compara a precisão da previsão com o método de *Naive*, que usa a última observação registada como previsão para o período seguinte. Caso o índice apresente valores inferiores a 1, o método apresenta uma maior precisão do que o método de *Naive*. Registando-se um valor superior a uma unidade, o método de *Naive* deve ser usado em detrimento do método em questão. Assim, quanto mais próximo de zero se encontrar o índice, melhor será a previsão por comparação com a que atribui o valor da observação anterior ao valor a prever (Makridakis *et al*, 1998).

$$U\text{-theil} = \frac{\sqrt{\frac{1}{m} \sum_{t=1}^m (Y_t - \hat{Y}_t)^2}}{\sqrt{\frac{1}{m} \sum_{t=1}^m Y_t^2 + \frac{1}{m} \sum_{t=1}^m \hat{Y}_t^2}}$$

2.4.3 Regressão linear

A regressão linear insere-se no grupo dos métodos de previsão quantitativos explicativos e tem como principal objetivo avaliar a relação entre várias variáveis independentes e a variável que se pretende prever.

Quando existe apenas uma variável independente, a regressão linear simples é a apropriada. Se mais do que uma variável explicarem a variável dependente, a regressão linear múltipla necessita de ser utilizada. Assim, pode-se dizer que a regressão linear simples é um caso especial da regressão linear múltipla na medida em que apenas existe uma variável independente que explica as variações da variável dependente (Makridakis *et al*, 1998).

A regressão linear múltipla assume que existe uma relação linear entre uma variável Y , a variável dependente, e um certo número k de variáveis independentes (X_1, X_2, \dots, X_k). As variáveis independentes são também chamadas de regressores, uma vez que são utilizadas para explicarem a variação da variável dependente.

A formulação geral da regressão múltipla é:

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k + e$$

A relação entre as variáveis independentes e a variável dependente pode assumir diferentes formas: linear, exponencial, quadrática, cúbica, entre outras. O facto de se apenas poder estabelecer relações lineares é uma limitação que pode ser suprida pelo facto de muitas relações não lineares poderem ser transformadas em funções lineares (Makridakis *et al*, 1998).

Com o objetivo de se avaliar se efetivamente as variáveis independentes explicam as variações da variável dependente, existem algumas métricas que devem ser analisadas.

O coeficiente de correlação, r , é uma medida relativa de associação entre duas variáveis. Pode variar de -1 a 1, sendo que se assumir um valor nulo, significa que não existe qualquer correlação entre as variáveis. Caso $r = 1$ ou $r = -1$, as variáveis encontram-se perfeitamente correlacionadas. O sinal do coeficiente indica se se está perante uma correlação positiva ou negativa. Se se estiver perante uma regressão linear múltipla, este designa-se por coeficiente de correlação múltipla e é representado por R .

O r^2 (ou R^2 para a regressão múltipla) traduz, assim, a proporção da variação da variável dependente que é explicada pela equação de regressão, traduzindo quão bem as observações se ajustam à linha da regressão. Assume um valor que se situa no intervalo]0,1].

Contudo, o coeficiente de determinação apresenta algumas limitações. Como refere Barrett (1974) um maior valor de R^2 não implica necessariamente que o modelo de regressão traduza um melhor ajustamento. A adição de uma variável ao modelo aumenta, inevitavelmente, o valor deste coeficiente visto que a soma de quadrados da regressão também aumenta. No entanto, ignora-se a perda na informação devido a uma possível redução nos graus de liberdade, podendo assim tornar o modelo menos eficiente.

Nesse sentido, o coeficiente de determinação ajustado visa corrigir o R^2 , ajustando o numerador e denominador através dos respetivos graus de liberdade. De forma contrária ao R^2 , o coeficiente de determinação ajustado pode assumir um valor inferior se a contribuição da variável adicional na explicação da variação total for inferior ao impacto que essa adição acarreta nos graus de liberdade (Srivastava, 1995).

Os testes à significância são outra ferramenta que deve ser usada de forma a avaliar a significância estatística dos resultados da regressão antes destes serem usados para realizar a previsão. Assim, o objetivo é determinar a confiança que se pode colocar na regressão calculada e a sua aplicabilidade à restante população de possíveis valores (Makridakis *et al*, 1998).

Os *testes-F* avaliam a significância global da equação da regressão, indicando se a relação entre a variável dependente e a combinação das variáveis independentes é estatisticamente significativa.

Por sua vez, os *testes-t* à significância dos coeficientes são também eles relevantes, permitindo avaliar as variáveis que são estatisticamente significativas.

Para além da linearidade, a regressão linear parte de mais pressupostos, que devem ser verificados de forma a validar a utilização da regressão no processo de *forecasting*:

- Independência dos resíduos. A violação desta condição pode significar que os resíduos se encontram correlacionados. Apesar de não influenciar a estimativa dos parâmetros, afeta os testes à significância e o valor do R^2 .
- Homeostacidade. Os resíduos devem apresentar homogeneidade nas variâncias.
- Normalidade dos resíduos. Assume-se que os resíduos seguem uma distribuição normal.

A multicolineariedade deve também ser um tópico a ter em consideração uma vez que na sua presença, os resultados da previsão podem sair bastante afetados, causando erros graves na previsão. Esta é detetada quando uma ou mais variáveis independentes não se revelam estatisticamente significativas, apesar de os valores de R^2 e do teste-F serem bastante elevados. Desta forma, uma das variáveis que se encontra correlacionada deve ser removida (Makridaki *et al*, 1998).

3 Descrição do projeto

Neste capítulo, é apresentada uma análise mais pormenorizada das variáveis macro do problema: as vendas brutas e os *gross-adds* bem como o *churn* voluntário e involuntário, com o objetivo de se perceber quais os principais fatores que as influenciam.

3.1 Funil de Vendas

O funil de vendas abrange o processo que se inicia numa venda bruta, quando existe confirmação da vontade de um consumidor em aderir aos serviços da empresa e que finda com um novo cliente em parque na empresa, ao qual se dá o nome de *gross addition* ou de forma abreviada, *gross-add*. O processo encontra-se representado na Figura 2.

Como etapas intermédias tem-se a venda líquida, que se diferencia da venda bruta por já existir um agendamento da ativação dos serviços, seguida, naturalmente, da própria ativação, a partir da qual o cliente começa a usufruir dos serviços contratados. É importante realçar que nem todas as ativações são, efetivamente, novos clientes para a empresa, visto que algumas dessas ativações correspondem a transferência de morada ou alterações de titularidade, e, portanto, não se trata de um novo cliente para a empresa. Por essa razão, apenas algumas ativações chegam à última etapa do processo, e se convertem em *gross-adds*.

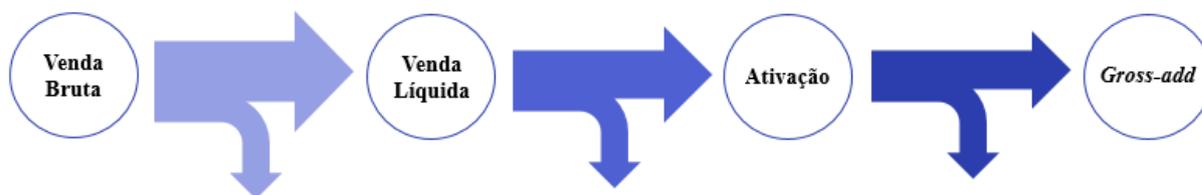


Figura 2 – Funil de vendas

Em todas as fases do processo, existem sucessivas perdas e *timings* que necessitam de ser estimados para que se efetue uma previsão diária do número de *gross-adds*.

3.1.1 Vendas Brutas

A NOS apresenta no seu portfólio uma grande variedade de serviços disponíveis que se podem agrupar de acordo com as necessidades dos clientes, originando os denominados pacotes convergentes. O consumidor pode adquirir desde um serviço de apenas TV até *bundles* que combinam cinco serviços diferentes: TV, Internet fixa e móvel, Voz fixa e móvel.

Os serviços que a NOS tem à disposição dos seus clientes podem também dividir-se em três grupos de acordo com a tecnologia utilizada para a transmissão do sinal:

- Cabo coaxial – em inglês: *Hybrid Fiber Coaxial (HFC)*;
- Fibra - em inglês: *Fiber to the home (FTTH)*;
- Satélite – em inglês: *Direct-to-home satellite broadcasting (DTH)*.

A tecnologia HFC, comumente chamada de cabo, é a que apresenta um maior número de clientes em parque e representa uma grande parte das vendas efetuadas pela empresa todos os meses.

As vendas da tecnologia DTH têm vindo a diminuir ao longo dos anos face à natural substituição desta tecnologia por uma transmissão de sinal por cabo, capaz de assegurar um serviço de maior fiabilidade. Assim, tem-se assistido a um crescente número de migrações de clientes DTH para HFC ou FTTH.

A tecnologia FTTH é a que assegura uma maior qualidade do serviço ao cliente, garantindo uma maior capacidade de transmissão de dados e conexões mais rápidas, estáveis e seguras. Por outro lado, implica um grande investimento por parte das operadoras e por isso, o seu crescimento tem vindo a ser feito de forma sustentada e controlada.

As vendas de FTTH estão bastante dependentes das aberturas de novas zonas, que se iniciam algumas semanas antes da abertura em regime de pré-venda. As vendas diretas e por Outbound, canais de venda que serão explicados mais à frente, assumem aqui um papel importante, sendo os principais responsáveis por estas vendas.

Série temporal

Pela análise da série temporal das vendas brutas e usando um mês ilustrativo representado na Figura 3, denota-se, desde logo, a existência de uma clara sazonalidade semanal, que se justifica pelo facto de ao fim de semana se realizarem um número de vendas brutas bastante inferior aos dias úteis.

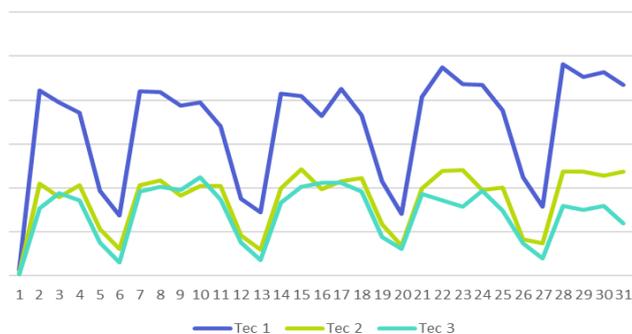


Figura 3 - Série temporal de vendas brutas para um mês ilustrativo

Para além disso, no próprio mês, existe também alguma sazonalidade. Verifica-se que, nos últimos dias do mês, as vendas tendem a aumentar o nível em relação ao do resto do mês, o que pode ser explicado pela necessidade de cumprimento de objetivos que levam a uma maior produtividade por parte dos vendedores no final do mês.

Deste modo, a sazonalidade semanal e mensal são aspetos que necessitam de ser considerados no modelo preditivo de vendas.

Rede de fibra

Como já se tornou perceptível, a fibra é a tecnologia que requer uma especial atenção na previsão de vendas brutas. A fibra ótica foi um grande avanço tecnológico, representando uma revolução na forma como se transmite informação e contribuindo para o desenvolvimento das telecomunicações.

Havendo uma intensa concorrência comercial no mercado português, as operadoras intensificaram o desenvolvimento de infraestruturas nos últimos anos. Sendo uma tecnologia que requer um grande investimento, a expansão da sua cobertura tem de ser feita de uma forma gradual e controlada. Foi nesse sentido que, no segundo semestre de 2017, a NOS e a

Vodafone celebraram uma parceria estratégica com o objetivo de partilhar a rede fibra em mais de cerca de 2.6 milhões de casas, tendo sido partilhadas 530 mil casas em 2018.

No final de 2018, a cobertura da NOS chegava a cerca de 4.4 milhões de casas, das quais mais de 1 milhão de casas são fornecidas pela tecnologia FTTH. A Figura 4 mostra a expansão da cobertura nacional nos últimos três anos e apresenta um crescente número de lares com cobertura FTTH.

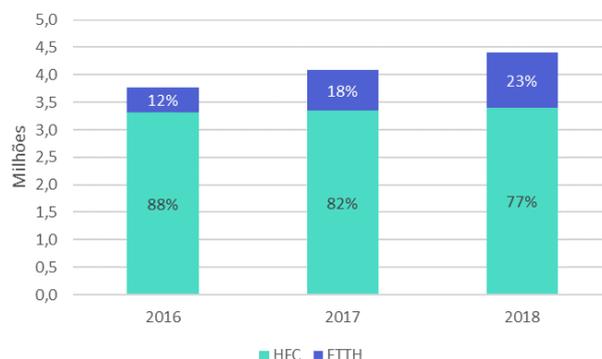


Figura 4 - Expansão da rede de FTTH in Relatório e Contas 2018 - NOS

A empresa pretende continuar a implementação de FTTH em novas zonas de tal forma que em 2022 se espere uma taxa de penetração de 70% de FTTH na rede fixa da empresa, com uma cobertura de cerca de 3,2 milhões de lares. Assim, é expectável que a abertura de novas células seja cada vez mais frequente, reforçando a necessidade de a previsão de vendas brutas seguir uma abordagem distinta das restantes tecnologias, em que o momento da abertura e a respetiva cobertura sejam tidas em consideração. Por célula entenda-se uma unidade geográfica à qual estão associadas várias unidades de alojamento (UA).

Assim sendo, a rede de FTTH da NOS pode-se desagrupar em três grupos:

- Fibra DST, proveniente de um grossista português de fibra ótica – DSTelecom.
- Fibra NOS, detida pela própria empresa.
- Fibra VDF, que é detida pela Vodafone e partilhada com a NOS no âmbito da parceria estratégica de partilha de fibra.



Figura 5 - Logótipos das empresas detentoras de rede fibra

Para a rede fibra NOS e VDF, é possível ainda distinguir dois tipos. A fibra NOS do tipo 1 é partilhada com a Vodafone no momento da abertura das células, sendo explorada por ambas as operadoras em simultâneo. Por outro lado, a fibra NOS tipo 2 já foi previamente explorada pela empresa antes de ser realizada a partilha.

A fibra VDF do tipo 1 é detida pela subsidiária do grupo britânica e, de igual forma, partilhada com a NOS no momento da abertura da mesma, enquanto que a fibra do tipo 2 já foi inaugurada no passado pela Vodafone, sendo agora partilhada com a NOS.

Canais de venda

As vendas podem ser efetuadas através de diferentes canais, que se podem dividir em reativos, através dos quais a empresa reage a um contacto efetuado pelo consumidor e proativos, através dos quais a operadora tenta alcançar potenciais novos clientes.

Dos canais reativos podem-se distinguir:

- Retalho, onde se destacam as lojas próprias e franchisadas, as grandes superfícies comerciais como *Worten*, *Fnac* e *The Phone House* e as pequenas superfícies comerciais como quiosques e papelarias.
- Telemarketing *inbound*, através do qual os interessados em aderir a um serviço, contactam a operadora pelo serviço de apoio ao cliente.
- Web e loja online.

Quanto aos canais proativos destacam-se:

- Telemarketing *outbound*, a partir do qual a empresa aborda proativamente os consumidores com o objetivo de vender os seus serviços.
- Vendas diretas, também conhecidas, como vendas porta a porta.

As vendas apresentam, naturalmente, comportamentos diferentes entre os vários canais. O canal *outbound* merece uma especial atenção pelo facto de estar sujeito a campanhas comerciais que ocorrem com o objetivo de aumentar o número de vendas. Por essa razão, é comum que neste canal surjam variações bruscas no nível de vendas brutas, devendo assim ser tratado de uma forma diferente dos restantes.

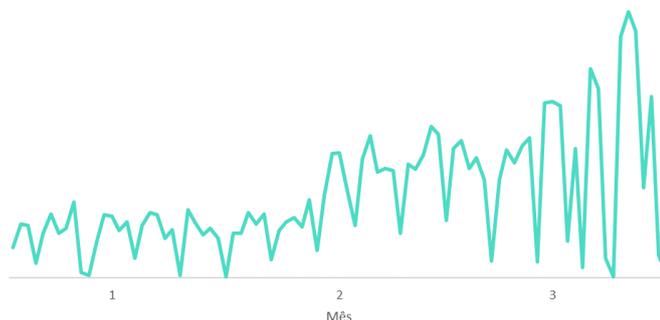


Figura 6 - Série temporal - Campanha comercial

Na Figura 6, encontram-se representadas as vendas por *outbound* num espaço temporal de três meses. É possível identificar-se uma campanha comercial que teve como objetivo a angariação de um maior número de clientes TV. A empresa abordou proativamente estes clientes, apresentando propostas de serviços associados à tecnologia DTH. Como é natural, a campanha teve influência direta nas vendas brutas desta tecnologia, que se traduziu num aumento considerável do nível de vendas efetuadas por *outbound*. Deste modo, a previsão de vendas através deste canal necessita de uma maior reatividade, pelo que, assim que seja notado o efeito da campanha nas vendas, o modelo deve-se adaptar e reagir de forma rápida. Ao mesmo tempo, é de igual forma importante a recolha de informação do departamento comercial no sentido de tomar conhecimento da campanha antecipadamente bem como a duração prevista de forma a aumentar a precisão da previsão.

3.1.2 *Gross-Adds*

Como referido anteriormente, apenas uma parte das vendas brutas se convertem em adições brutas. Assim sendo, verifica-se que as taxas de conversão de uma venda bruta num *gross-add* variam, consideravelmente, de acordo com a tecnologia da célula a que está associada a venda.

Pela Figura 7, verifica-se que a tecnologia 1 é a que apresenta, em média, uma taxa de conversão mais elevada sendo também aquela que contribui com um maior número de novos clientes para a empresa. Do mesmo modo, repara-se que quanto maior o peso que a tecnologia assume no total dos *gross-adds*, maior a taxa de conversão associada à mesma.

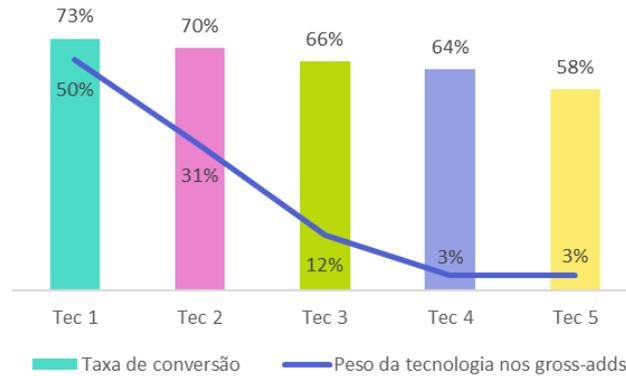


Figura 7 - Taxas de conversão de venda bruta em *gross-add* por tecnologia

Uma vez mais, é reforçada a necessidade de desagrupar a previsão de *gross-adds* por tecnologia.

Tempo entre venda e *gross-add*

O tempo que decorre desde que se efetua a venda bruta até ao momento em que esta se torna um *gross-add* é uma variável de enorme importância para a previsão, porque permitirá identificar o dia em que o *gross-add* será registado, ou seja, o dia em que um novo cliente entra para o parque da empresa. Este tempo compreende o intervalo entre a venda bruta e o agendamento da ativação, e o tempo entre o agendamento e a própria ativação.

Na Figura 8, podem-se reparar nas diferenças que existem entre as diferentes tecnologias, alertando para o facto de os seus valores não serem reais, mas serem proporcionais aos mesmos (encontram-se em base 1).

Assim, verifica-se novamente que este tempo apresenta também diferenças acentuadas de acordo com a tecnologia. As vendas da tecnologia 2 são as que se convertem num menor espaço de tempo em *gross-adds*, enquanto que as vendas da tecnologia 3 demoram mais do dobro do tempo em relação a esta tecnologia.

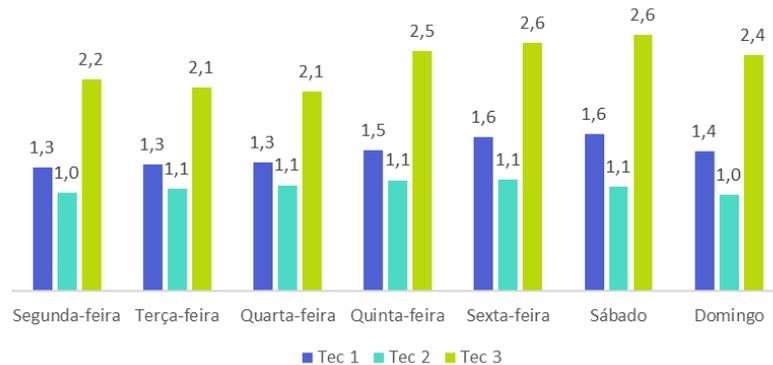


Figura 8 - Tempos entre venda bruta e *gross-add* por dia da semana e tecnologia

Para cada dia da semana em que se efetua a venda bruta, o tempo é também diferente. De uma forma geral, à medida que se aproxima o fim de semana, o valor desta variável vai aumentando. Assim sendo, a sexta-feira e o sábado são os dias que apresentam o maior tempo entre a venda bruta e o *gross-add* pelo facto de grande parte das ativações das vendas correspondentes apenas se iniciarem na semana subsequente.

