

CADERNOS DO IME – Série Estatística

Universidade do Estado do Rio de Janeiro - UERJ
ISSN on-line 2317-4536 / ISSN impresso 1413-9022 - v.48, p. 36 - 49, 2020
DOI: 10.12957/cadest.2020.55671

UTILIZAÇÃO DOS MODELOS ARIMA PARA PREVISÃO DA TAXA DE CHURN: ESTUDO DE CASO PARA UMA EMPRESA DE E-COMMERCE

Eduardo Dabul Torres
Universidade do Estado do Rio de Janeiro – UERJ
eduardodabul@gmail.com

João Marcos Gomes da Silva
Universidade do Estado do Rio de Janeiro – UERJ
gs.jmarcos@gmail.com

Marcello Montillo Provenza
Universidade do Estado do Rio de Janeiro – UERJ
mprovenza@gmail.com

Igor Campos de Almeida Lima
Universidade do Estado do Rio de Janeiro – UERJ
ilima@hotmail.com.br

Jorge Luiz de Jesus Goulart
Universidade do Estado do Rio de Janeiro – UERJ
jorge_goulart@outlook.com

Resumo

A taxa de Churn, ou simplesmente Churn, calcula o número de usuários que se desconectam dos serviços de uma empresa em um período de tempo específico. Para alguns setores, esta é uma métrica básica para avaliar o sucesso do negócio, já que apresenta impacto direto no faturamento. Neste trabalho, projeta-se a curto prazo o Churn de uma empresa de e-commerce com base no histórico de seus dados. Para isso, utilizam-se as séries temporais para a previsão desses dados, o modelo Autorregressivo Integrado Médias Móveis (ARIMA). O trabalho passou por todas as etapas do ciclo iterativo de um processo de previsão dos dados, começando do estudo e análise da base de dados, passando pela escolha e validação dos parâmetros do modelo até chegar a projeção dos dados. O teste Dickey-Fuller mostrou que a série é estacionária, o melhor modelo encontrado foi o AR(1) e os resíduos seguem uma distribuição normal.

Palavras-chave: Taxa de Churn, Séries Temporais, ARIMA, Previsão de Dados.

1. Introdução

O *Churn Rate*, ou taxa de *Churn*, é uma métrica ligada a rotatividade de clientes em um negócio e está relacionada ao índice de evasão ou cancelamento de serviços. Este é um indicador muito importante para empresas que prestam serviços ligados diretamente ao consumidor final, como um companhia de TV por assinatura ou uma empresa de vendas pela internet, como para empresas que fornecem serviços para outras empresas, como serviços de limpeza e transportadoras.

Apesar de ser importante, seu cálculo é bem simples. Para se calcular a taxa de *Churn* basta dividir o número de clientes que cancelaram o serviço ao final do período de análise pela quantidade total de clientes que existiam no início do período de análise. É geralmente mais caro conquistar novos clientes do que manter os já existentes. Em muitos setores, o custo de aquisição de novos clientes pode ser cinco vezes superior ao seu custo de retenção (KURTZ & CLOW, 1998).

As empresas de serviços possuem, em sua maioria, uma assinatura mensal e quando o cliente deixa de pagar o mesmo é considerado como um *Churn* de sua base, esse cliente sai da base de clientes ativos da empresa para aquele determinado mês, esse é o método mais simples para se calcular o *Churn* para essas empresas onde os dados são contínuos.

Para as empresas de *e-commerce*, que tem como objetivo a venda de produtos, esses dados não são considerados contínuos, mas sim discretos, pois o cliente não necessariamente tem que comprar todo mês para ser considerado ativo. Por isso, para essas empresas, é utilizado o período de 12 meses para o cliente se tornar *churn*. Por exemplo, o cliente realiza uma compra em jan/12. Para esse cliente ser considerado como *Churn* (saiu da base) ele não pode ter compra nos próximos 12 meses (365 dias). Ele será um cliente *churn* da base de jan/12 no período de jan/13, por isso são necessários 12 meses para analisar esses clientes.