Por outro lado, o domingo e a segunda-feira, apresentam os tempos mais baixos uma vez que a maioria das ativações se efetuam na própria semana, não sofrendo o impacto negativo do fim de semana.

Relativamente aos diferentes tipos de fibra, verificam-se também diferenças significativas que se encontram representadas na Figura 9.

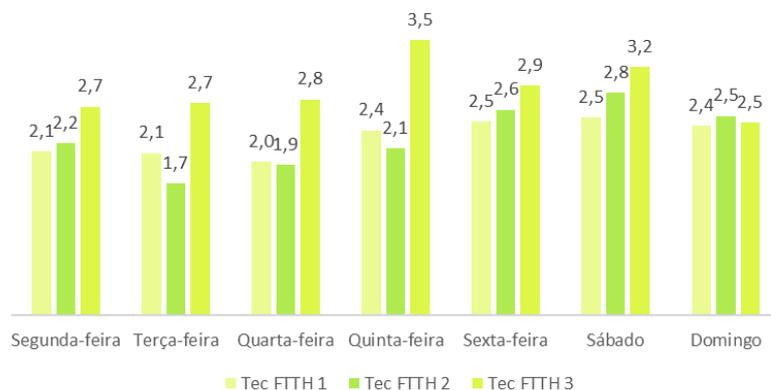


Figura 9 - Tempos entre venda bruta e gross-add por dia da semana e tipo de FTTH

Para além do efeito que o dia da semana apresenta no momento em que uma venda se torna um gross-add, verifica-se ainda, para todas as tecnologias, que à medida que o mês decorre, o tempo entre a venda bruta e o gross-add vai diminuindo de forma gradual. Enquanto que na primeira metade do mês o tempo se mantém constante, na segunda parte do mês, à medida que se aproxima o final do mês esse tempo vai apresentando valores mais baixos. À semelhança do que acontece nas vendas realizadas também aqui se pode explicar por uma aceleração do número de ativações por dia na reta final do mês.

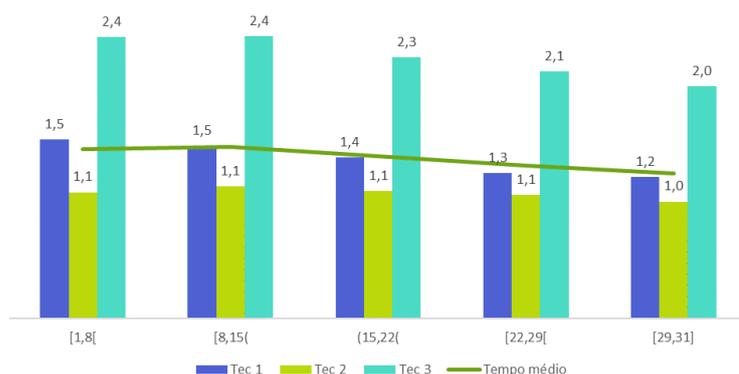


Figura 10 - Tempo entre venda bruta e gross-add por momento do mês

Em suma, há três fatores essenciais que necessitam de ser tidos em conta na estimativa do tempo que cada venda bruta toma até se tornar um gross-add: a tecnologia do serviço, o dia da semana em que ocorre a venda bruta e o respetivo momento do mês.

Dias especiais

Ao longo do ano existe uma série de dias especiais que têm impacto no processo de vendas. Estes são, principalmente, feriados nacionais ou municipais, ou dias especiais para a empresa, que influenciam os valores gross-adds quer pela quebra de vendas que provocam quer pela influência que têm no tempo que decorre entre uma venda bruta e um gross-add, face à menor capacidade da empresa em realizar ativações nestes dias.

O impacto que cada um desses dias deve ser avaliado de forma individual uma vez que apresentam impactos diferentes nos *gross-adds*. O dia anterior e o dia seguinte devem também ser considerados nesta análise uma vez que podem apresentar um valor diferente do nível historicamente apresentado.

3.2 Funil de retenção

O funil de retenção abrange o processo que se inicia quando um cliente manifesta a sua intenção de cancelar os seus serviços - diz-se que o cliente entra em retenção - desencadeando uma série de processos internos que têm como objetivo manter o cliente em parque.

Existem três canais através do qual o cliente pode manifestar a sua intenção de cancelar a subscrição dos serviços contratados e assim, entrar em retenção:

- *Inbound*. Estas entradas em retenção referem-se às chamadas realizadas pelos clientes para o serviço de apoio ao cliente.
- Em loja. O cliente dirige-se a uma loja NOS.
- Carta. A entrada em retenção é feita pelo envio de uma carta registada pelo cliente.

Ao entrar em retenção, quer por *inbound* quer em loja, o cliente é abordado por uma primeira equipa de retenção que tenta perceber as razões da intenção do cliente em abandonar os serviços da empresa e conseqüentemente, irá fazer uma primeira proposta de refidelização. Caso a mesma não seja aceite pelo cliente, este será transferido para uma segunda equipa de retenção, detentora de uma maior autonomia, que aborda a entrada em retenção e apresenta uma segunda proposta de refidelização. Caso o cliente se encontre em loja, é colocado em contacto com a segunda equipa de retenção por via telefónica.

Na hipótese de o cliente ainda manter a sua decisão de cancelar os seus serviços, terá de formalizar o processo através do envio de uma carta. Após a receção da mesma por parte da empresa, é criada uma encomenda de desligamento.

Caso a entrada em retenção seja realizada pelo envio de carta, a empresa tenta contactar esses clientes por via telefónica com o mesmo objetivo de os refidelizar. No entanto, as entradas em retenção por carta apresentam uma dificuldade acrescida para a empresa em ser bem sucedida na retenção do cliente pelo facto de, em muitas das situações, ser já uma decisão definitiva do mesmo. Para além disso, é mais provável que estes clientes já possuam um novo serviço de uma operadora concorrente e dessa forma, tornam o contacto telefónico, por vezes, impossível.

A cada etapa do processo de retenção, o cliente pode ser considerado como não recuperado, caso não tenha aceite a proposta de refidelização. Se, por outro lado, a proposta for aceite, o cliente é tratado como recuperado, sendo que, é comum que a proposta implique, por exemplo, uma alteração do pacote subscrito pelo cliente ou a renovação de equipamentos.

Neste ponto, é importante referir que apenas no momento em que se dá o desligamento, o cliente é considerado como *churn*. O desligamento de serviços associados à tecnologia HFC necessita do recurso a ordens de trabalho, enquanto que o desligamento de tecnologias FTTH e DTH é realizado de forma remota. Tal implica que para estas tecnologias, em muitas situações, o desligamento seja feito na data de fim de faturação.

Aqui, é oportuno também mencionar que ao domingo não são efetuados quaisquer desligamentos com recurso a ordens de trabalho e ao sábado, a força de trabalho é mais reduzida em relação a um dia útil comum, sendo, por isso, o dia da semana necessariamente um fator a ter em conta.

3.2.1 Retenção por Canal

As entradas em retenção variam de acordo com o canal bem como as respetivas taxas de retenção. Como se pode observar no gráfico da Figura 11, o *inbound* é o canal que apresenta, em média, um maior número de entradas em retenção que se explica pela maior facilidade e comodidade em efetuar uma chamada para o serviço de apoio ao cliente do que se dirigir a uma loja ou enviar uma carta registada. Assim, também é compreensível que este seja o canal com a maior taxa de retenção.



Figura 11- Entradas em retenção por canal

Por oposição, o envio da carta registada é, como seria de prever, o que regista o menor número de entradas em retenção e as taxas de *churn* mais elevadas pelo facto de o cliente estar a pagar o envio da mesma e provavelmente ter já a decisão definitiva.

É importante também não esquecer que as idas à retenção por canal são influenciadas pelo dia da semana, visto que os consumidores apresentam rotinas diferentes de acordo com o dia da semana em questão. Assim, aos fins de semana, são mais os clientes que entram em loja por comparação com os restantes canais, apesar do total de entradas em retenção ser inferior para todos os canais. A distribuição do número médio de entradas em retenção por dia da semana e canal é apresentado no Anexo A.

É interessante também reparar que os fins de semana apresentam taxas de retenção ligeiramente superiores aos dias úteis, pelo que o dia da semana é um fator a ter em conta não só na estimativa do número de idas à retenção, mas também na estimativa das taxas de retenção.

3.2.2 Período de fidelização

O período de fidelização é um instrumento de enorme importância no mercado das telecomunicações. Este é definido pela ANACOM como uma condição contratual pela qual o cliente se compromete a não cancelar o contrato definido com a operadora, sob pena de ter de suportar encargos de acordo com o momento do período de fidelização em que se situa.

Assim, de forma a fidelizar os consumidores, as operadoras oferecem condições mais vantajosas, como descontos na mensalidade, equipamentos mais baratos, oferta do valor da instalação do serviço, oferta de canais *premium*, pacotes de chamadas gratuitas, entre outros.

Em agosto de 2016, as operadoras de telecomunicações passaram a ter a obrigatoriedade de oferecer a todos os utilizadores a possibilidade de celebrarem contratos sem período de fidelização, bem como contratos com 6 e 12 meses de fidelização, sendo a duração máxima permitida de 24 meses.

No entanto, atualmente, esses contratos implicam o pagamento de uma quantia referente à ativação dos serviços e uma mensalidade superior quando comparado com o mesmo serviço

com um período de fidelização de 24 meses. Assim, os contratos sem período de fidelização ou com um período de fidelização de 6 e 12 meses são muito pouco apelativos em termos económicos face às grandes vantagens que os contratos de 24 meses oferecem.

Em qualquer um dos períodos de fidelização escolhidos, caso o cliente pretenda cancelar o seu contrato antes do fim do contrato, terá de pagar uma quantia que deve ser proporcional à vantagem que lhe foi oferecida pela cláusula de fidelização, não podendo exceder esse valor.

Para os clientes com contrato de 24 meses, o valor dos encargos é considerável ao longo de praticamente todo o período de fidelização e por essa razão, o *churn* voluntário apresenta valores bastante reduzidos até aos últimos meses do período de fidelização. Nos últimos dois meses, esses valores aumentam de uma forma bastante acentuada, como se pode reparar na Figura 12, pelo facto de os custos que implica o cancelamento do contrato já serem mais baixos. Uma vez mais, os valores percentuais encontram-se representados em base 100. No eixo horizontal, estão representados os diferentes escalões do período de fidelização com intervalos de três meses com exceção dos três meses que antecedem e sucedem o final do período de fidelização devido às diferenças significativas que existem entre eles.

É também possível verificar-se que a taxa de não retidos mantém-se praticamente constante para os diferentes escalões sem período de fidelização, sendo consideravelmente superiores para a tecnologia DTH. É normal que clientes com serviços de DTH que terminem o seu período de fidelização tenham possibilidade de subscrever um serviço de melhor qualidade fornecido pelas operadoras concorrentes.

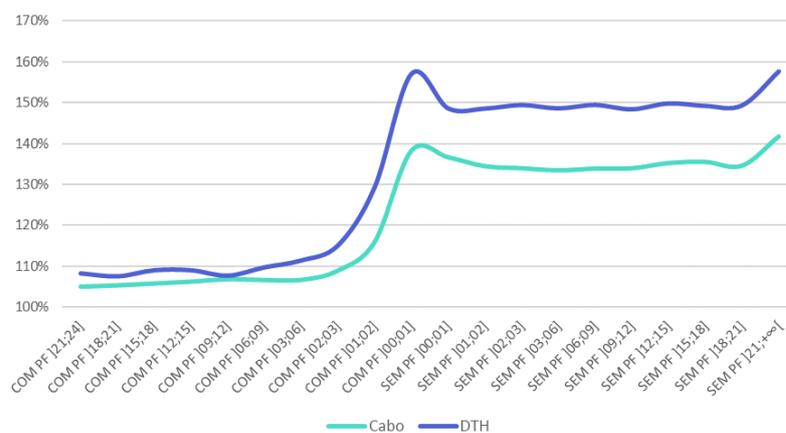


Figura 12 - Taxas de não retidos por escalão do período de fidelização

Ao longo do período de fidelização, a empresa apresenta diferentes abordagens para com os clientes de acordo com o momento do contrato em que se situam. No início do período de fidelização o objetivo da empresa é impulsionar as suas receitas, assumindo uma estratégia de *up-sell* e *cross-sell*. Aqui, é comum incentivarem-se os clientes a consumirem serviços e conteúdos adicionais bem como a adquirirem equipamentos acessórios.

A partir de determinado momento, o objetivo da empresa passa por refidelizar o cliente. Como Rosenberg e Czeipel (1984) referem, manter o cliente implica, a certo momento, oferecer mais do que aquilo que inicialmente o atraiu. Assim, à medida que o cliente se aproxima do final de contrato, a operadora vai assumindo uma estratégia cada mais agressiva de forma a não deixar o cliente escapar para uma operadora concorrente.

3.2.3 Série Temporal

Pela análise da série temporal do *churn* voluntário de um mês ilustrativo, representada na Figura 13, é desde logo bastante notório o elevado valor de *churn* no primeiro dia do mês que

se justifica pelos desligamentos que são realizados em clientes DTH nesse mesmo dia. O dia 1 de cada mês representa, assim, uma parte considerável do *churn* voluntário. É também de reparar que no último dia do mês, o *churn* voluntário para DTH volta a apresentar um valor superior em relação aos restantes dias do mês.

Quanto ao *churn* de clientes FTTH, é de notar que para além de ser inferior em relação às outras tecnologias apresenta uma irregularidade justificada pelo ciclo de faturação dos respetivos clientes. Aos dias 1, 8, 16 e 23 de cada mês, o *churn* para esta tecnologia apresenta valores superiores quando comparados com os restantes dias. Estes dias correspondem às datas de fecho do ciclo de faturação dos respetivos clientes.

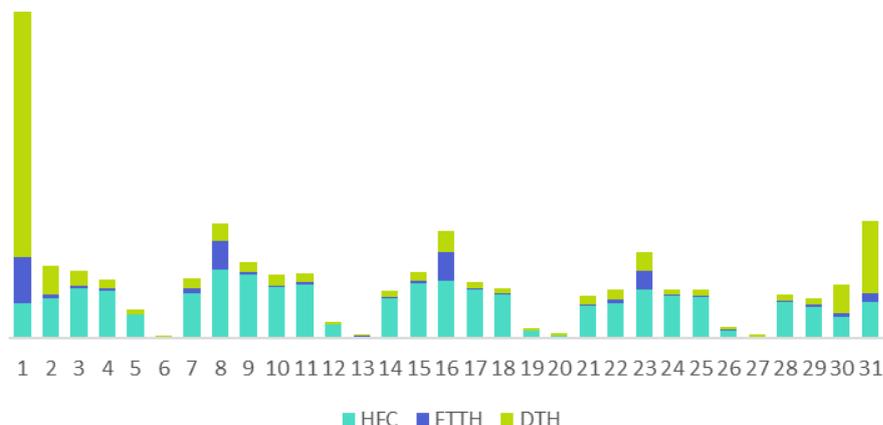


Figura 13 - Distribuição diária do *churn* voluntário mensal

Quanto ao *churn* para a tecnologia HFC e, uma vez que os desligamentos não são realizados remotamente, já se verifica uma distribuição mais semelhante entre os dias do mês com destaque para a sazonalidade semanal tal como acontecia nos *gross-adds*. Aos sábados, a força de trabalho para a realização de desligamentos é mais reduzida e ao domingo é dia de descanso, justificando, assim, os valores inferiores que se verificam nestes dias.

A distribuição diária do *churn* para o mesmo mês ilustrativo é apresentada, de forma individual para cada tecnologia, no Anexo B.

3.3 Churn involuntário

No que diz respeito ao processo de *churn* involuntário, este desenrola-se, naturalmente, de forma diferente do *churn* voluntário.

A régua de *dunning*, esquematizada na Figura 14, descreve todas as etapas e *timings* do processo. Este inicia-se quando existe um atraso no pagamento da fatura, que leva a que a empresa notifique o cliente da situação através de uma SMS. Caso o cliente não regularize a situação num prazo de 5 dias, será aplicado uma taxa de 1,50€ na fatura a pagar, sendo enviado um novo aviso nos dias que se seguem.

Ao fim de 35 dias após a data limite de pagamento (DLP), a empresa tenta contactar o cliente por chamada telefónica (via *outbound*) com o objetivo de perceber o motivo do atraso e normalizar a situação, sendo esta a primeira etapa que acarreta custos significativos para a empresa. Nesta fase do processo, é provável que surjam acordos de pagamento entre o cliente e a empresa, de forma a suprir dificuldades financeiras que o cliente possa ter.

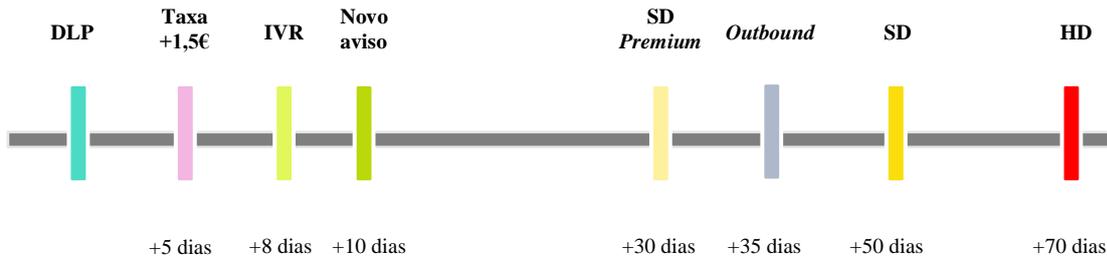


Figura 14 - Régua de *dunning*

No entanto, mantendo-se o incumprimento, a empresa procede à suspensão de alguns serviços digitais, os canais *premium* numa primeira fase e posteriormente, os serviços de internet. Ao fim de 70 dias após o atraso de pagamento, a empresa executa o *hard disconnect* (HD). Se se tratar de um serviço por satélite ou fibra, este é suspenso de imediato e os clientes tornam-se *churners*. Caso se trate de um serviço com a tecnologia HFC, as encomendas de desligamento são efetuadas e apenas quando os técnicos se deslocam a casa do cliente é que este é considerado como *churn*.

Os acordos de pagamentos realizados entre a empresa e o cliente influenciam o processo que este último toma. Caso o cliente efetivamente cumpra o acordo de pagamento, o processo dá-se por concluído. No entanto, o cliente pode entrar num novo incumprimento se não efetuar o pagamento até à data acordada com a operadora e nesta situação, 16 dias depois da data limite de pagamento, a empresa procede ao *soft disconnect* e os mesmos 20 dias mais tarde, o *hard disconnect*. O processo pormenorizado encontra-se na Figura 15.

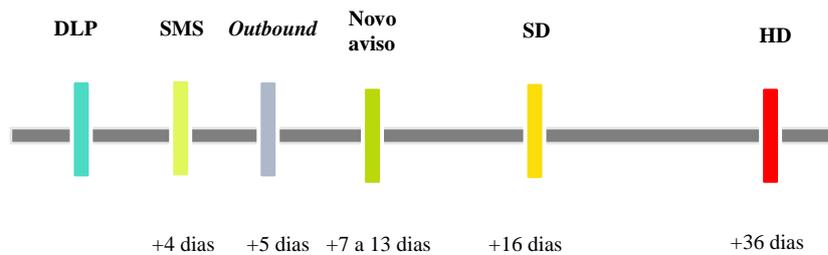


Figura 15 - Régua de *dunning* - Acordo de pagamento

A qualquer momento de ambos os processos, o cliente pode regularizar a situação e este termina automaticamente. Caso o cliente o faça após lhe ser aplicado o *hard disconnect*, há uma reversão no *churn* e o cliente continua em parque na empresa.

3.3.1 Série temporal

Pela análise da série temporal de um mês ilustrativo, representada na Figura 16, repara-se desde logo na existência de alguns picos para os valores de *churn* de clientes DTH. As datas de limite de pagamento já se encontram definidas e conseqüentemente, é possível identificar-se os momentos de cada uma das restantes etapas do processo de *churn* involuntário. Como a cada ciclo de faturação está associada uma determinada data de limite de pagamento, cada pico que se verifica no *churn* involuntário corresponde aos clientes de um dado ciclo de faturação.

Cada cliente de TV por subscrição tem um ciclo de faturação associado, existindo um total de nove ciclos de faturação. A cada um deles está associada uma data de início e fim de faturação e conseqüentemente, uma data a partir do qual o cliente se encontra em incumprimento, a DLP. Do total de ciclos de faturação, alguns deles partilham a mesma data

de fecho, pelo que há assim, cinco datas de fecho de ciclo diferentes que, portanto, explicam os cinco picos registados para DTH.

O mesmo não se verifica para a tecnologia HFC pelo facto de, após a execução do *hard disconnect*, ainda ser necessária a deslocação de um técnico a casa do cliente para concluir o cancelamento dos serviços e apenas aí, tal como no *churn* voluntário, o cliente é considerado *churn*. Este tempo entre o HD e o *churn* necessita de ser estimado no modelo de previsão.

Para serviços de fibra, o processo é o mesmo que ocorre para o DTH, visto que o desligamento é realizado exatamente 70 dias após a data limite de pagamento. Apesar de não ser possível apresentar os valores de cabo repartidos por FTTH e HFC, os picos que se registam para DTH também se verificam para FTTH. Ainda assim, é visível um ligeiro aumento nos valores de *churn* involuntário de cabo nesses mesmos dias.

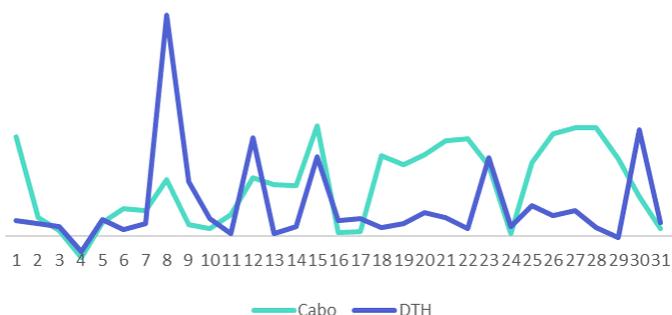


Figura 16 - Distribuição diária do churn involuntário mensal

Nota-se também, especialmente, na segunda metade do mês o efeito da sazonalidade semanal pelo facto de não serem efetuados desligamentos ao domingo.

É importante também referir que no dia 4 do mês em questão, o *churn* apresenta valores negativos que se justificam pelas reversões que ocorrem no processo de *churn* involuntário. Um cliente mesmo após ser considerado como *churn*, ou seja, após ser executado o HD para os serviços de DTH e FTTH ou após o técnico efetuar o desligamento em casa do cliente para HFC, ainda pode efetuar o *pagamento* da(s) fatura(s) em falta até se dar o *final disconnect*, um processo administrativo a partir do qual é impossível reverter o processo.

4 Modelo preditivo da evolução da base de clientes

O presente capítulo tem como objetivo apresentar detalhadamente o modelo preditivo desenvolvido.

O primeiro subcapítulo é dedicado à previsão diária do número de *gross-adds* e à forma como o modelo foi construído.

Numa segunda fase, é abordado o *churn* voluntário e involuntário, detalhando os *triggers* preditivos que impactam estas variáveis e de igual forma, todas as etapas da construção de ambas as previsões.

Este projeto exigiu o processamento de grandes quantidades de informação, sendo esta controlada por diferentes áreas da empresa. Assim, foi necessário ao longo do projeto procurar e recolher a informação junto de diferentes utilizadores de diferentes departamentos da empresa.

Uma grande parte da informação apenas é acessível usando o SQL Server, sendo por isso, tratada e explorada no SQL Server e posteriormente exportada para o Excel. De uma forma geral, para toda a informação consultada é possível aceder a todo o histórico. No entanto, visto se tratar de um mercado em constante transformação foi usada essencialmente informação do ano em questão e do anterior.

O modelo de previsão está construído em Microsoft Excel, sendo a sua informação extraída e atualizada automaticamente do SQL Server.

4.1 Previsão dos *gross-adds*

A primeira tarefa para a realização da previsão dos *gross-adds* foi a previsão das vendas brutas. Tendo em conta as diferentes características e especificidades das três tecnologias vendidas pela empresa, foi efetuada uma previsão das vendas brutas para cada uma delas, sendo que, para HFC e DTH são realizadas duas previsões de acordo com os canais através dos quais as vendas se efetuam. De forma a se isolar as vendas resultantes de campanhas comerciais, é realizada uma previsão para as vendas brutas por *outbound* e uma outra para os restantes canais. Estas previsões são realizadas com recurso ao método de previsão *Holt Winters*.

No que diz respeito à tecnologia FTTH, são realizadas diferentes previsões de acordo com a antiguidade da célula. Tratando-se de uma célula antiga, as vendas a esta associada já se encontram mais estabilizadas e por essa razão, é plausível aplicar-se novamente o *Holt Winters*. No entanto, se se tratar de células com uma baixa antiguidade, as vendas apresentam uma maior irregularidade e por isso, são estimadas com base na conversão do número de lares associados a uma célula em vendas brutas e de acordo com o número de semanas que já decorreram após a abertura da célula. As diferentes previsões de vendas brutas encontram-se sumarizadas na Figura 17.

O método *Holt Winters* utilizado em todas as previsões foi o multiplicativo pelo facto de a sazonalidade ser do tipo multiplicativa, dado que se a série apresentar alguma tendência a componente sazonal irá variar de acordo com a mesma.

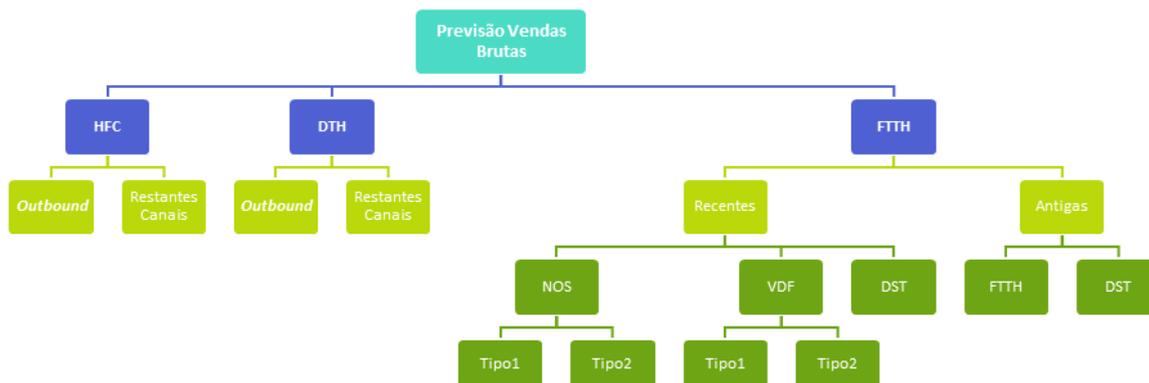


Figura 17 - Esquema da previsão de vendas brutas

O objetivo do projeto é ter-se o modelo com o máximo de automatismos possível. Por essa razão, no início de cada mês, com recurso à ferramenta Solver do Excel, são calculados os parâmetros alfa, beta e gama pela minimização do erro quadrático médio das observações do mês anterior. Com o decorrer do mês, os parâmetros devem ser monitorizados e alterados caso o decisor assim o entenda.

Os valores iniciais atribuídos seguem a proposta de Makridakis (1998) apresentada no segundo capítulo.

De forma a obter-se a previsão do mês em questão, é realizada uma previsão num horizonte de 7 dias numa primeira fase. Em seguida, os valores previstos para os mesmos 7 dias são assumidos como os valores reais e incorporados na previsão dos próximos 7 dias, e assim sucessivamente até a previsão do mês se encontrar completa. Os valores reais são introduzidos no modelo diariamente, permitindo que a previsão se adapte de acordo com as vendas do próprio mês.

4.1.1 Vendas brutas HFC

Para a tecnologia HFC e excluindo o canal *outbound* é visível, pela análise da série temporal, que as vendas apresentam alguma estabilidade ao longo do tempo. Na Figura e Figura , são apresentados os valores de vendas brutas de dois meses e, com base nessa informação, é realizada a previsão para o mês seguinte.

No que diz respeito ao canal *outbound*, pode-se afirmar que, mesmo havendo uma maior irregularidade do que nas vendas dos restantes canais agregados, não se registou para esta tecnologia qualquer campanha comercial que impactasse o respetivo nível de vendas.

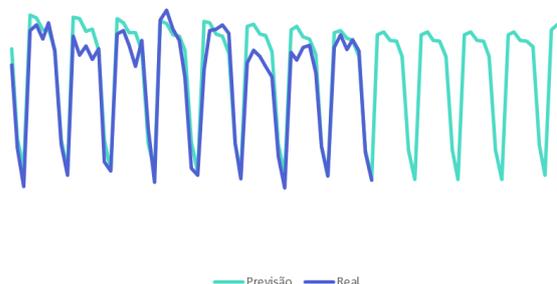


Figura 18 - Vendas brutas HFC – Sem *outbound*

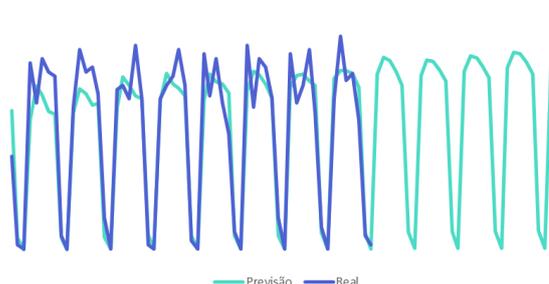


Figura 19 - Vendas brutas HFC – *Outbound*

Tratando-se de um método baseado exclusivamente em informação histórica, a componente subjetiva assume um papel importante pelo que a troca de impressões com o departamento comercial deve ser realizada de forma a se introduzir pequenos ajustes às componentes se assim for necessário.

4.1.2 Vendas brutas DTH

A previsão das vendas brutas para a tecnologia DTH seguiu a mesma abordagem das vendas brutas de HFC. Em relação ao canal de *outbound* e contrariamente ao que se verificava na tecnologia HFC, é visível a presença de uma campanha que provocou um aumento notório no nível de vendas realizadas. Assim, sempre que tal aconteça deve ser recolhida informação junto do departamento comercial, percebendo-se qual o alcance da campanha em curso e o tempo estimado para a sua realização. Dessa forma, podem ser feitos os ajustes necessários às componentes de forma a obedecer às expectativas da campanha.

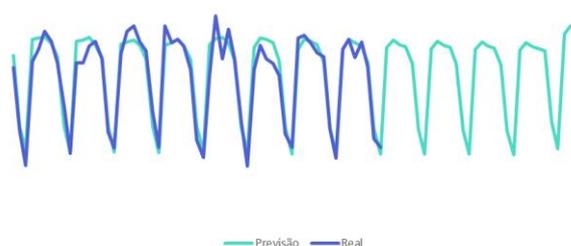


Figura 20 – Vendas brutas DTH – Sem *outbound*

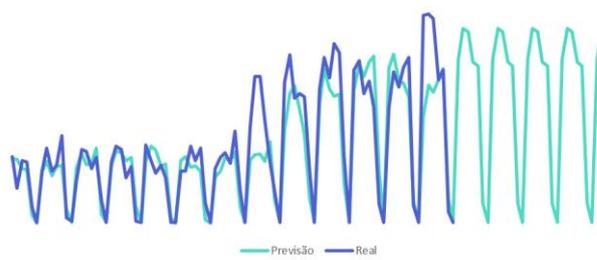


Figura 21 - Vendas brutas DTH - *Outbound*

De forma a se ter uma visão mais clara acerca do modelo, no Anexo C, é apresentado como a previsão de vendas DTH sem *outbound* se encontra construída.

Pela análise dos parâmetros e tendo sido obtido $\beta=1$, pode-se desde logo concluir que a componente da tendência para cada dia será calculada pela diferença entre os valores do nível da observação do próprio dia e do anterior, não sendo incluída neste cálculo a componente da tendência do dia anterior. Por outro lado, os baixos valores dos parâmetros $\alpha=0.0004$ e $\gamma=0.0697$ indicam que a informação mais recente apresentará um menor peso no cálculo das restantes componentes. Assim, no cálculo da componente do nível, a soma do valor da mesma componente no dia anterior com a componente da tendência assume uma maior importância. De igual forma, no cálculo da componente sazonal, o valor da mesma componente do homólogo semanal anterior acarreta uma elevada preponderância.

4.1.3 Vendas brutas FTTH

Células recentes

Por células recentes consideram-se aquelas cujas aberturas foram realizadas há menos de 36 semanas, inclusive. As instalações de fibra em novas zonas são planeadas pelo departamento comercial, sendo por isso, possível identificar-se com a antecedência necessária as datas de abertura de células previstas bem como o número de casas associadas à mesma.

Do número de UAs que cada célula cobre, apenas uma parte, naturalmente, irá ser uma venda bruta para a empresa. Esta conversão do número de lares em vendas brutas varia conforme o número de semanas que já decorreram desde o momento de abertura da mesma, sendo a própria semana de abertura aquela em que se regista uma taxa de captação superior, apresentando valores sucessivamente inferiores a partir de então.

A pré-venda, ou seja, a venda dos serviços antes das aberturas das células assume aqui um papel importante na estratégia comercial da empresa. Por norma, inicia-se um mês antes da abertura prevista da célula, no entanto, as vendas apenas se tornam significativas 3 dias antes da abertura.

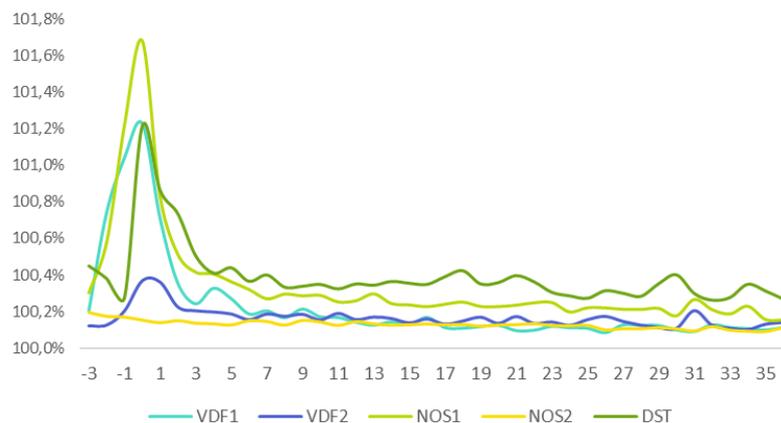


Figura 22 - Taxas de captação de FTTH

O tipo de célula a que está associada a venda é mais uma variável que, como pode ser observado na Figura, influencia as taxas de captação, pelo que se torna necessário realizar uma previsão de vendas para os diferentes tipos de FTTH. É importante referir que os valores do eixo vertical encontram-se em base 100 por questões de confidencialidade.

A semana de abertura da célula é a que apresenta uma maior taxa de captação, sendo as células do tipo NOS1 as que registam a maior taxa de captação na mesma. Nas semanas que se seguem, a taxa desce de forma acentuada mantendo-se depois relativamente estável ao longo do tempo para todos os tipos de célula.

As células que já foram previamente exploradas pela Vodafone (VDF2) ou pela NOS (NOS2) apresentam, como seria de esperar, taxas de captação mais baixas do que as células que estão a ser pela primeira vez exploradas por ambas as empresas.

Assim, de acordo com esta estimativa das taxas de conversão de lares em vendas brutas, pelos valores médios de cada semana, e tendo-se a informação relativa à data de previsão de abertura das células e o respetivo número de lares, é possível realizar-se uma previsão semanal do número de vendas brutas para cada tipo de célula, multiplicando-se o número de UAs em cada semana pela respetiva taxa de captação.

$$\text{Vendas FTTH} = \text{Taxa conversão UAs em Vendas} \times \text{Nr. UAs}$$

De forma a obter-se uma previsão diária do número de vendas realizadas, o valor da previsão semanal é depois distribuído pelos dias da semana, de acordo com o peso que cada dia da semana apresentou no mês anterior ao que se está a realizar a previsão.

Aqui é importante referir que as taxas de captação são atualizadas de forma automática. À medida que as unidades de alojamento se vão convertendo em vendas brutas, as taxas de captação vão-se atualizando considerando toda a informação histórica disponível até à data.

A informação quanto à data prevista de abertura de novas células e o respetivo número de lares associados é exportado do SQL Server e comparada com os agendamentos previstos pelo departamento comercial.

Células antigas

As células cuja abertura foi realizada há mais de 36 semanas são consideradas como células antigas. Estas apresentam uma maior estabilidade nos valores de vendas brutas em relação às

células mais recentes visto já não terem uma influência significativa das vendas diretas e do canal *outbound*. Desta forma, sendo possível identificar-se sazonalidade e tendência, recorreu-se, uma vez mais, ao método de previsão *Holt Winters*.

Verificou-se também que seria oportuno agregar os diferentes tipos de FTTH, optando-se por desagregar apenas a fibra DST.

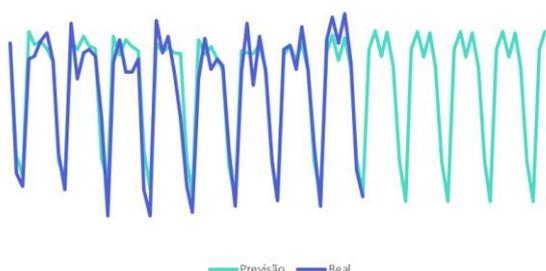


Figura 23 - Vendas brutas FTTH - Células antigas

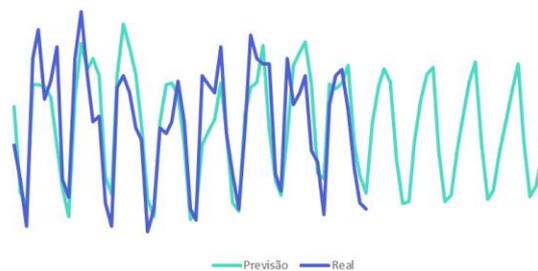


Figura 24 - Vendas brutas DST - Células antigas

As vendas da fibra DST são as que apresentam uma maior irregularidade que se traduz numa maior dificuldade em obter uma boa *accuracy* quando se compara com as restantes previsões. No entanto, é importante referir que estas vendas representam uma pequena parte das vendas de fibra e por essa razão, uma previsão com um maior erro associado será atenuado por esse facto.

4.1.4 Feriados e dias especiais

No modelo, para todas as tecnologias, é possível parametrizar cada dia do mês, de duas formas diferentes. Ao colocar o dia na tabela da Figura 25, o valor correspondente ao dia é substituído pela respetiva previsão, obtida pelo *Holt Winters*. Esta tabela é especialmente útil para tratar os feriados ou dias especiais que já aconteceram e que servem de base ao *Holt Winters*.

DIAS ESPECIAIS	
Dia	Descrição
...	...
01/05/2018	Dia do Trabalhador
10/06/2018	Dia de Portugal
13/06/2018	Dia de Santo António
24/06/2018	Dia de São João
31/05/2018	Corpo de Deus
15/08/2018	Assunção Nossa Srª.
05/10/2018	Implantação da República
01/11/2018	Dia de Todos os Santos
01/12/2018	Restauração da Independência
08/12/2018	Nossa Srª. Da Conceição
24/12/2018	Véspera de Natal
25/12/2018	Natal
31/12/2018	Véspera de Ano Novo
01/01/2019	Ano Novo
04/03/2019	Segunda-feira de Carnaval
...	...

Figura 25 – Dias especiais

Por outro lado, na tabela da Figura 26, é possível parametrizar o valor das vendas para qualquer dia e tecnologia. Sempre que surjam problemas nos sistemas de informação da empresa ou picos que não se pretendam substituir pela previsão do dia, pode-se definir qual o

valor das vendas para os dias afetados, escolhendo a tecnologia pretendida. Os valores apresentados na tabela não são reais.

DIAS ESPECIAIS							
Dia	Descrição	FTTH>36	DST>36	HFC OUTB	HFC S/ OUTB	DTH OUTB	DTH S/ OUTB
...
06/04/2019	Problema X	54	17				
14/05/2019	Problema Y	70	29	50	329	79	168
17/05/2019	Pico A	63	27				
19/05/2019	Problema Z	46					
27/05/2019	Pico B			40			
...

Figura 26 - Dias especiais por tecnologia

Esta é uma tarefa essencial de forma a que nos valores históricos que estão a ser considerados no método *Holt Winters* não sejam encontrados *outliers* que afetem negativamente a precisão do modelo.