Neste trabalho, será analisada uma base com a taxa de *Churn* de uma empresa de vendas *e-commerce*, para entender melhor seu compartimento e para que possa ser prevista a curto prazo. São usados os modelos Auto Regressivos Integragos Médias Móveis (ARIMA) de séries temporais como método de previsão da taxa de *Churn*. Foi utilizada a base histórica de clientes compradores e clientes que deixaram de comprar no período de janeiro de 2012 até dezembro de 2015 para treino e, para projeção, usou-se os

próximos 18 meses (janeiro de 2016 até junho de 2017). Foi utilizada a aplicação de modelos de séries temporais para a previsão desses dados, de forma a identificar aquele que melhor representa o comportamento desta variável e, portanto, seria indicado para a realização das estimativas.

2. Metodologia

As séries temporais (ou séries históricas) são conjuntos de medidas de uma mesma grandeza, relativas a vários períodos consecutivos (DOWNING & CLARK, 2006).

Uma série temporal é um conjunto de observações ordenadas no tempo (não necessariamente igualmente espaçadas) e que apresentam dependência serial (isto é, dependência entre instantes de tempo). A notação usada aqui para denotar uma série temporal será $Y_1, Y_2, Y_3, \dots, Y_t, \dots, Y_T$ que representa uma série de tamanho T . O instante T geralmente indica o último instante disponível (BARROS, 2005).

Os modelos utilizados para descrever séries temporais são processos estocásticos, isto é, processos controlados por leis probabilísticas. Qualquer que seja a classificação que se faça para os modelos de séries temporais, pode-se considerar um número muito grande de modelos diferentes para descrever o comportamento de uma série particular. A construção destes modelos depende de vários fatores, tais como o comportamento do fenômeno ou o conhecimento a priori, que se tem de sua natureza e do objetivo da análise. Na prática, depende também da existência de métodos apropriados de estimação e da disponibilidade de programas adequados (MORETTIN & TOLÓI, 2006).

Para elaboração deste trabalho foi usado o *software* JupyterLab, que é a última versão para notebook (JUPYTER, 2020).

2.1 Componentes da série temporal

Uma série temporal é composta por quatro elementos, sendo eles tendência, ciclo, sazonalidade e aleatório.

A tendência de uma série indica a velocidade das mudanças dos dados ao longo do tempo, sentido de deslocamento, e como esses dados se comportam, crescendo, decrescendo ou permanecendo estáveis. Os casos mais comuns de tendência são Constante, Linear e Quadrático.

Os ciclos são caracterizados pelas oscilações dos dados, seja de subida ou de queda nas séries, de forma suave e repetida, ao longo da componente de tendência. A sazonalidade, assim como o ciclo, também se caracteriza pelas oscilações dos dados. O grande diferencial entre os dois é o comportamento previsível da sazonalidade, ocorrendo em intervalos regulares de tempo enquanto o ciclo tende a ser irregular nesse ponto. O componente aleatório, também conhecido como ruído branco ou processo puramente randômico, é um processo estocástico que consiste de uma sequência de variáveis randômicas independente e identicamente distribuídas, ou seja, é a parcela que a parte determinística do modelo não consegue captar. Variância constante, correlação nula, distribuição normal e média zero são características dos componentes aleatórios (BOX, JENKINS & REINSEL, 1994).

2.2 Classificação da série temporal

As séries temporais podem ser classificadas como estacionárias ou não estacionárias. A estacionariedade de uma série temporal significa que os dados oscilam sobre uma média constante, independente do tempo, com a variância das flutuações permanecendo essencialmente a mesma. Em outras palavras, uma série temporal é estacionária quando suas características estatísticas, como média, variância e autocorrelação, são constantes ao longo do tempo.

Existem testes que comprovam a estacionariedade de uma série temporal, como teste de Dickey-Fuller, teste de Phillips-Perron e teste KPSS, porém uma inspeção visual já pode dar indícios sobre sua classificação.

2.3 Teste de Dickey-Fuller

O teste de Dickey-Fuller é o mais antigo e famoso método formal para verificar se uma série temporal é ou não estacionária. Esse teste foi introduzido em uma versão básica pelos trabalhos de Fuller (1976), Dickey (1976) e Dickey e Fuller (1979). Posteriormente, foi objeto de uma generalização no trabalho de Dickey e Fuller (1981). A versão generalizada é conhecida como teste de Dickey-Fuller Aumentado e constitui a modalidade deste teste que passou a ser usada desde então (FRANCO, REISEN & BARROS, 2004).