No que diz respeito aos dias especiais que ainda estejam para acontecer, sabe-se que eles influenciam os valores das vendas brutas e diretamente o número de *gross-adds*. Assim, no modelo existe uma tabela na qual se devem inserir todas as datas que tenham historicamente uma redução prevista no número de vendas ou então, qualquer dia que, por indicação do departamento comercial, irá ter um desempenho abaixo do previsto.

Relativamente aos feriados, a taxa de redução é calculada automaticamente pela redução que aconteceu no último feriado em dia útil face à média dos últimos três homólogos semanais. Esta taxa é estimada para a tecnologia DTH e cabo devido ao diferente impacto que esses dias apresentam nas vendas de ambos os serviços. As taxas são aplicadas às previsões de vendas resultantes do *Holt Winters*.

Em termos de *gross-adds*, tem-se uma tabela com exatamente a mesma funcionalidade, sendo que as taxas de redução são aplicadas diretamente aos respetivos valores. Uma vez mais, as taxas são meramente ilustrativas.

Data	Dia da Semana	Descrição	DTH	CABO
...
19/04/2019	sexta-feira	Sexta Feira Santa	-82%	-91%
21/04/2019	domingo	Páscoa	-85%	-78%
22/04/2019	segunda-feira	Segunda-feira de Páscoa	-24%	-19%
29/04/2019	segunda-feira	Aumento - Final do mês	120%	112%
30/04/2019	terça-feira	Aumento - Final do mês	132%	125%
25/04/2019	quinta-feira	Dia da Liberdade	-91%	-91%
01/05/2019	quarta-feira	Dia do Trabalhador	-82%	-96%
10/06/2019	segunda-feira	Dia de Portugal	-90%	-97%
...

Figura 27 - Dias especiais - Taxas de redução

Caso o feriado ocorra ao fim de semana, o valor das vendas e *gross-adds* não sofre historicamente uma redução significativa não devendo por isso, ser incluído na lista. Como exceções verificam-se o dia de Páscoa e de Natal, visto que, quer se realize num dia útil ou não, é sempre verificada uma queda nos valores das vendas e *gross-adds* ainda que, a redução seja diferente se se tratar de um dia útil ou fim de semana.

As tabelas da Figura 26 e Figura 27 podem também ser usadas no sentido de incrementar o número de vendas previsto, assim que tal seja expectável. Os últimos dias do mês apresentam, por vezes, um valor superior aos restantes dias do mês e, portanto, devem também aqui ser considerados, bem como a respetiva taxa prevista de aumento.

4.1.5 Taxas de conversão de vendas brutas em *gross-adds*

Como foi referido no capítulo anterior, apenas uma parte das vendas brutas realizadas se tornam novos clientes para a operadora. Assim, assume-se necessário realizar uma estimativa dessa taxa de conversão para cada tecnologia a que o serviço subscrito se encontra associado.

Para tal, foi considerada informação histórica desde o início do segundo semestre de 2018, sendo recolhida a informação quanto ao número de vendas brutas de cada tecnologia para cada dia e obtidos os valores de *gross-adds* a que essas mesmas vendas deram origem. Assim, e uma vez que as taxas se mantêm idênticas de mês para mês, consegue-se obter uma estimativa das taxas de conversão por tecnologia.

O modelo atualiza de forma automática as taxas de conversão sempre que as tabelas a partir das quais é realizada a estimativa sejam também atualizadas, sendo possível selecionar os meses ou semanas que servem de base para esta estimativa.

No final desta fase, é aplicada a taxa de conversão estimada à previsão diária das vendas brutas de forma a se obter o número de *gross-adds* correspondentes a cada dia de vendas.

4.1.6 Tempo entre venda e *gross-add*

Uma vez previsto o número de *gross-adds* resultantes das vendas realizadas, interessa agora estimar o tempo que decorre entre o dia em que se dá a venda e aquele em que esta se converte numa adição para a empresa.

Sendo a tecnologia, o dia da semana e o momento do mês em que a venda se realiza os principais fatores a ser considerados, foi construída uma matriz na qual se estima a percentagem de novos clientes que se tornam *gross-adds* em cada dia após aquele em que se realizou a venda.

A Figura 28 serve de ilustração e como se pode observar, uma grande parte dos *gross-adds* dão-se nos primeiros dias após a venda, estando esta distribuição sempre dependente do impacto do fim de semana, que atrasa o tempo até o cliente ser considerado efetivamente um *gross-add*.

Dia e Semana	Dias													
	0	1	2	3	4	5	6	7	8	9	10	11	12	...
domingo1	0%	19%	34%	12%	5%	6%	5%	0%	5%	2%	1%	1%	2%	...
domingo2	1%	28%	26%	10%	9%	4%	5%	0%	3%	1%	1%	1%	2%	...
domingo3	3%	27%	25%	14%	6%	5%	4%	0%	2%	3%	2%	1%	2%	...
domingo4	3%	32%	26%	12%	5%	5%	5%	0%	4%	1%	0%	0%	0%	...
segunda-feira1	4%	33%	26%	10%	6%	4%	0%	3%	2%	1%	2%	1%	1%	...
segunda-feira2	6%	38%	20%	8%	6%	4%	0%	4%	1%	1%	1%	1%	2%	...
segunda-feira3	5%	40%	20%	11%	5%	4%	0%	4%	2%	2%	1%	2%	1%	...
segunda-feira4	7%	41%	20%	14%	6%	3%	0%	2%	1%	0%	0%	0%	0%	...
terça-feira1	5%	41%	17%	8%	5%	0%	7%	3%	2%	1%	1%	1%	0%	...
terça-feira2	5%	40%	19%	11%	4%	0%	5%	2%	2%	1%	1%	1%	0%	...
terça-feira3	7%	38%	20%	10%	5%	0%	4%	3%	2%	3%	2%	1%	0%	...
terça-feira4	7%	38%	25%	13%	3%	0%	4%	1%	1%	1%	1%	1%	0%	...
quarta-feira1	4%	43%	21%	6%	0%	10%	3%	2%	1%	1%	1%	0%	1%	...
quarta-feira2	7%	39%	21%	3%	0%	9%	3%	2%	3%	2%	2%	0%	1%	...
quarta-feira3	6%	41%	22%	6%	0%	7%	3%	4%	3%	2%	2%	0%	1%	...
quarta-feira4	8%	42%	22%	5%	0%	7%	4%	3%	1%	1%	1%	0%	1%	...
quinta-feira1	6%	36%	8%	0%	0%	20%	8%	5%	3%	2%	2%	0%	2%	...
quinta-feira2	7%	43%	8%	0%	16%	6%	4%	2%	2%	2%	0%	3%	0%	...
quinta-feira3	8%	40%	10%	0%	16%	8%	3%	3%	3%	2%	0%	1%	1%	...
quinta-feira4	10%	39%	10%	0%	16%	6%	4%	3%	1%	1%	0%	2%	0%	...
sexta-feira1	3%	10%	0%	33%	21%	9%	4%	4%	3%	0%	2%	1%	1%	...
sexta-feira2	4%	12%	0%	45%	14%	6%	3%	2%	3%	0%	1%	1%	1%	...
sexta-feira3	8%	17%	0%	34%	15%	7%	3%	3%	3%	0%	2%	1%	2%	...
sexta-feira4	9%	21%	1%	33%	16%	7%	5%	2%	2%	0%	1%	1%	0%	...
sábado1	3%	0%	26%	31%	10%	6%	3%	5%	0%	4%	1%	1%	1%	...
sábado2	4%	1%	34%	26%	7%	4%	5%	5%	0%	3%	2%	1%	0%	...
sábado3	3%	0%	32%	21%	11%	6%	4%	6%	0%	3%	1%	2%	1%	...
sábado4	5%	0%	26%	27%	16%	8%	5%	5%	0%	1%	0%	0%	0%	...

Figura 28 - Distribuição dos *gross-adds* por dia da semana e semana do mês

Uma matriz idêntica é construída para cada tecnologia, sendo que para FTTH, são construídas três matrizes diferentes dadas as diferenças que se verificam entre a fibra detida pela empresa, a que é partilhada pela Vodafone e a que é comprada ao grossista DStelecom.

À semelhança da estimativa das taxas de conversão de vendas brutas em *gross-adds*, também os valores destas matrizes se atualizam automaticamente assim que nova informação seja acrescentada às tabelas que as alimentam. De igual forma também, os meses ou semanas que servem de calculo a esta estimativa são selecionáveis.

Após realizadas as previsões das vendas brutas, a estimativa das taxas de conversão das mesmas em *gross-adds* e do respetivo tempo, é possível ter-se uma previsão diária do número de *gross-adds*.

Na Figura 29, pode-se observar como é realizada, em termos práticos, a previsão final do número de *gross-adds* diários para cada tecnologia e tipo de FTTH, sendo os valores apresentados em base 100 por questões de confidencialidade. Ao número de vendas brutas previsto para cada dia, representado a verde claro, é aplicada a estimativa da taxa de conversão de venda bruta em *gross-add* para a respetiva tecnologia, seguido da distribuição destes pelos dias seguintes à venda. Assim, para cada linha consta o número de *gross-adds* a que cada dia de vendas deu origem, sendo a previsão do número de *gross-adds* de um certo dia dada pela soma de cada diagonal no sentido ascendente, exemplificado a negrito para o dia 3 de abril.

Dia	Data	Dia da Semana	Vendas brutas	Gross-Adds	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
...
23	23/03/2019	sábado4	100	146	4	0	59	58	32	17	3	5	0	2	1	0	1	1	...
24	24/03/2019	domingo4	69	6	1	24	50	27	11	11	10	0	4	3	0	0	1	4	...
25	25/03/2019	segunda-feira4	286	244	16	21	15	26	13	7	0	4	0	0	1	0	1	0	...
26	26/03/2019	terça-feira4	313	232	16	81	13	9	6	0	8	3	2	1	1	1	0	3	...
27	27/03/2019	quarta-feira4	287	232	17	99	46	3	0	14	8	6	2	2	0	0	3	1	...
28	28/03/2019	quinta-feira4	292	240	21	84	22	0	8	4	7	7	2	3	0	1	0	3	...
29	29/03/2019	sexta-feira4	278	230	19	45	1	78	34	4	4	4	4	0	3	1	0	0	...
30	30/03/2019	sábado4	131	137	5	0	56	56	37	16	3	4	0	2	1	0	1	0	...
31	31/03/2019	domingo4	94	5	2	30	53	25	11	12	9	0	3	3	0	0	1	4	...
1	01/04/2019	segunda-feira1	284	228	9	23	26	21	14	9	0	7	1	1	3	2	2	0	...
2	02/04/2019	terça-feira1	298	208	12	89	12	8	10	0	15	7	5	1	1	2	0	3	...
3	03/04/2019	quarta-feira1	273	215	8	95	44	4	0	20	7	5	3	1	1	0	2	2	...
4	04/04/2019	quinta-feira1	286	192	12	72	18	0	14	7	11	6	4	4	0	0	1	2	...
5	05/04/2019	sexta-feira1	274	174	6	20	0	73	43	6	3	9	7	0	5	3	1	1	...
6	06/04/2019	sábado1	149	100	3	1	54	62	21	12	2	5	0	8	2	3	2	0	...
7	07/04/2019	domingo1	90	1	0	21	69	26	11	14	11	0	4	4	3	1	4	1	...
8	08/04/2019	segunda-feira2	305	240	14	26	22	17	12	7	0	9	1	1	1	2	3	1	...
9	09/04/2019	terça-feira2	326	246	12	90	13	12	7	0	10	4	4	1	1	3	0	3	...
10	10/04/2019	quarta-feira2	304	211	17	93	47	2	0	17	6	5	6	3	2	0	2	2	...
11	11/04/2019	quinta-feira2	280	181	14	96	19	1	10	7	8	3	3	4	0	2	0	3	...
12	12/04/2019	sexta-feira2	268	217	9	24	1	107	30	4	4	4	7	0	3	3	1	1	...
13	13/04/2019	sábado2	146	108	4	1	69	59	17	9	3	6	0	5	4	3	1	1	...
14	14/04/2019	domingo2	88	5	1	30	51	20	19	11	11	0	3	2	3	2	3	4	...
15	15/04/2019	segunda-feira3	309	297	12	27	22	23	10	10	0	9	1	2	3	4	1	0	...
16	16/04/2019	terça-feira3	314	218	15	90	13	11	11	0	10	7	4	2	3	3	0	2	...
17	17/04/2019	quarta-feira3	301	201	12	95	51	4	0	15	6	8	7	4	1	0	1	2	...
...

Figura 29 - Cálculo do número de *gross-adds* diários

Após se seguir a mesma abordagem para cada tecnologia e tipo de FTTH, é feito uma agregação dos resultados de forma a se obter a previsão do total de *gross-adds* para cada dia do mês.

4.2 Previsão do *churn*

Encontrando-se concluída a parte do modelo que permite prever o número de clientes captados para a empresa para o mês em questão, é necessário agora permitir a previsão do movimento de clientes que se regista no sentido inverso e, que assim, abandonam a empresa.

Essa previsão insere-se num contexto diferente daquele em que é feito a previsão dos *gross-adds*. A informação encontra-se, neste processo, do lado da empresa e apesar da vontade de desistir de um serviço partir do próprio cliente, o tempo entre o momento em que o cliente comunica a sua intenção até que efetivamente se torna *churn* é considerável, permitindo, assim, a incorporação de informação real do processo. De igual forma, no *churn* involuntário, um cliente que abandone o parque de clientes também tem de percorrer uma série de etapas até se tornar *churn*.

4.2.1 Previsão do *churn* voluntário

No processo do *churn* voluntário, a retenção é a primeira fase pela qual, em regra, todos os clientes terão de passar, pelo que o primeiro passo para se realizar uma previsão do *churn* voluntário será efetuar uma estimativa do número de clientes que entra em retenção.

Mesmo sabendo que, no início de cada mês, uma parte considerável do *churn* que se vai registar no próprio mês já se encontra em retenção, mantém-se necessário realizar uma previsão das idas à retenção. Assim, ainda que o número de clientes que entram em retenção e se tornam *churners* na própria semana ou na semana seguinte seja bastante reduzido, este necessita naturalmente de ser estimado. Para além disso, trata-se de informação com valor acrescentado uma vez que permite perceber como pode variar o *churn* num horizonte mais alargado.

Idas à retenção

O principal fator que influencia significativamente as idas à retenção é o período de fidelização do cliente. Clientes que se aproximam do fim do seu período de fidelização têm uma maior tendência para entrar em retenção e conseqüentemente, manifestar a sua vontade de suspender os seus serviços.

Assim, foi realizada uma estimativa das idas à retenção recorrendo a uma regressão múltipla, na qual a variável dependente é o número de idas à retenção e as variáveis independentes são o parque de clientes nos vários momentos do período de fidelização, com intervalos de 3 meses durante o período de fidelização e após o fim do mesmo, com exceção dos 3 meses que antecedem e se seguem ao fim do contrato, em que o intervalo passa a ser de 1 mês. Os resultados das ANOVAs são apresentados no Anexo D.

Na regressão, foram utilizadas 58 observações, sendo que cada uma delas representa o número de idas à retenção numa dada semana. A cada observação corresponde o respetivo parque de clientes no mesmo intervalo do período de fidelização. Uma vez mais, as idas à retenção necessitam de ser estimadas para cada tecnologia, sendo por isso realizada uma regressão múltipla para cada uma delas.

Assim, inicialmente, foi considerado o máximo de escalões do período de fidelização possível, sendo deixados de parte aqueles que apresentam um menor número de entradas em retenção. Foram-se realizando sucessivas regressões, retirando-se as variáveis que não se verificaram estatisticamente significativas, apresentado um alto valor de prova no teste-*t*. O valor do R^2 não sofreu, portanto, uma redução significativa, por oposição ao R^2 ajustado, que apresentou um valor cada vez mais satisfatório.

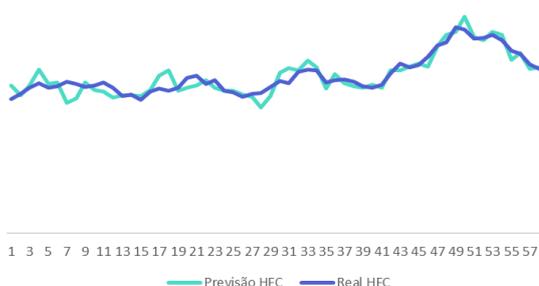


Figura 30 - Regressão das idas à retenção - HFC

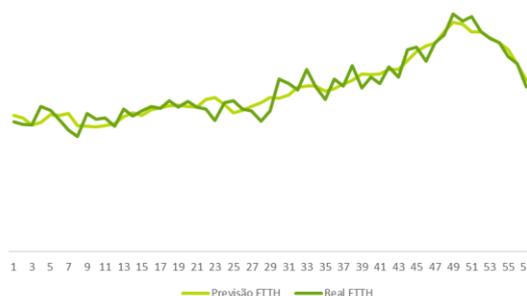


Figura 31 - Regressão das idas à retenção - FTTH

Como se pode observar na Figura 30, o número de idas à retenção previstas para a tecnologia HFC, ajusta-se bastante bem aos valores reais, resultando num $R^2 = 0,85$ e um R^2 ajustado = 0,83. A significância global da equação da regressão é validada por um valor de prova praticamente nulo.

As variáveis independentes que explicam as variações das entradas em retenção são os intervalos do período de fidelização que se encontram mais próximos do seu fim bem como os que se seguem ao fim do mesmo, confirmando a ideia de que estes são os clientes que mais entram em retenção.

No que diz respeito à tecnologia FTTH, o número de idas à retenção apresenta uma evolução temporal semelhante à registada para HFC. De igual forma, a regressão é capaz de explicar as variações das idas á retenção com base no parque de clientes, registando um $R^2 = 0,91$. De referir que as idas à retenção dos clientes fibra representa cerca de 10% a 15% do total de clientes com serviços cabo.

A tecnologia DTH, por outro lado, apresentou os resultados menos satisfatórios com um $R^2 = 0,64$, não tendo sido encontrada uma justificação consistente para explicar uma correlação mais fraca por comparação com as outras tecnologias.

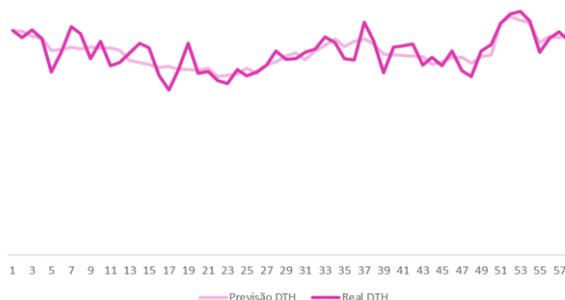


Figura 32 - Regressão das idas à retenção - DTH

O próximo passo consistiu em realizar uma previsão do número de clientes em parque em cada momento do seu período de fidelização no futuro. Sendo possível se obter informação quanto à data de início e fim do período de fidelização de cada cliente, é possível saber-se em que momento do contrato o cliente se encontrará em cada semana no futuro.

Não é, no entanto, possível quantificar o parque total de clientes no futuro pelo que se assume o pressuposto de que este é igual ao que se regista na última semana com tal informação disponível, não sendo por isso considerado o número de clientes que entra e sai do portfolio da empresa. Ainda assim, esse pressuposto apresenta pouca influência na previsão das idas à retenção, visto que o parque de clientes ronda os 1.4 milhões e os movimentos do parque em cada um dos sentidos rondam 1% desse valor.

Em seguida, com base no parque de clientes para cada escalão do período de fidelização e por tecnologia, multiplicam-se os coeficientes obtidos com recurso às regressões múltiplas, pelo parque correspondente aos escalões com influência estatisticamente significativa.

O número de entradas em retenção por semana é depois distribuído pelos dias da semana que a constituem bem como pelos três canais através dos quais se pode entrar em retenção, tendo por base o peso que cada dia da semana e que cada canal apresentam historicamente. Nesta estimativa é considerada a informação mais recente até ao dia em questão.

Uma vez mais, todo este processo encontra-se automatizado, sendo que com o decorrer do tempo, as previsões e estimativas vão sendo substituídas pelos valores reais de forma automática. Quanto à regressão múltipla é aconselhável que espontaneamente, por exemplo no início de cada mês, seja novamente executada de forma a manter uma boa precisão dos resultados.

Não Retidos

Após se ter uma estimativa diária do número de idas de retenção por canal e, sabendo que uma grande parte do *churn* que está a ser previsto já se encontra em retenção, é realizada uma previsão dos clientes que não irão ser retidos neste processo e assim, se tornarão *churners*.

Desde logo, a tecnologia do serviço do cliente influencia de forma notória a taxa de retenção do cliente, pelo que se torna obrigatório realizar uma previsão do número de clientes não retidos para cada tecnologia: HFC, DTH e FTTH.

Para além disso e uma vez mais, o período de fidelização é um fator determinante na retenção do cliente. Um cliente que já se encontra sem período de fidelização ou que se aproxima do fim de contrato apresenta uma maior probabilidade de vir a ser *churn* do que um cliente que acabou de contratar um serviço. O canal pelo qual o cliente entrou em retenção também influencia a taxa de não retidos, como se constatou no subcapítulo 3.2.1.

Deste modo, para cada escalão do período de fidelização, canal de entrada em retenção e dia da semana, é obtida uma estimativa da taxa de não retidos. Na Figura 33, é ilustrado uma dessas tabelas que diz respeito ao canal *inbound* para uma dada tecnologia, sendo que para as restantes, tem-se uma tabela idêntica com valores naturalmente diferentes. Os valores encontram-se, uma vez mais, mascarados em base 100.

INBOUND			
Escalão PF	Útil	Sábado	Domingo
COM PF]00;01]	122,0%	119,5%	119,8%
COM PF]01;02]	110,8%	114,1%	116,3%
COM PF]02;03]	104,7%	102,5%	105,8%
COM PF]03;06]	102,5%	100,8%	102,0%
COM PF]06;09]	101,8%	101,8%	103,0%
COM PF]09;12]	101,6%	102,8%	101,1%
COM PF]12;15]	101,3%	101,1%	101,8%
COM PF]15;18]	101,4%	101,4%	101,3%
COM PF]18;21]	101,2%	101,5%	101,0%
COM PF]21;24]	101,3%	101,4%	101,1%
SEM PF]00;01]	118,4%	112,5%	115,0%
SEM PF]01;02]	114,5%	111,5%	108,2%
SEM PF]02;03]	115,0%	110,9%	111,4%
SEM PF]03;06]	117,3%	108,8%	121,9%
SEM PF]06;09]	116,8%	115,1%	117,0%
SEM PF]09;12]	115,8%	116,5%	112,1%
SEM PF]12;15]	115,1%	113,9%	111,8%
SEM PF]15;18]	118,0%	124,3%	100,0%
SEM PF]18;21]	118,4%	115,9%	117,6%
SEM PF]21;+=[123,0%	116,1%	121,3%

Figura 33 - Taxas de não retidos por escalão do período de fidelização - *Inbound*

Todas as tabelas são atualizadas à medida que nova informação é incorporada nas suas fontes, sendo possível, uma vez mais, selecionar as semanas ou meses que servem de cálculo às mesmas de forma a que se adequem da melhor forma à realidade atual.

No final desta etapa, é obtida uma previsão do número de idas à retenção para cada dia e o respetivo número de clientes que não é retido pela empresa. O próximo passo é, por isso, estimar quando é que esses clientes irão ser considerados como *churn*, uma vez que o desligamento não é realizado no momento imediato ao que o cliente é considerado como não recuperado.

HFC

Quanto os clientes HFC, o desligamento implica o deslocamento de um técnico a casa do cliente. Desta forma, numa primeira fase, é realizada uma estimativa do tempo que decorre desde a entrada em retenção até ao *churn*, recorrendo-se ao histórico mais recente. Esta estimativa encontrar-se-á sempre atualizada, assumindo a nova informação que vai surgindo com o decorrer do tempo.

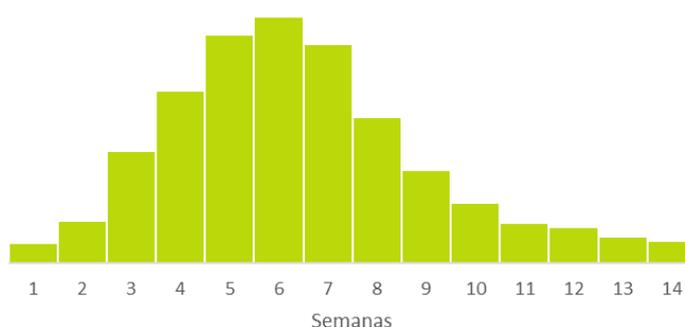


Figura 34 - Estimativa do tempo entre entrada em retenção e *churn* - HFC

Na Figura 34, pode ser observada a distribuição em termos percentuais desse mesmo tempo. Verifica-se que a maior parte dos clientes que entram em retenção tornam-se *churn* 3 a 8 semanas depois e que, nas primeiras duas semanas após a entrada, essa percentagem é bastante reduzida. Refira-se que estas percentagens foram estimadas com base no histórico desde o último trimestre do último ano até à data e que as mesmas se verificaram bastante semelhantes para cada semana considerada.

DTH e FTTH

Como se referiu no Capítulo 3, os clientes DTH e FTTH veem os seus serviços serem desligados de forma remota. Dessa forma, foi necessário perceber-se quais os critérios usados para definir o dia em que se dá o desligamento e o cliente ser, assim, considerado *churn*. Desta forma, verificou-se que o principal fator que determina o dia do desligamento é o ciclo de faturação do cliente. Uma considerável parte dos desligamentos efetuados para estes clientes coincide com o dia de fecho do ciclo de faturação.

Assim, atendendo à existência de vários picos que se registam durante o mês, torna-se extremamente difícil utilizar a mesma abordagem que a tecnologia HFC e distribuir um valor semanal pelo peso de cada dia da semana visto que este irá variar de semana para semana. Por esse motivo, é realizada uma previsão mensal do *churn*, sendo depois distribuída pelo peso que cada dia assume no mês.

Aqui, é importante referir que o peso que cada dia assume é muito semelhante entre os diferentes meses do ano. Assim, dependendo do número de dias que o mês contém, a estimativa do peso assume os valores reais do último mês com o mesmo número de dias. De

referir também que para feriados e dias especiais, é considerado que não apresentam qualquer impacto.

De igual forma, é estimada uma distribuição da percentagem de clientes que se tornam *churn* em cada mês após aquele em que entram em retenção, considerando-se a informação histórica desde o segundo semestre do ano anterior. As percentagens revelam-se relativamente constantes mês após mês pelo que se torna válido realizar-se uma estimativa deste modo.



Figura 35 - Estimativa do tempo entre entrada em retenção e *churn* - HFC e DTH

Churn sem retenção

Apesar de, em regra, todos os clientes que pretendam cancelar os seus serviços, terem obrigatoriamente de passar pelo processo de retenção, nem sempre tal acontece. Todos os meses existe uma pequena parte dos *churners* para os quais não é registada a sua entrada em retenção, não havendo uma razão consistente para explicar o sucedido.

Relativamente a estes clientes, foi efetuada uma análise com o objetivo de se identificar algum tipo de característica comum que ajudasse a explicar esse facto e permitisse, de alguma forma, prever o número desses clientes para cada mês. No entanto, tal não foi conseguido. Não se verifica qualquer dado comum entre estes, sendo por isso, difícil efetuar-se uma estimativa dos clientes que não vai à retenção e que se tornam *churn*.

Ainda assim, foi identificado que o peso dos clientes que não passa pelo processo é bastante reduzido e relativamente constante todos os meses, verificando-se também que a tecnologia tem influência na proporção destes clientes no total.

Por essa razão, aos valores de *churn* previstos, é aplicada uma taxa estimada para cada tecnologia com base no que se registou nos últimos meses.

4.2.2 Previsão do *churn* involuntário

O processo de *churn* encontra-se bem definido com os tempos entre cada fase delimitados. Por essa razão, a estratégia para a previsão do *churn* involuntário consiste em prever o número de clientes em cada fase do processo e as respetivas taxas de conversão.

Sendo o objetivo principal prever os valores de *churn* involuntário no final do mês, ou seja, num horizonte máximo de 31 dias, não será necessário prever todas as etapas do processo, visto que o tempo que decorre entre a data limite de pagamento e o *churn* é no mínimo 70 dias.

A previsão dividiu-se em dois ramos: um que diz respeito aos clientes que seguem a régua de *dunning* normal e outro que é relativo à régua de *dunning* que deriva dos acordos de pagamento realizados entre a empresa e o cliente.

Desta forma, a primeira tarefa para a previsão do *churn* involuntário sem acordos de pagamento consistiu na previsão do número de *soft disconnects* que é realizado com base no número de clientes que se encontram na fase anterior do processo, o *outbound*. Assim, é possível prever o *churn* num horizonte de 35 dias dado ser o período de tempo mínimo entre o

momento em que a empresa contacta o cliente para regularizar a dívida e aquele em que este se torna *churn*.

Aqui, devido às diferenças que se verificam entre os diferentes ciclos de faturação, é estimada uma taxa de conversão para cada um deles, sendo que ciclos de faturação com a mesma data de fecho são agregados num só, visto não se registarem discrepâncias relevantes entre eles.

Posteriormente, é realizada uma estimativa da taxa conversão entre a presente etapa e a seguinte, o *hard disconnect*, seguindo-se a mesma abordagem. Uma vez mais, verificou-se que existem diferenças significativas nos valores da taxa de conversão para os vários ciclos de faturação pelo que cada um deles terá associada uma taxa de conversão estimada com base na informação dos últimos três meses.

Por fim, de forma a obter-se os valores de *churn* previstos para clientes que não efetuam acordo de pagamento durante o processo, é realizada uma estimativa das taxas de conversão do número de clientes a que lhes é aplicado o *hard disconnect* e se tornam *churn*, para cada tecnologia.

Aqui, é relevante mencionar que com o decorrer do tempo, o número de clientes em cada fase do processo vai sendo sucessivamente atualizado. Na tabela que diz respeito ao *outbound*, apresentada na Figura 36, apenas se encontram valores reais, que servem de base para a previsão da etapa seguinte. Por outro lado, as tabelas relativas ao número de clientes aos quais se aplicam *soft disconnects* e/ou *hard disconnects* contêm os valores reais disponíveis até à data, sendo os restantes valores previstos. Isto permite que com o decorrer do mês, a previsão se torne cada vez mais precisa, dada a substituição dos valores previstos por valores reais. De referir que estes se encontram mascarados por questões de confidencialidade.

DLP	Outbound	DTH	FTTH	HFC	SD	DTH	FTTH	HFC	HD	DTH	FTTH	HFC
12/12/2018	16/01/2019	100	89	496	31/01/2019	59	46	226	20/02/2019	24	16	70
19/12/2018	23/01/2019	123	91	428	07/02/2019	67	41	183	27/02/2019	30	17	73
28/12/2018	01/02/2019	363	75	501	16/02/2019	154	30	202	08/03/2019	67	13	92
01/01/2019	05/02/2019	347	86	596	20/02/2019	134	31	211	12/03/2019	39	8	59
04/01/2019	08/02/2019	105	87	546	23/02/2019	62	35	225	15/03/2019	28	13	75
12/01/2019	16/02/2019	116	102	543	03/03/2019	64	48	228	23/03/2019	33	20	85
19/01/2019	23/02/2019	111	82	391	10/03/2019	63	39	167	30/03/2019	34	18	68
28/01/2019	04/03/2019	342	76	471	19/03/2019	165	34	215	08/04/2019	70	12	85
01/02/2019	08/03/2019	316	76	533	23/03/2019	122	27	180	12/04/2019	38	9	50
04/02/2019	11/03/2019	101	85	525	26/03/2019	67	40	250	15/04/2019	30	15	80
12/02/2019	19/03/2019	110	96	536	03/04/2019	57	39	200	23/04/2019	29	16	81
19/02/2019	26/03/2019	121	96	433	10/04/2019	67	43	200	30/04/2019	25	17	79
26/02/2019	02/04/2019	306	71	471	17/04/2019	148	30	209	07/05/2019	57	12	86
01/03/2019	05/04/2019	307	80	551	20/04/2019	127	28	183	10/05/2019	34	8	54
04/03/2019	08/04/2019	104	93	528	23/04/2019	66	47	254	13/05/2019	26	17	72
12/03/2019	16/04/2019	116	107	548	01/05/2019	65	45	220	21/05/2019	27	16	71
19/03/2019	23/04/2019	143	110	508	08/05/2019	64	35	154	28/05/2019	29	13	59
28/03/2019	02/05/2019	376	93	533	17/05/2019	141	29	191	06/06/2019	58	11	80
01/04/2019	06/05/2019	339	86	597	21/05/2019	121	26	185	10/06/2019	35	8	52
04/04/2019	09/05/2019	108	93	527	24/05/2019	54	36	204	13/06/2019	23	13	64
12/04/2019	17/05/2019	111	100	513	01/06/2019	62	45	211	21/06/2019	28	17	74
19/04/2019	24/05/2019	128	101	437	08/06/2019	67	43	177	28/06/2019	31	18	70

Figura 36 - Previsão do *churn* involuntário - Etapas intermédias

Quanto à previsão do número de clientes que são *churn* devido ao incumprimento do acordo de pagamento, a abordagem é semelhante. Contudo, em todos os dias do mês são realizados novos acordos de pagamento, sendo assim o dia acordado entre o cliente e a empresa para o pagamento da dívida, aquele que define os *timings* do novo processo.

Assim, a principal diferença da previsão do *churn* que deriva dos acordos de pagamento encontra-se nos *timings* do mesmo e nas taxas de conversão entres as várias fases que o constituem.

De igual forma, o modelo recolhe a informação mais atualizada do número clientes em *outbound*, e com base nessa informação e na estimativa das taxas de conversão para este processo, obtém-se uma previsão do número de clientes aos quais será aplicado o *soft*

disconnect. Com a mesma abordagem, em seguida, é realizada uma estimativa dos *hard disconnects* e posteriormente, do *churn*.

Por fim, obtém-se uma estimativa dos valores de *churn* para cada um dos ramos e por tecnologia. Para os clientes DTH e FTTH, sabe-se que o *churn* acontece no mesmo dia em que a empresa aplica o HD, sendo apenas necessário proceder à soma do número de clientes que foram *churn* sem efetuar qualquer acordo de pagamento com o número daqueles que o fizeram.

Por outro lado, relativamente à tecnologia HFC, é necessário ainda estimar o tempo que decorre desde o momento em que foi efetuado o *hard disconnect* até que se verifica o *churn*. A estimativa é apresentada na Figura 37.

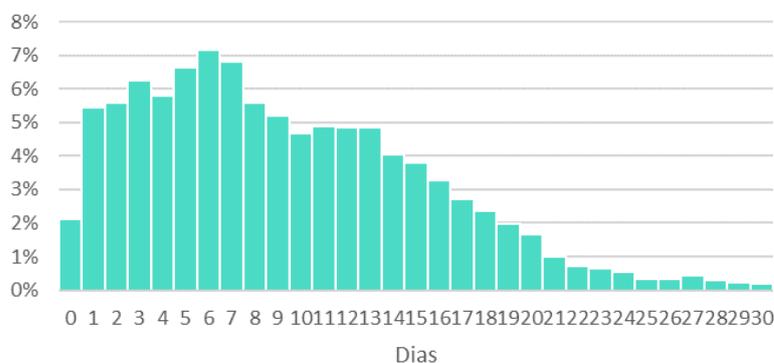


Figura 37 - Estimativa do tempo entre HD e churn - HFC

Uma vez finalizada a previsão do número de clientes que abandonam o portfolio da empresa e a previsão do número de novos clientes captados pela empresa, é possível estimar as adições líquidas, ou em inglês *net additions*, no início de cada mês. Neste cálculo, é usual incluírem-se também as migrações de clientes, ou seja, os movimentos que ocorrem dentro da base de clientes da empresa, em especial, do setor empresarial para o *consumer*, para os quais já existe uma previsão. Deste modo, tendo-se uma estimativa do número de adições líquidas é possível se antever como irá evoluir a base de clientes *Pay-TV* da empresa no curto-prazo.



Figura 38 - Cálculo de *net-adds*

5 Resultados

O presente capítulo serve para analisar os principais resultados do modelo preditivo. Sendo a previsão dos *gross-adds* a primeira etapa a estar completa, foi possível realizar-se uma previsão para os meses de abril e maio, sendo os resultados registados e monitorizados ao longo de ambos.

Quanto à previsão de *churn* voluntário e involuntário, o mês de maio foi usado como teste, sendo apenas tida em consideração a informação histórica até ao final do mês de abril. Durante o mês, o modelo tornou-se operacional permitindo, assim, uma estimativa dos valores do fecho.

5.1 Previsão dos *gross-adds*

É certo que o objetivo principal passa por prever com a maior precisão possível o número de *gross-adds* mensal, no entanto, garantir que a previsão de vendas apresenta pequenos desvios face ao real é o primeiro passo para se obter resultados satisfatórios na previsão final dos *gross-adds*. Assim, torna-se relevante fazer uma análise dos desvios das previsões de vendas brutas face aos valores reais.

Abril

À medida que se aproxima o final do mês é, naturalmente, expectável que a previsão se aproxime cada vez mais do valor real devido à incorporação de informação real no modelo. Desta forma, ao longo do mês de abril foram realizadas previsões em três momentos diferentes do mês, dia 1, 10 e 21, podendo-se assumir cada vez mais com uma maior confiança os resultados do fecho do mês.

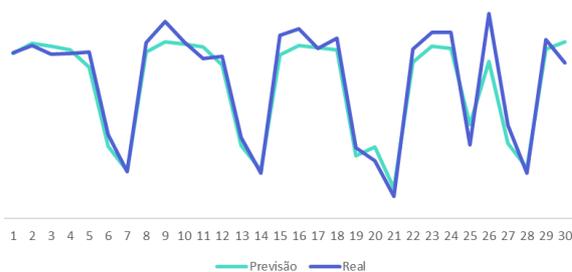


Figura 39 - Primeira previsão de vendas brutas

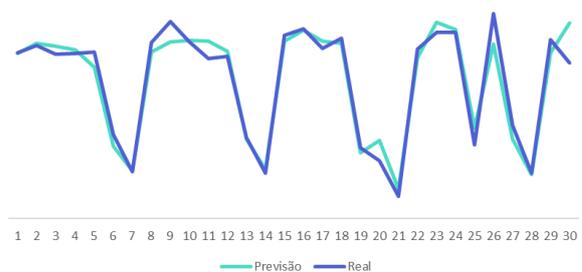


Figura 40 - Previsão de vendas brutas mais atualizada

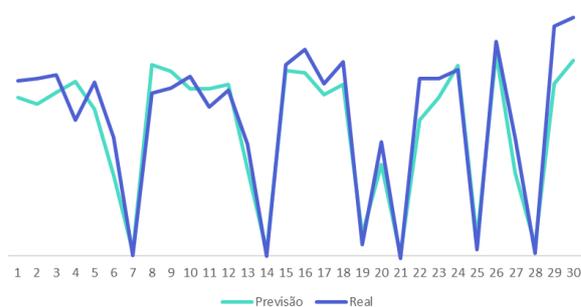
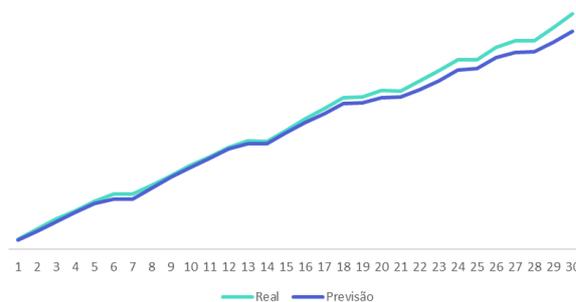
Observando-se, pelo gráfico da Figura 39, a comparação da previsão diária face aos valores reais, pode-se concluir que a previsão realizada no início do mês se encontrou abaixo dos valores reais. Este facto é, em boa parte, explicado pela campanha comercial que não estava a ser considerada na previsão (por falta de conhecimento), que teve início no curso do mês e levou a aumento do número de vendas diário de DTH.

Tabela 1 – MAPE – Previsão das vendas brutas de abril

	MAPE			Desvio - Total Vendas
	1-30	10-30	21-30	
01/04/2019	10,0%	11,9%	13,8%	-3,7%
10/04/2019	-	10,1%	11,2%	4,9%
21/04/2019	-	-	11,2%	0,1%

Contudo, na segunda e terceira previsões realizadas, a mesma campanha já é tida em consideração, resultando numa previsão de vendas que se veio a verificar mais ajustada à realidade, como se comprova na Figura 40.