O desenvolvimento do teste de Dickey-Fuller foi motivado pela necessidade de verificar se uma série precisa ser diferenciada para se tornar estacionária. Essa verificação

é a primeira etapa da metodologia proposta por Box e Jenkins para construção de modelos Autorregressivos Integrados Médias Móveis (ARIMA) para séries temporais.

O primeiro aspecto a destacar sobre o teste Dickey-Fuller é que ele procura verificar se uma série é estacionária ou não testando se ela possui uma raiz unitária. Antes de analisar a fórmula da equação do teste de Dickey-Fuller, deve-se retornar para alguns conceitos e fórmulas de processos estocásticos e teste de Raiz Unitária (MARGARIDO & DE MEDEIROS JUNIOR, 2006).

Simplificando a fórmula do processo estacionário, de forma que $\psi_0 = \psi_1 = 0$, implicando em $a = b = 0$, tem-se a seguinte fórmula para o teste de Dickey-Fuller:

$$Y_t = \rho Y_{t-1} + u_t$$

As premissas utilizadas pelo teste de Dickey-Fuller são $H_0: \rho = 1$ e $H_1: \rho < 1$, onde H_0 é não estacionária. Com isso, o teste Dickey-Fuller nos permite chegar à conclusão de que uma série é não estacionária, no sentido de possuir uma ou mais raízes unitárias, ou estacionária.

2.4 Métodos de previsão

A maioria dos métodos de previsão se baseia na ideia de que observações passadas contém informações sobre o padrão de comportamento da série temporal e, que este padrão será preservado no tempo futuro. O propósito dos métodos é distinguir o padrão que possa estar contido nas observações e então usar esse padrão para prever valores futuros da série (MORETTIN, 1987).

Deve-se passar por algumas etapas até chegar ao processo de prever e estimar valores futuros e tomar decisões apropriadas. Conforme Morettin e Tolói (2006) e Kunst (2009), tem-se quatro objetivos principais quando se estuda séries temporais, que resumidamente são: descrição, modelagem, previsão e controle de processos. Inicia-se estudando os dados e descrevendo o padrão de comportamento da série, como tendência, ciclo e variações sazonais. Parte-se para a construção de um modelo que estime os parâmetros e que avalie a qualidade estatística dos ajustes. A partir disso, pode-se estimar valores futuros, variando de curto a longo prazo, com base nos dados históricos que se possui. E, para garantir a integridade das informações, deve-se monitorar o modelo para identificar às periodicidades na série e assegurar a relevância do modelo adotado.

Os métodos de modelagem de séries temporais podem ser agrupados em três categorias, sendo elas os métodos de decomposição de séries temporais, métodos de alisamento e métodos Box-Jenkins.

2.5 Modelo ARIMA

O modelo ARIMA possui características flexíveis, permitindo ajustes e adaptações em seus parâmetros, contribuindo para que diferentes tipos de séries temporais sejam representadas, assumindo assim, uma estrutura de correlação linear entre os valores das séries temporais. Estes processos são descritos pelos modelos Autorregressivos, Integrados (número de diferenciações), e Médias Móveis de ordens p , d , e q , respectivamente, que podem ainda, ter incluído um operador sazonal, representado então por SARIMA(p , d , q)x(P , D , Q) (MAKRIDAKIS, 1998). A fórmula do modelo ARIMA é dada por:

$$\varphi(B)(1 - B)^d Z_t = \theta(B)\varepsilon_t$$

Onde $\varphi(B)$ é o polinômio autorregressivo, $\theta(B)$ é o polinômio de médias móveis, ε_t é um ruído branco com média zero e variância σ^2 e d é a ordem de integração, isto é, o número de diferenças aplicadas à série para que esta se torne estacionária.

Quando adicionamos a parte sazonal no modelo ARIMA, passamos a ter o modelo SARIMA, com a fórmula:

$$\varphi(B)\Phi(B)(1 - B)^d(1 - B^s)^D Z_t = \theta(B)\Theta(B)\varepsilon_t$$

Onde (p, d, q) são as ordens do modelo referentes à dinâmica ordinal e (P, D, Q) são às ordens da parte sazonal.

2.6 Identificação e seleção de um modelo

A identificação particular de um modelo ARIMA a ser ajustado aos dados pode ser considerado uma das fases mais críticas ao se utilizar uma modelagem ARIMA. A escolha do modelo a ser utilizado é feita principalmente com base nas Funções de Autocorrelações (FAC) e Autocorrelações Parciais (FACP) estimadas (WERNER & RIBEIRO, 2003).