No que respeita aos valores de *gross-adds*, na Figura 41 e Figura 42, pode-se observar a previsão efetuada no primeiro dia do mês de abril e a respetiva comparação com o real. Na primeira metade do mês, o modelo segue em linha com os valores reais, apresentando uma precisão de 97,5% até ao 15º dia. No entanto, na segunda metade do mês, os valores previstos afastam-se dos valores reais, sendo que o resultado final previsto do mês apresenta um desvio de cerca de 7% face ao real.

Figura 41 - Primeira previsão diária de *gross-adds*Figura 42 – Primeira previsão acumulada de *gross-adds*

O dia 22 de abril, segunda-feira que se sucede à Páscoa, é considerado um dia de especial na medida em que o histórico transmite uma redução no número de vendas e ativações. No entanto, tal apenas se verificou em termos de vendas. No que diz respeito aos *gross-adds* de serviços cabo, não se verificou uma diminuição face aos homólogos semanais, o que impactou negativamente a precisão do modelo.

Para além disso, no final do mês, especialmente nos últimos dois dias, verifica-se um aumento considerável do número de *gross-adds*, apesar de não ter sido registado qualquer aumento significativo nas vendas que pudesse levar a um maior número de novos clientes. Desta forma, a variação positiva pode ser justificada por um aumento das taxas de conversão de vendas brutas em *gross-adds* e uma aceleração do número de ativações por dia.

Por outro lado, se compararmos os resultados da previsão mais recente a cada dia do mês com os valores reais, na Figura 43 e Figura 44, pode-se reparar que a *accuracy* do modelo aumenta consideravelmente. A segunda previsão permitiu que os resultados da previsão e do real se encontrassem bastante próximos, com um desvio percentual de 2,7%.

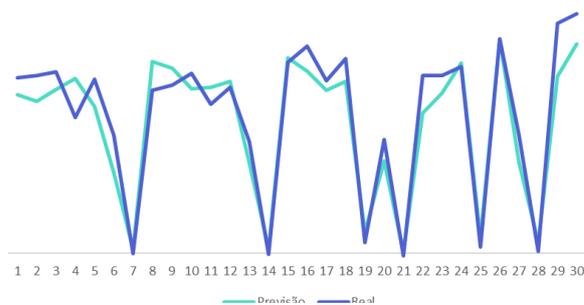


Figura 43 - Previsão diária mais atualizada de *gross-adds*

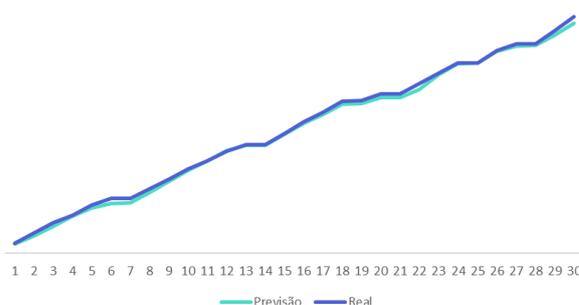


Figura 44 – Previsão acumulada mais atualizada de *gross-adds*

A última previsão do mês foi realizada no dia 21 de abril e manteve-se em linha com o real até os últimos dois dias do mês de abril. O modelo tinha em conta uma ligeira subida do nível de *gross-adds* após o feriado de dia 25, no entanto, esse acréscimo não foi suficiente face aos valores que se registaram.

Quando se comparam as tecnologias, verifica-se que é no DTH que surge o maior desvio em termos percentuais relativamente ao valor de fecho do mês. Este justifica-se principalmente por na primeira previsão não ser considerada a campanha comercial já referida cujo impacto se resume apenas a esta tecnologia. Assim sendo, enquanto que o desvio do valor do final do mês para DTH foi de 12,6%, para cabo o mesmo foi de 4,8%. Contudo, a precisão do modelo aumenta consideravelmente quando a mesma passou a ser considerada, resultando numa precisão final de 97,3%. Como se pode verificar na tabela, quanto mais próxima a data da previsão se encontra do último dia do mês, menor o desvio face ao real. Os resultados por tecnologia encontram-se a partir do Anexo E.

Tabela 2 - Desvio percentual – Previsão dos *gross-adds* de abril

	Desvio Percentual		
	Cabo	DTH	Total
01/04/2019	-4,8%	-12,6%	-7,3%
10/04/2019	-2,8%	-6,7%	-4,1%
21/04/2019	-2,5%	-2,9%	-2,7%

Após o final do mês de abril, foram introduzidas algumas melhorias ao modelo em termos de taxas de conversão de vendas brutas em *gross-adds* e do respetivo tempo. Foi também detetado que, de um modo geral, o tempo que, em média, cada venda bruta levava a tornar-se um *gross-add* tinha vindo a baixar ligeiramente ao longo dos meses. Por essa razão, na estimativa do tempo e das taxas de conversão foram apenas considerados os meses de 2019, ao contrário, do que acontecia anteriormente, em que o segundo semestre de 2018 era também tido em conta.

Maio

Relativamente ao mês de maio, pela comparação da previsão diária de vendas face aos valores reais, pode-se reparar, desde logo numa quebra acentuada nas vendas reais ao dia 14 do mês. Este deveu-se a um problema ocorrido que manteve alguns sistemas de informação da empresa em baixo, o que teve naturais consequências negativas em termos de vendas e *gross-adds*.

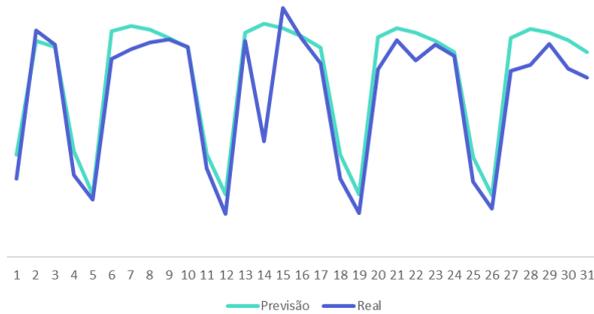


Figura 45 - Primeira previsão de vendas brutas

Para além disso, nota-se que nas últimas duas semanas de maio, o número de vendas previstas encontra-se sempre acima do real pelo facto de as vendas terem baixado o seu nível quando comparadas com as semanas anteriores. Esse desvio acontece, principalmente, nas vendas de serviços por cabo, em que o número de vendas foi, de um modo geral, mais baixo. A juntar a este facto, as aberturas planeadas de FTTH e respetiva conversão em vendas brutas ficaram aquém do que estava planeado, sendo um considerável número de células adiadas para o mês de junho.

Em termos de DTH, a previsão manteve-se em linha com os valores reais, não se verificando desvios consideráveis, com exceção do dia afetado pelo problema ocorrido.

Como se pode reparar na Tabela 3, uma vez mais, quanto menor o horizonte da previsão, menores são os valores do erro médio absoluto percentual, aferindo-se com uma maior precisão o valor das vendas brutas e por consequência, o número de *gross-adds*.

Tabela 3 - MAPE - Previsão das vendas brutas de maio

	MAPE					Desvio - Total Vendas
	1-31	10-31	16-31	20-31	27-31	
01/05/2019	17,3%	19,6%	15,9%	14,3%	14,3%	11,0%
10/05/2019	-	21,2%	18,1%	16,1%	16,0%	10,5%
16/05/2019	-	-	14,2%	12,7%	12,9%	5,7%
20/05/2019	-	-	-	15,1%	16,7%	5,5%
27/05/2019	-	-	-	-	6,9%	1,3%

Analisando agora os resultados do modelo em termos de número de *gross-adds*, pode-se reparar no desvio ocorrido nos primeiros dias do mês. Não foi registado, no entanto, um aumento considerável das vendas brutas nos últimos dias de abril que pudesse explicar o sucedido, pelo que a explicação passa por um maior número de ativações nos próprios dias.

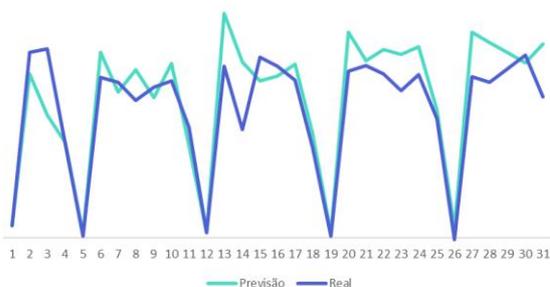


Figura 46 - Primeira previsão diária de *gross-adds*

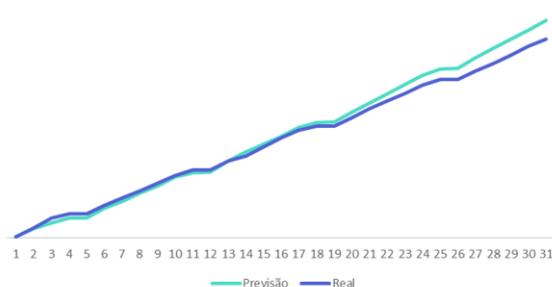


Figura 47 - Primeira previsão acumulada de *gross-adds*

Ao dia 14 de maio, repara-se também num desvio acentuado da previsão face ao real devido ao problema já descrito, que para além de ter influência nos *gross-adds* do próprio dia, faz

também baixar ligeiramente os *gross-adds* nos dias seguintes devido à quebra registada nas vendas.

A juntar-se a este facto um nível de vendas mais baixo nas últimas duas semanas do mês, ajuda a justificar uma previsão de *gross-adds* acima dos valores reais na segunda metade de maio.

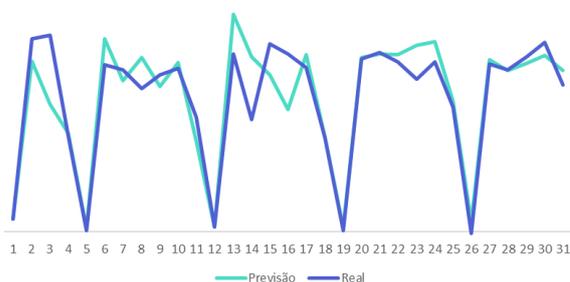


Figura 48 - Previsão diária mais atualizada de *gross-adds*

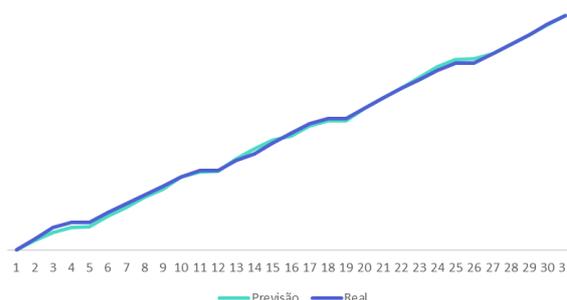


Figura 49 - Previsão acumulada mais atualizada de *gross-adds*

Por outro lado, se analisarmos os resultados tendo em conta a previsão mais atualizada, pode-se reparar que os desvios dos valores previstos face aos reais são bastante inferiores. A última previsão realizada no início da última semana de maio já tem em conta a redução do número de vendas brutas do próprio mês e assim, permitiu uma previsão do final do mês com uma *accuracy* perfeita.

Tabela 4 - Desvio percentual - Previsão dos *gross-adds* de maio

	Cabo	DTH	Total
01/05/2019	11,6%	4,3%	9,3%
10/05/2019	9,9%	7,2%	9,1%
16/05/2019	6,2%	2,2%	4,9%
20/05/2019	7,1%	1,8%	5,5%
27/05/2019	0,5%	-1,0%	0,0%

Esta última previsão demonstra, acima de tudo, uma boa precisão das taxas de conversão de vendas brutas em *gross-adds* e do respetivo tempo, tendo em conta que os valores das vendas reais até à data irão representar uma parte considerável dos *gross-adds* dos últimos dias do mês.

Mais uma vez, verifica-se que o modelo apresenta uma maior precisão do número de *gross-adds* à medida que o mês vai decorrendo visto que o erro associado à previsão das vendas brutas é cada vez menor.

Em suma, a previsão dos *gross-adds* mostra dois meses com resultados muito distintos. Em abril, um mês com poucos dias úteis e impactado negativamente pelos vários feriados do mês, verificou-se uma previsão abaixo dos valores reais. Por outro lado, no mês de maio, em que existe um maior número de dias úteis, a previsão inicial ficou bastante acima do real, visto que o nível de vendas nas últimas semanas não se manteve.

5.2 Previsão do *churn* voluntário

Relativamente ao *churn* voluntário, interessa, em primeiro lugar, analisar os resultados da previsão do número de clientes que entra em retenção por tecnologia.

Como se pode reparar na Tabela 5, os desvios percentuais da previsão do número total de clientes são baixos, com exceção da semana 20 em que esse número é sobrestimado.

Outra métrica importante para avaliar o impacto que esta previsão pode ter nos valores do *churn* é o erro médio absoluto percentual da previsão de todos os escalões do período de fidelização. Essa métrica é, especialmente, elevada para a tecnologia DTH, explicada por um valor de R^2 mais baixo na regressão realizada.

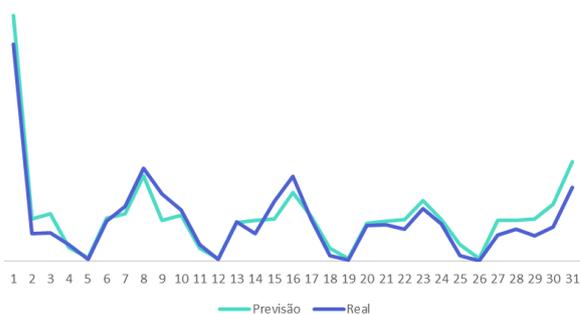
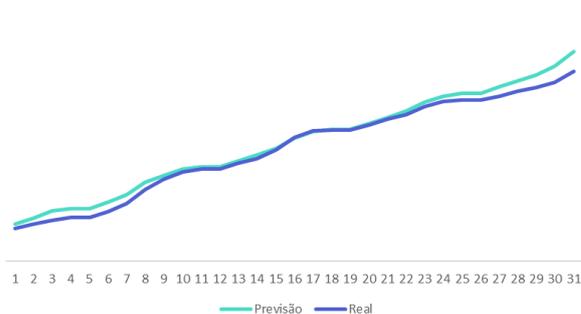
Tabela 5 – Desvio percentual e MAPE da previsão do número de entradas em retenção

Semana	HFC		FTTH		DTH		TOTAL
	Desvio - Total	MAPE	Desvio - Total	MAPE	Desvio - Total	MAPE	
201918	1,8%	10,0%	-2,6%	13,0%	6,0%	21,4%	2,2%
201919	3,8%	7,0%	-3,0%	13,1%	0,5%	13,4%	2,4%
201920	14,2%	14,4%	6,4%	19,2%	12,1%	32,6%	13,0%
201921	1,9%	12,5%	-5,0%	15,1%	-3,9%	16,6%	-0,1%
201922	5,7%	25,4%	-1,5%	24,7%	1,8%	33,0%	4,1%

Assim, juntando o facto de os valores de MAPE serem um pouco elevados aos baixos desvios registados nos valores totais, pode-se concluir que, apesar dos desvios registados por escalão, o número de idas à retenção total previsto encontra-se bastante próximo dos valores reais, amenizando assim o impacto negativo na previsão de *churn*.

Neste ponto, é importante também lembrar que as semanas com maior influência no *churn* de maio são as do mês de abril, sendo que as primeiras semanas de maio apresentam ainda algum peso, mas apenas no *churn* do final do mês.

Quanto ao *churn* voluntário, foi realizada uma previsão em três momentos diferentes do mês de maio. A primeira previsão do mês foi realizada tendo em conta a informação disponível até ao dia 1 desse mês apesar do modelo se encontrar operacional apenas durante o próprio mês e encontra-se representada na Figura 50 e Figura 51.

Figura 50 - Primeira previsão diária de *churn* voluntárioFigura 51 - Primeira previsão acumulada de *churn* voluntário

Como é possível reparar, no primeiro dia do mês houve um pequeno desvio da previsão face ao real, principalmente explicado por uma estimativa do peso que o primeiro dia do mês apresenta acima do peso real.

As próximas duas semanas mostram um ligeiro desvio negativo da previsão diária em relação aos valores reais, enquanto que os maiores desvios surgem na última semana do mês, o que fez baixar os valores finais de *churn* face à previsão.

Os mesmos podem ser justificados pelos erros associados à previsão do número de clientes que entram em retenção, especialmente para a tecnologia DTH. Enquanto que nas primeiras semanas de maio, as entradas em retenção com maior impacto nestes valores são as do mês de abril (para as quais se já se tem os valores reais), na última semana as entradas em retenção do próprio mês já apresentam um peso significativo.

Deste modo, á medida que o mês decorre, é natural que a previsão do final do mês se aproxime do valor real visto que se tira partido de informação real quanto às entradas em retenção e as respetivas taxas do mês passam a ser incluídas na estimativa. Ainda assim, como se sabe, devido ao facto de as idas á retenção com maior influência no *churn* serem as que se verificaram 3 a 8 semanas antes da própria semana, a previsão que é realizada no início de cada mês nunca irá sofrer alterações muito significativas durante o resto do mês.

Tabela 6 – Desvio percentual das previsões de *churn* voluntário

	Cabo	DTH	Total
01/05/2019	5,6%	13,0%	7,7%
12/05/2019	4,2%	5,7%	4,6%
19/05/2019	5,4%	4,5%	5,2%

É importante também referir que o processo de automatização do modelo decorreu no mês de maio, sendo que no modelo final é possível consultar diariamente os valores da previsão de *churn*, o que não acontecia no momento em que se realizaram as previsões acima descritas.

5.3 Previsão do *churn* involuntário

Relativamente ao *churn* involuntário, é relevante fazer uma análise dos desvios da previsão do número de *soft disconnects* e *hard disconnects* executados face aos valores reais.

No que diz respeito á previsão de *soft disconnects*, apenas é necessário a previsão dos que ocorreram no dia 1 e 8 de maio. Os desvios verificados nessas previsões seguem-se na Tabela 7.

Tabela 7 - Desvio percentual da previsão de *soft disconnects*

Data SD	DTH	FTTH	HFC
01/05/2019	-0,8%	10,7%	3,4%
08/05/2019	-10,2%	9,3%	-9,0%

Quanto aos *hard disconnects* e tendo em conta que aqueles que são aplicados aos dias 7, 10 e 13 de maio, são previstos com base nos valores reais do número de *soft disconnects*, verifica-se que estes apresentam desvios inferiores comparativamente com os que são previstos com base nas previsões de *soft disconnects*.

Tabela 8 - Desvio percentual da previsão de *hard disconnects*

Data HD	DTH	FTTH	HFC
07/05/2019	-8,5%	2,2%	4,0%
10/05/2019	6,9%	-8,6%	-5,1%
13/05/2019	-18,7%	3,8%	-9,9%
21/05/2019	14,1%	16,4%	3,0%
28/05/2019	-27,0%	6,3%	-24,0%

Assim, quando as previsões do número de SDs são substituídas por valores reais, a previsão do *churn* torna-se mais precisa, como se pode observar na Tabela 9. Deste modo, a previsão ao dia 12 de maio apresenta um desvio de apenas 3,1% face ao valor do final do mês, cerca de metade do desvio verificado na previsão no primeiro dia do mês.

Tabela 9 - Desvio percentual da previsão de *churn*

Data Previsão	Cabo	DTH	Total
01/05/2019	-5,4%	-8,8%	-6,5%
12/05/2019	-5,1%	1,0%	-3,1%

Na Figura 52 e Figura 53, pode-se reparar na evolução do *churn* involuntário ao longo do mês, com especial destaque para o afastamento, nos últimos dias do mês, da previsão face ao real, que se deveu ao desvio verificado no número de *hard disconnects* efetuados nos dias 28 e 29 de maio.

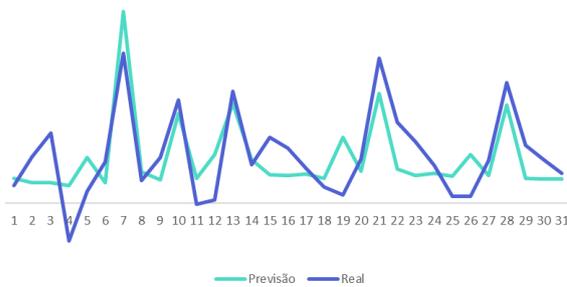


Figura 52 - Primeira previsão diária de *churn* involuntário

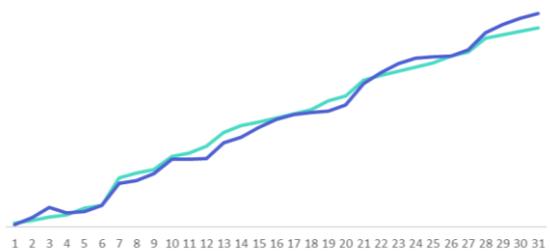


Figura 53 - Primeira previsão acumulada de *churn* involuntário

Os grandes desvios que se verificam ao dia devem-se à tecnologia cabo. A distribuição desse *churn* está dependente da performance diária das equipas de desligamento, tendo sido identificadas diferenças significativas entre os diferentes dias, sem conclusões sólidas quanto à razão pela qual acontecem. Ainda assim, a qualidade da previsão mensal do *churn* involuntário não se encontra comprometida.

6 Conclusões

O projeto de dissertação desenvolvido permitiu identificar os principais fatores que impactam o processo de captação e de *churn*, voluntário e involuntário.

Em relação às vendas, constatou-se a existência de sazonalidade e alguma regularidade nos seus valores que tornou possível obter-se uma boa *accuracy* com recurso ao método de previsão *Holt Winters*.

Devido à constante instalação de fibra em novas zonas, verificou-se a necessidade de a previsão de vendas de fibra seguir uma abordagem distinta. Assim, as vendas associadas a células cuja abertura ocorreu há menos de 36 semanas foram previstas com base nos momentos de instalações de fibra e a respetiva estimativa das taxas de captação.

No que diz respeito ao *churn* voluntário, a entrada em retenção do cliente é o primeiro indício de que um cliente vai ser *churn*. Assim, foi realizada uma estimativa no número de clientes que entra em retenção para cada tecnologia, obtendo-se um bom ajustamento entre os vários momentos do período de fidelização e o número de entradas em retenção. As taxas de retenção foram estimadas para cada tecnologia, canal e escalão do período de fidelização e permitem, de acordo o número de entradas em retenção reais e previstas, prever o número de clientes que vão ser *churn* no curto prazo.

Quanto ao *churn* involuntário, foi realizada uma previsão do número de clientes em cada fase do processo a partir da estimativa das taxas de conversão entre as diferentes fases, que registam diferenças significativas entre os vários ciclos de faturação.

Pela análise dos resultados obtidos do modelo preditivo, conclui-se que, apesar da primeira previsão realizada ter uma margem de erro acima do pretendido, as previsões vão sendo renovadas com o decorrer do mês, verificando-se um aumento considerável da precisão do modelo.

Deste modo, a sua utilização como ferramenta para determinar o valor de *gross-adds* final do mês deve ser tida em consideração quando já se percebe o nível de vendas que o próprio mês tem apresentado. Por outro lado, no início de cada mês, pode ser encarada como uma ferramenta que permite avaliar como será a performance da empresa mantendo-se o mesmo nível de vendas das últimas semanas. No caso de *churn*, a previsão durante o mês não sofre grandes variações visto que os clientes que abandonam o portfolio da empresa no próprio mês já se encontram com os seus processos de *churn* voluntário ou involuntário a decorrer.

O facto da previsão de *gross-adds* mostrar dois meses com resultados muito distintos, ou seja, um mês em que se a previsão se encontrou abaixo dos valores reais e outro em que se sobrestimou o número de clientes captados, leva-nos a concluir que, para além de todos os fatores que são tidos em consideração no modelo preditivo há sempre outros que se revelam difíceis de modelar. Estes encontram-se, essencialmente, relacionados com o fator humano, mais concretamente, com a performance das equipas de vendas e das de ativações. No caso do *churn*, a performance das equipas de desligamentos também influencia o *churn* e, de igual forma, é difícil de estimar o seu impacto.

Deste modo, a componente subjetiva deve assumir um papel relevante na execução das previsões, sendo por isso importante os *inputs* subjetivos das pessoas que estão inseridas num contexto mais operacional.

De um modo geral, pode-se afirmar que os objetivos do projeto foram cumpridos, uma vez que há agora um conhecimento total dos processos que permitem a captação e a perda de clientes bem como dos principais fatores que impactam os mesmos. Para além disso, o modelo desenvolvido permite uma previsão com uma *accuracy* satisfatória, especialmente, durante o decorrer do próprio mês. Para além disso, o modelo encontra-se automatizado de forma a que com pouca interação humana se possam obter as respetivas previsões.

Por outro lado, os resultados permitem também concluir que existe ainda potencial de melhoria do modelo preditivo desenvolvido. Assim, como trabalho futuro, a incorporação do agendamento das ordens de trabalho poderá assumir-se como uma ferramenta relevante na previsão do *churn*, voluntário e involuntário. Estas referem-se ao agendamento dos desligamentos a realizar pela empresa, sendo a última etapa antes de ocorrer o *churn* para a tecnologia HFC que, como se sabe, representa a sua maioria. Identificando-se com alguma antecedência o número de desligamentos previstos para cada dia, pode-se saber com toda a certeza o número de *churners* nesse mesmo dia. No entanto, para que tal aconteça, é necessária uma maior qualidade da informação no que diz respeito às datas de agendamento. Pela informação recolhida, apenas uma pequena parte dos desligamentos se realizam efetivamente na data prevista, sendo os restantes realizados no prazo de 7 ou 14 dias, o que torna difícil tirar proveito dessa informação. Por essa razão, deve ser investigado com uma maior profundidade a razão pela qual isso acontece e se é possível melhorar a qualidade da informação. A incorporação desta informação iria conferir uma maior autenticidade ao modelo contribuindo, certamente, para uma melhoria da precisão do modelo de *churn*.

Para além disso, é possível melhorar a *accuracy* do modelo se mais variáveis forem consideradas. Assim, variáveis como o pacote subscrito/número de serviços subscritos, o número de chamadas efetuadas para o *call center* e dados pessoais como a idade do cliente ou morada devem também ser exploradas.

A influência dos dias especiais que tem no *churn* não é comparável com a que os mesmos apresentam nos *gross-adds*. Ainda assim, estes podem também ser tidos em consideração de forma a aumentar a precisão das previsões. Estes apresentam influência nas idas à retenção pelo facto de os consumidores alterarem as suas rotinas nesses dias, fazendo variar não só o número de entradas em retenção, mas também a sua distribuição pelos diferentes canais.

Para além disso, estes dias especiais têm também influência no tempo que decorre desde que o cliente é considerado como não recuperado até que é executado o desligamento, uma vez que a força de trabalho das equipas de desligamento é também mais reduzida, pelo que a incorporação da informação quanto ao agendamento dos desligamentos no modelo permitiria resolver esta questão.

Quanto à previsão dos *gross-adds* e atendendo a que o modelo foi sendo aperfeiçoado ao longo do tempo, há uma menor margem para a introdução de melhorias. Ainda assim, é essencial que a estimativa de conversão de vendas brutas em *gross-adds* apresente a maior precisão possível, sendo assim importante que a informação real quanto ao número de *gross-adds* obtidos de acordo com as vendas realizadas, se encontra o mais atualizada possível.

Para além disso, podem ainda ser investigadas as diferenças que existem em termos de conversão de venda bruta em *gross-add* e do respetivo tempo entre os vários canais de vendas. Assim, pode-se aumentar a precisão da estimativa através de uma possível desagregação por canal, tendo em consideração o potencial aumento de complexidade do modelo que daqui pode resultar.

Referências

- Amina, A., Anwara, S., Adnan, A., Nawaz, M., Alawfi, K., Hussain, A., & Huang, K. (2017). Customer churn prediction in the telecommunication sector using a rough set approach. Em *Neurocomputing* (pp. 242-254). Elsevier.
- ANACOM - *Períodos de fidelização - Voz, Internet, TV*. (abril de 2019).
- Aronsson, H. (2015). *Modeling Strategies Using Predictive Analytics*.
- Barrett, J. P. (1974). The Coefficient of Determination—Some Limitations. *The American Statistician*.
- Batorsky, D. e. (2000). Telecom Operations Map. *TeleManagement Forum*.
- Berkhin, P. (2006). *Survey of Clustering Data Mining Techniques*.
- Chatfield, C. (1978). The Holt-Winters Forecasting Procedure. *Applied Statistics*.
- Chatfield, C., & Mohammad, Y. (1988). Holt-Winters forecasting: some practical issues. *The Statistician*.
- Crie, D. (1996). *Rétention de clientèle et fidélité des clients*.
- Ekholm, B. (2018). How 5G could speed up global growth. *World Economic Forum Annual Meeting*.
- Esling, P., & Agon, C. (2012). Time-series data mining. *ACM Computing surveys*.
- Gröne, F., Gupta, U., & Samad, R. A. (2018). *Grasping at differentiated strwas and Commoditization in the wireless telecom industry*. PricewaterhouseCoopers.
- Hadden, J., Tiwari, A., Roy, R., & Ruta, D. (2008). Churn Prediction: Does Technology Matter? *International Journal of Industrial and Manufacturing Engineering*.
- Han, J., Kamber, M., & Pei, J. (2017). *Data mining: concepts and techniques*. Elsevier.
- Harrison, P. J. (1967). Exponential Smoothing and Short-Term Sales Forecasting. *Management Science*.
- Huang, Y., Zhu, F., Yuan, M., Deng, K., & Li, Y. N. (2015). Telco Churn Prediction with Big Data.
- Jaroszewicz, S. (2008). Cross-selling models for telecommunication services. *Journal of telecommunications and information technology*.
- Juriscic, B., & Azevedo, A. (2011). Building customer–brand relationships in the mobile communications market: The role of brand tribalism and brand reputation. *Journal of Brand Management*.
- Jyh-Fu Jeng, D., & Bailey, T. (2012). *Assessing customer retention strategies in mobile telecommunications*.

- Kamakura, W. A. (2008). Cross-Selling: Offering the Right Product to the Right Customer at the Right Time. *Journal of Relationship Marketing*.
- Keramati, A. (2011). A Proposal Framework for investigating the impact of customer retention in e-commerce.
- Lemmens, A., & Croux, C. (2006). Bagging and Boosting Classification Trees to Predict Churn. *Journal of Marketing Research*.
- Livne, G., Simpson, A., & Talmor, E. (2011). *Do Customer Acquisition Cost, Retention and Usage Matter to Firm Performance and Valuation?*
- Makridakis, P. G., Wheelright, S. C., & Hyndman, R. J. (1998). *Forecasting: Methods and Applications*.
- Mattison, R. (2005). *The Telco Churn Management Handbook*.
- Mentzer, J. T., & Kahn, K. B. (1995). Forecasting Technique Familiarity, Satisfaction, Usage and Application. *Journal of Forecasting*.
- Morreale, P. (2000). *The Telecommunications Handbook*.
- Peppers, D., & Rogers, M. (2004). *Managing Customer Relationships: A Strategic Framework*.
- Rosenberg, L. J., & Czeipel, J. A. (2007). A Marketing Approach For Customer Retention. *The Journal Of Consumer Marketing*.
- Srivastava, A. K. (1995). The coefficient of determination and its adjusted version in linear regression models. *Econometric Reviews*.
- Strouse, K. G. (1999). *Marketing Telecommunications Services: New Approaches for a Changing Environment*.
- Tsai, C.-F., & Lu, Y.-H. (2009). Customer churn prediction by hybrid neural networks. *Elsevier*.
- Vafeiadis, T., Diamantaras, K. I., Sarigiannidis, G., & Chatzisavvas, K. (2015). *A comparison of machine learning techniques for customer churn prediction*.
- Ye, L., Qiuru, C., Haixu, X., Yijun, L., & Guangping, Z. (2013). Customer Segmentation for Telecom with the k-means Clustering Method. *Information Technology Journal*.

ANEXO A: Entradas em Retenção por dia da semana e canal

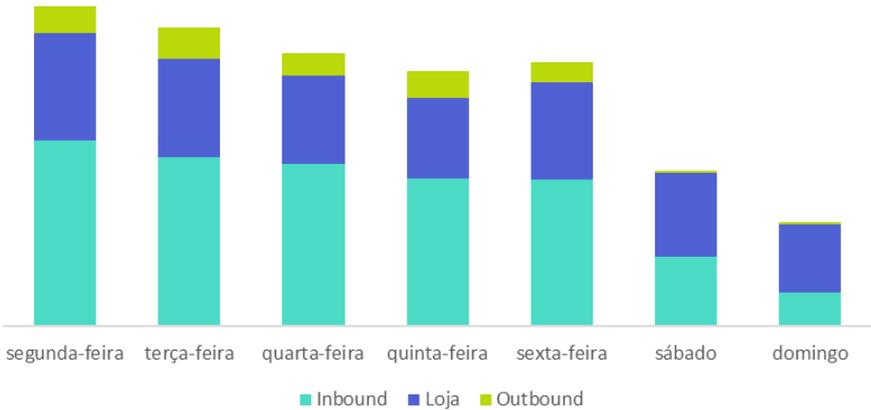


Figura 54 - Entradas em retenção por dia da semana e canal

ANEXO B: *Churn* voluntário para cada tecnologia

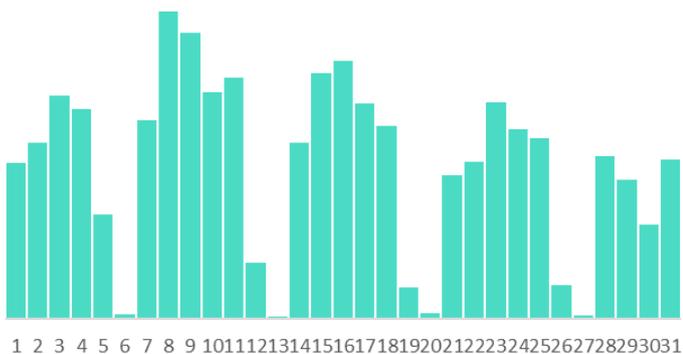


Figura 55 - *Churn* involuntário - HFC

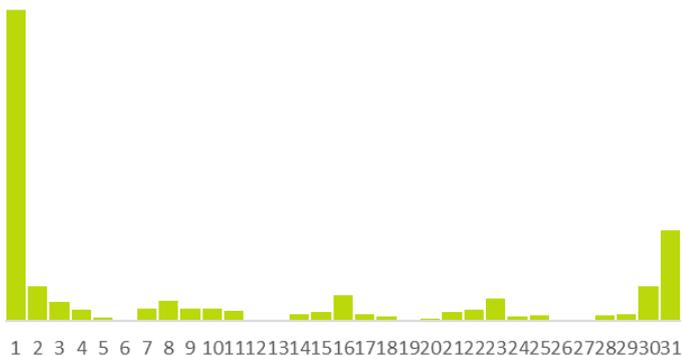


Figura 56 - *Churn* involuntário - DTH

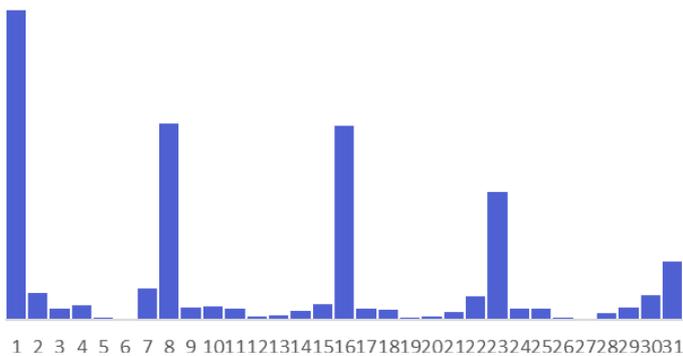


Figura 57 - *Churn* involuntário - FTTH

ANEXO C: *Holt Winters*

Data	Real	nt	bt	ft	Previsão
01/06/2018	100			1,1	
02/06/2018	55			0,6	
03/06/2018	43			0,5	
04/06/2018	116			1,3	
05/06/2018	106			1,2	
06/06/2018	108			1,2	
07/06/2018	99	89,7	0,0	1,1	
08/06/2018	124	89,7	0,0	1,1	
09/06/2018	57	89,7	0,0	0,6	
10/06/2018	59	89,8	0,0	0,5	
11/06/2018	118	89,8	0,0	1,3	
12/06/2018	123	89,8	0,1	1,2	
13/06/2018	117	89,8	0,1	1,2	
14/06/2018	100	89,9	0,1	1,1	99
15/06/2018	97	89,9	0,1	1,1	102
...
01/04/2019	95	79,9	-0,1	1,2	95
02/04/2019	99	79,9	-0,1	1,2	99
03/04/2019	97	79,8	-0,1	1,2	97
04/04/2019	95	79,8	-0,1	1,2	95
05/04/2019	86	79,7	-0,1	1,1	86
06/04/2019	49	79,7	-0,1	0,6	49
07/04/2019	35	79,6	-0,1	0,4	35
08/04/2019		79,6	-0,1	1,2	94
09/04/2019		79,5	-0,1	1,2	98
10/04/2019		79,5	-0,1	1,2	96
11/04/2019		79,4	-0,1	1,2	95
12/04/2019		79,4	-0,1	1,1	85
13/04/2019		79,3	-0,1	0,6	48
14/04/2019		79,3	-0,1	0,4	34
...

alfa	0,0004
beta	1
gama	0,0697
MSE	258,66
MAPE	7,03%
U theil	0,22

Figura 58 – *Holt Winters* DTH - Sem outbound

ANEXO D: Resultados Regressão - ANOVA

<i>Regression Statistics</i>	
Multiple R	0,924
R Square	0,854
Adjusted R Square	0,830
Standard Error	384,406
Observations	58

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	8	42252546,4	5281568,3	35,742	6,9473E-18
Residual	49	7240618,96	147767,734		
Total	57	49493165,4			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
Intercept	765,195	587,540	1,302	0,199	-415,512	1945,902	-415,512	1945,902
COM PF]0,4]	0,209	0,050	4,213	0,000	0,109	0,309	0,109	0,309
COM PF]12,24]	0,119	0,019	6,304	0,000	0,081	0,157	0,081	0,157
COM PF]4,8]	0,157	0,056	2,772	0,008	0,043	0,270	0,043	0,270
SEM PF [0,4]	-0,157	0,037	-4,181	0,000	-0,232	-0,081	-0,232	-0,081
SEM PF]8,12]	-0,119	0,059	-2,010	0,050	-0,238	0,000	-0,238	0,000
SEM PF]24,36]	0,343	0,057	6,019	0,000	0,229	0,458	0,229	0,458
COM PF]8,12]	-0,076	0,043	-1,765	0,084	-0,162	0,011	-0,162	0,011
COM PF]84,104]	-0,037	0,015	-2,513	0,015	-0,066	-0,007	-0,066	-0,007

Figura 59 - Regressão - HFC

<i>Regression Statistics</i>	
Multiple R	0,956
R Square	0,915
Adjusted R Square	0,903
Standard Error	63,326
Observations	58

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	7	2153240,46	307605,78	76,706	1,633E-24
Residual	50	200509,192	4010,18385		
Total	57	2353749,66			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
Intercept	602,430	125,779	4,790	0,000	349,796	855,065	349,796	855,065
COM PF]0,4]	0,283	0,076	3,715	0,001	0,130	0,436	0,130	0,436
COM PF]12,24]	0,074	0,015	5,072	0,000	0,045	0,104	0,045	0,104
COM PF]4,8]	0,318	0,094	3,379	0,001	0,129	0,507	0,129	0,507
SEM PF [0,4]	0,213	0,073	2,911	0,005	0,066	0,360	0,066	0,360
SEM PF]4,8]	-0,179	0,112	-1,597	0,117	-0,404	0,046	-0,404	0,046
SEM PF]12,24]	-0,305	0,114	-2,685	0,010	-0,533	-0,077	-0,533	-0,077
COM PF]8,12]	-0,162	0,087	-1,874	0,067	-0,336	0,012	-0,336	0,012

Figura 60 - Regressão - FTTH

<i>Regression Statistics</i>	
Multiple R	0,802
R Square	0,643
Adjusted R Square	0,576
Standard Error	158,342
Observations	58

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	9	2170396,87	241155,208	9,618	3,3056E-08
Residual	48	1203466,65	25072,222		
Total	57	3373863,52			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
Intercept	12267,713	3743,685	3,277	0,002	4740,529	19794,898	4740,529	19794,898
COM PF]0,4]	0,158	0,088	1,808	0,077	-0,018	0,334	-0,018	0,334
COM PF]12,24]	0,087	0,043	2,023	0,049	0,001	0,173	0,001	0,173
COM PF]4,8]	0,250	0,088	2,836	0,007	0,073	0,427	0,073	0,427
SEM PF [0,4]	0,255	0,066	3,865	0,000	0,122	0,387	0,122	0,387
SEM PF]12,24]	-0,303	0,152	-1,992	0,052	-0,610	0,003	-0,610	0,003
SEM PF]24,36]	-0,387	0,150	-2,575	0,013	-0,689	-0,085	-0,689	-0,085
COM PF]84,104]	-0,145	0,047	-3,094	0,003	-0,239	-0,051	-0,239	-0,051
SEM PF]36,48]	-0,502	0,157	-3,206	0,002	-0,817	-0,187	-0,817	-0,187
SEM PF]48,60]	-0,352	0,182	-1,937	0,059	-0,718	0,013	-0,718	0,013

Figura 61 - Regressão - DTH

ANEXO E: Previsão de *gross-adds* cabo - abril

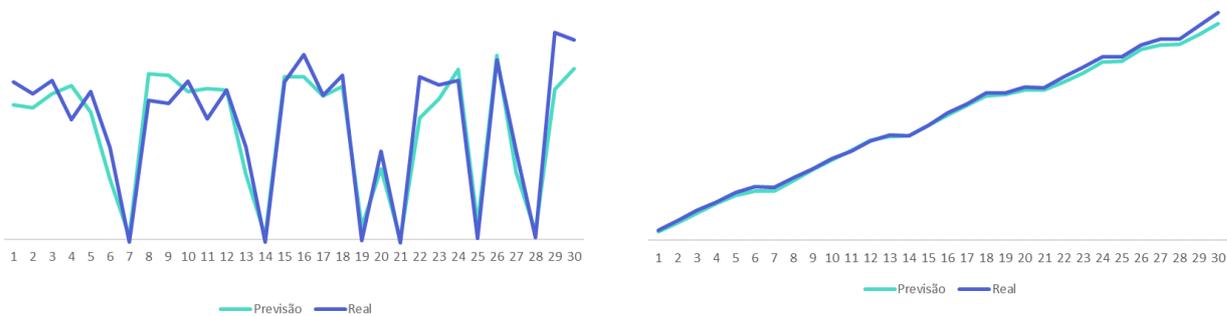


Figura 62 - Primeira previsão de *gross-adds* cabo - abril

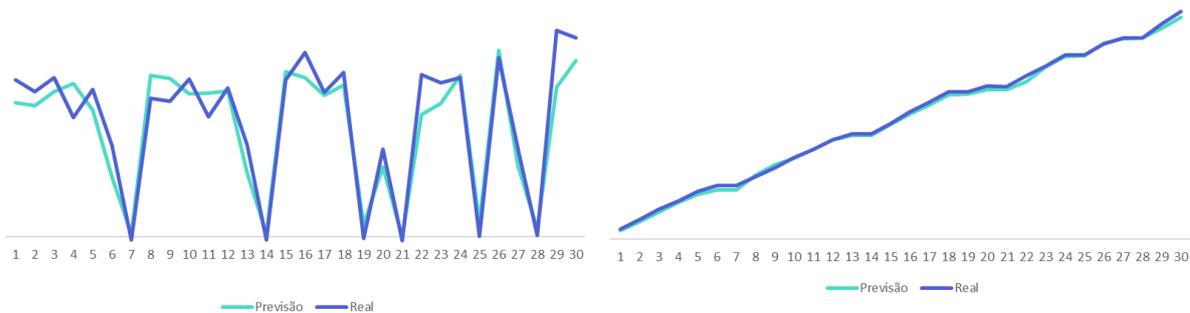


Figura 63 - Previsão mais atualizada de *gross-adds* cabo - abril

ANEXO F: Previsão de *gross-adds* DTH - abril

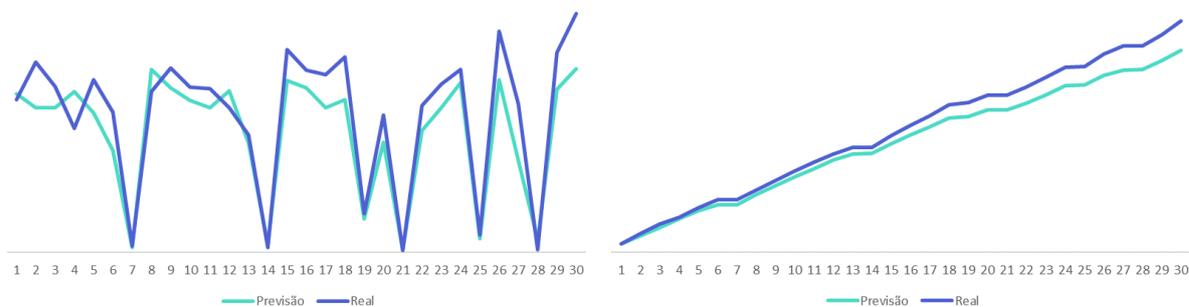


Figura 64 – Primeira previsão de *gross-adds* DTH - abril

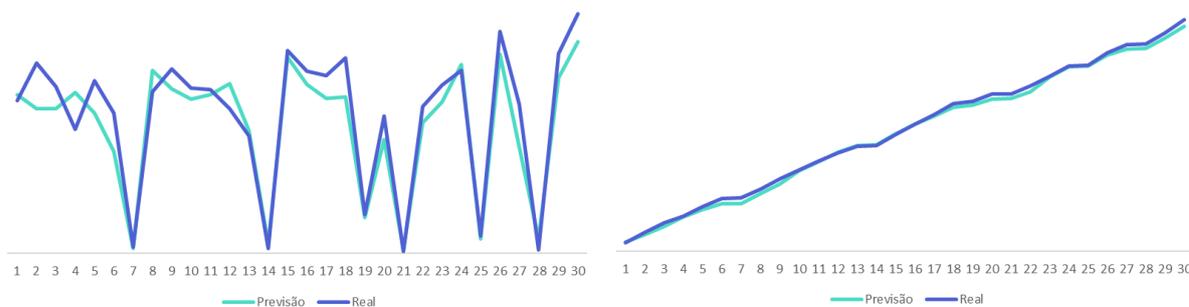


Figura 65 – Previsão mais atualizada de *gross-adds* DTH - abril

ANEXO G: Previsão de *gross-adds* cabo - maio

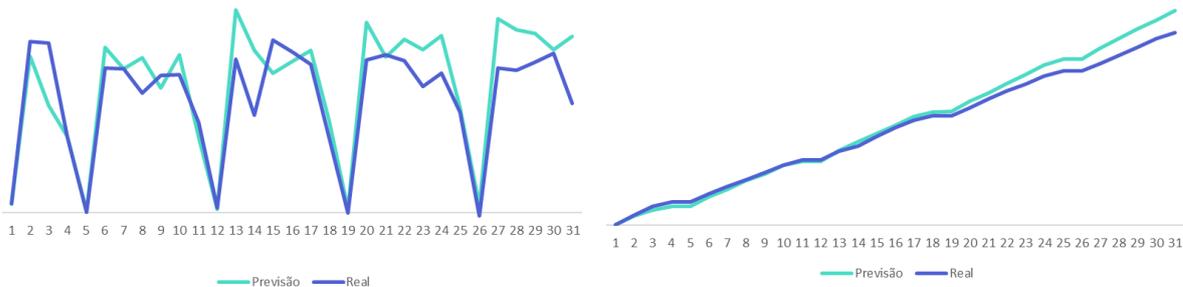


Figura 66 – Primeira previsão de *gross-adds* cabo - maio

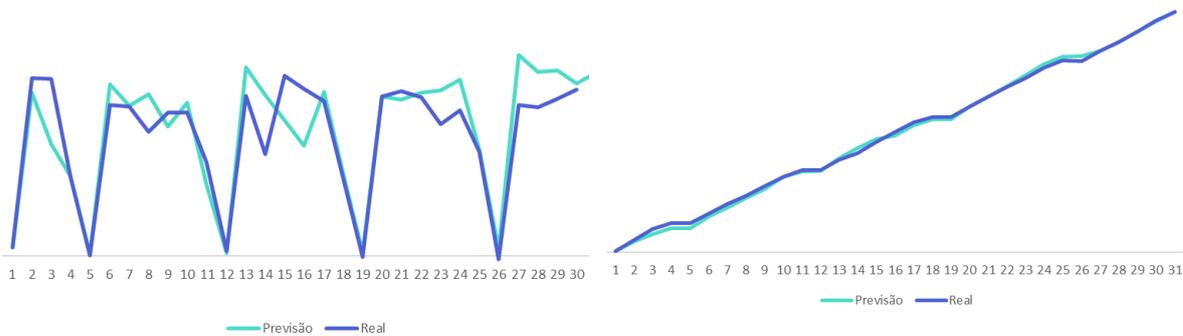


Figura 67 - Previsão mais atualizada de *gross-adds* cabo - maio

ANEXO H: Previsão de *gross-adds* DTH - maio

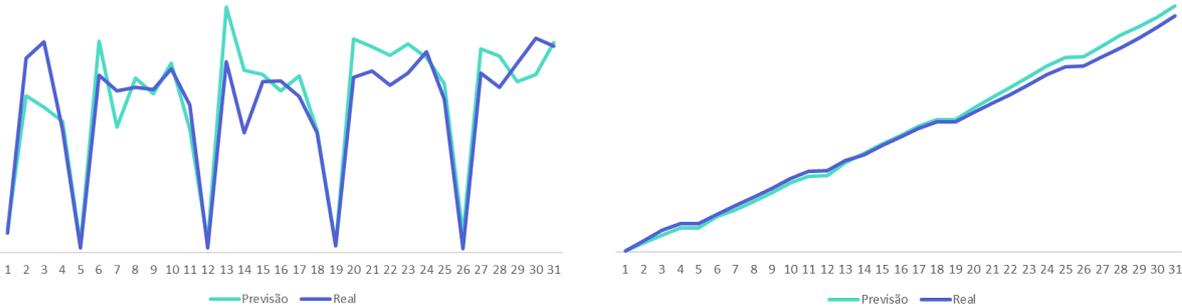


Figura 68 – Primeira previsão de *gross-adds* DTH - maio

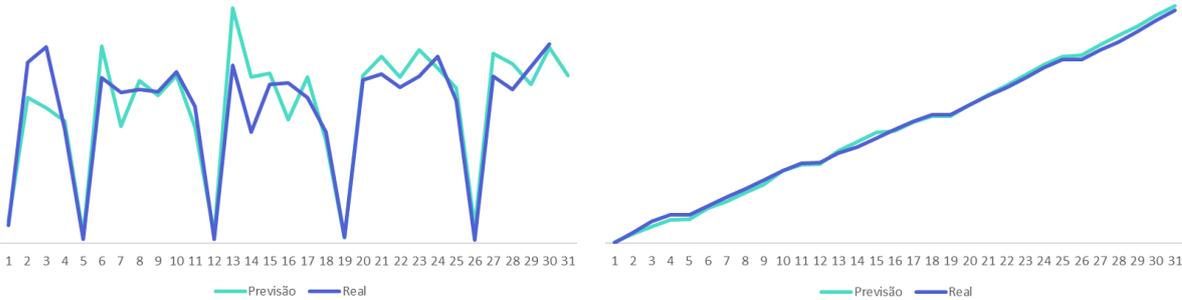


Figura 69 - Previsão mais atualizada de *gross-adds* DTH - maio

ANEXO I: Previsão do *churn* voluntário cabo

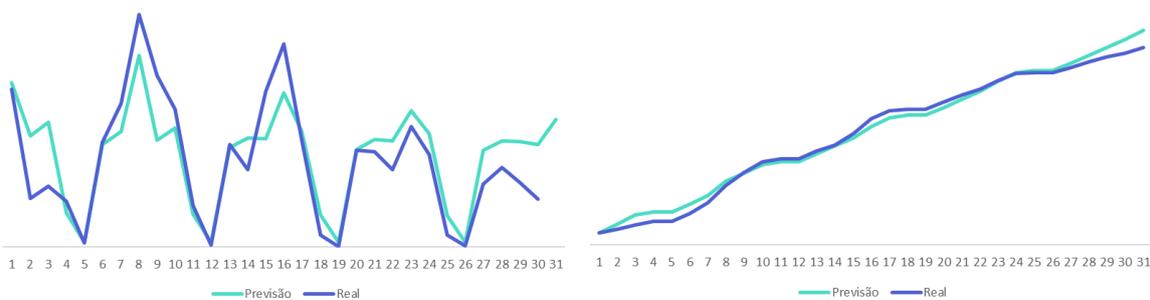


Figura 70 - Primeira previsão de *churn* voluntário cabo

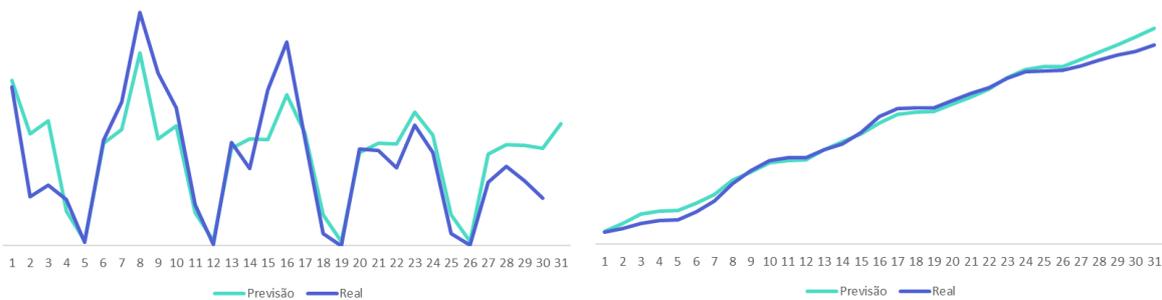


Figura 71 - Previsão mais atualizada de *churn* voluntário cabo

ANEXO J: Previsão do *churn* voluntário DTH

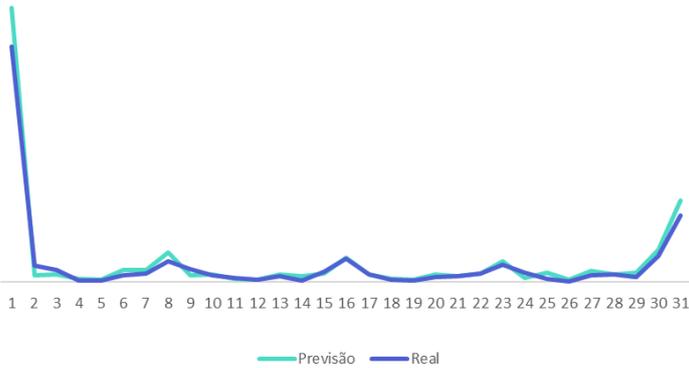


Figura 72 – Primeira previsão de *churn* involuntário DTH

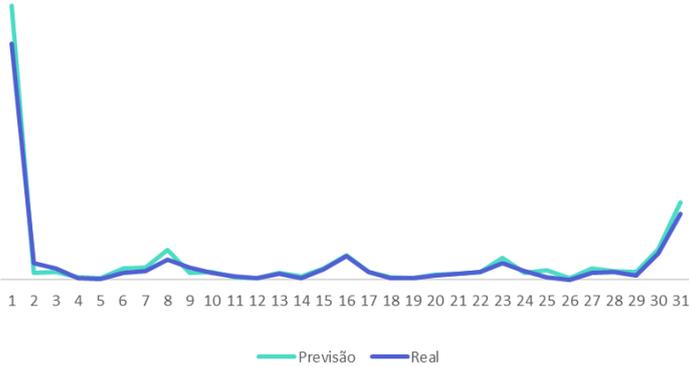


Figura 73 – Previsão mais atualizada de *churn* involuntário DTH

ANEXO K: Previsão do *churn* involuntário cabo

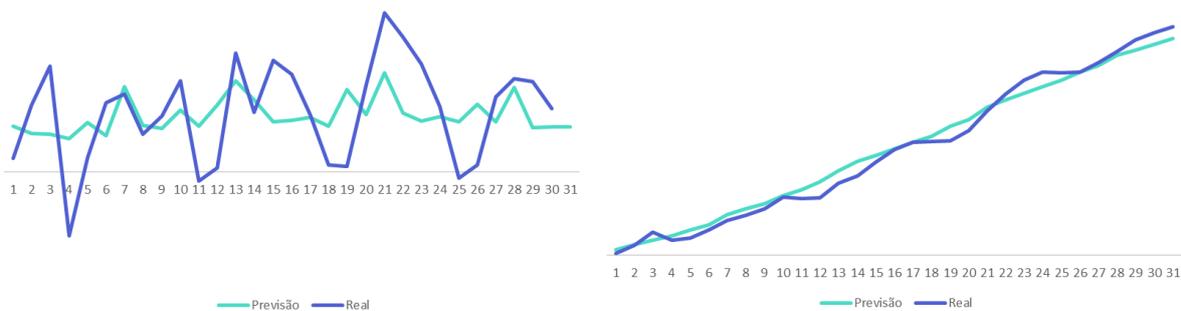


Figura 74 – Primeira previsão de *churn* involuntário cabo

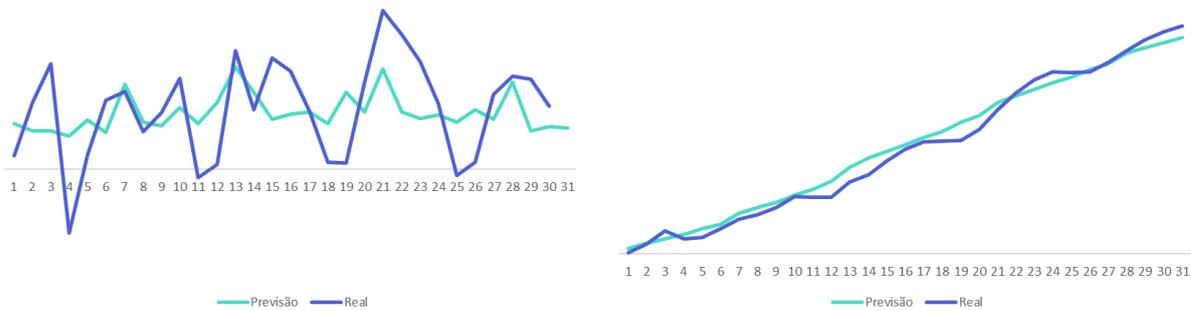


Figura 75 – Previsão mais atualizada de *churn* involuntário cabo

ANEXO L: Previsão do *churn* involuntário DTH

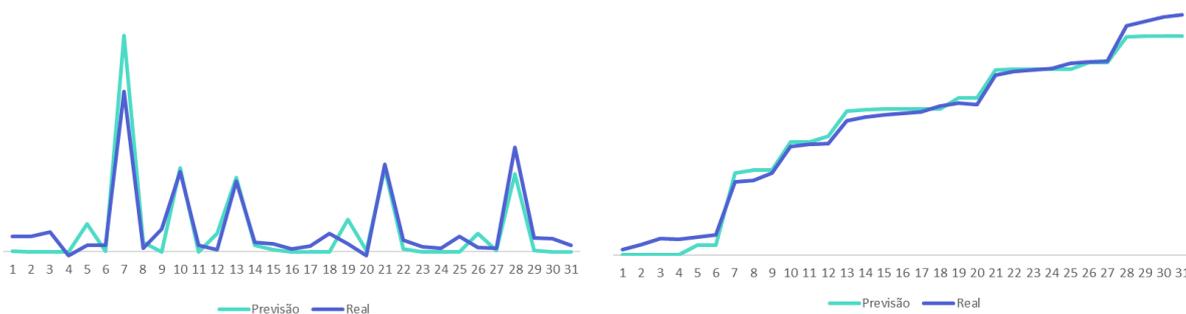


Figura 76 – Primeira previsão de *churn* involuntário DTH

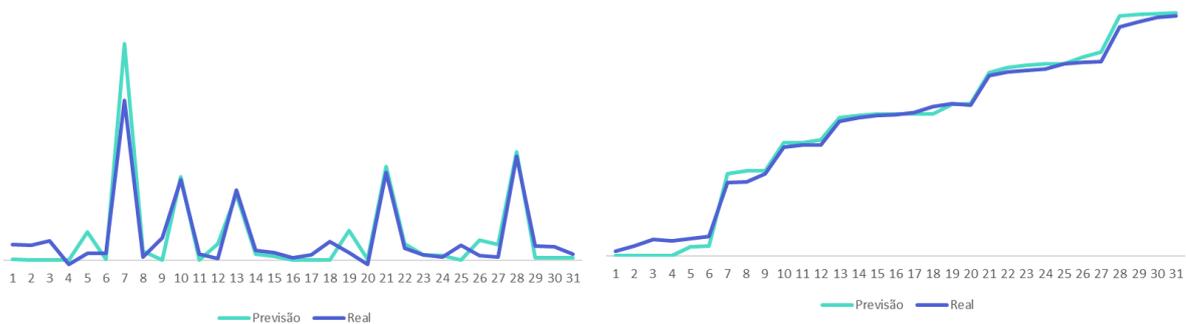


Figura 77 – Previsão mais atualizada de *churn* involuntário DTH