A análise destas funções permite a tomada de decisão acerca do número p de parâmetros auto-regressivos e do número q de parâmetros de médias móveis que devem

ser escolhidos de maneira a se obter uma estrutura a que tenha o menor número de parâmetros dentre todas as estruturas que se ajustem aos dados da série.

A função de autocorrelação (FAC) é dada por:

$$\rho_k = \frac{\gamma_k}{\gamma_0}; \quad k = 0, 1, 2, \dots$$

$$\gamma_0 = \text{var}(z_t)$$

A função de autocorrelação parcial (FACP) é dada por:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}$$

$$\phi(B) \cdot \rho_k = 0$$

Com o intuito de comparar e selecionar o melhor modelo, utilizamos do critério de seleção de Akaike (AIC) (AKAIKE, 1974). O AIC é um critério que avalia a qualidade do ajuste do modelo paramétrico, estimado pelo método da máxima verossimilhança.

Akaike encontrou uma relação entre a esperança relativa da K-L (Kullback-Leibler) informação e a função suporte maximizada, permitindo uma maior interação entre a prática e a teoria, em seleção de modelos e análises de conjuntos de dados complexos (BURNHAM & ANDERSON, 2002).

A fórmula do AIC é dada por:

$$AIC = -2 \log L(\hat{\theta}) + 2(k)$$

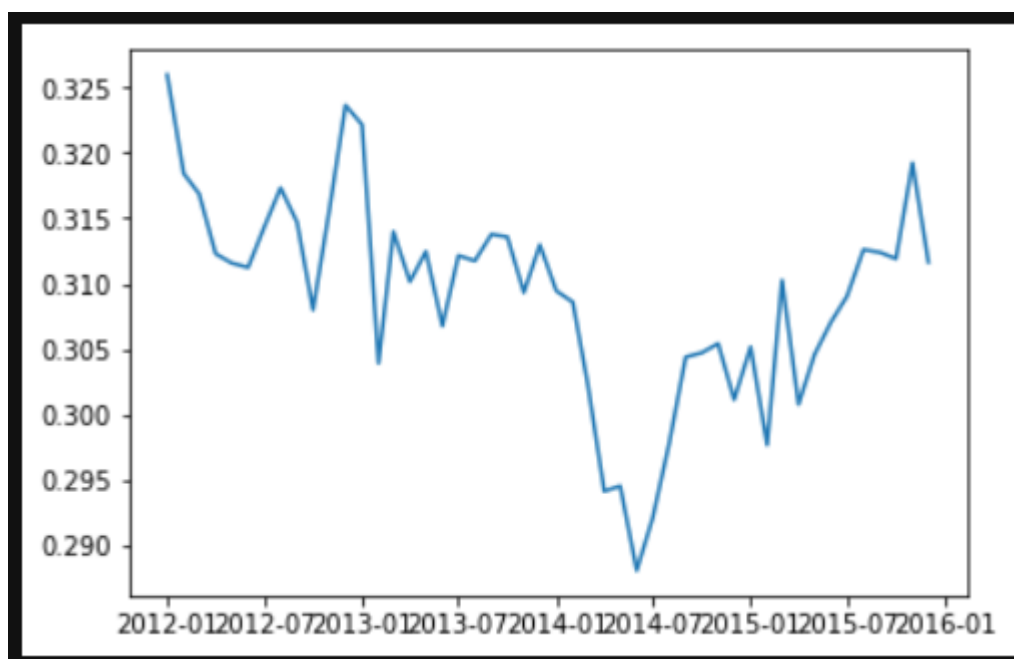
Onde $\log L(\theta)$ é a função de log-verossimilhança maximizada e k é o número de parâmetros do modelo. O termo $-2 \log L(\theta)$ é uma medida da qualidade do ajuste e $2k$ é a penalização que ocorre com a inclusão de parâmetros no modelo. Com esse critério, estimam-se vários modelos concorrentes e aquele com menor valor para o critério de informação é selecionado.

O AIC não é uma prova sobre o modelo, no sentido de testar hipóteses, mas uma ferramenta para a seleção de modelos; não é um teste de hipóteses, não há significância e nem p-valor. Dado um conjunto de dados e vários modelos concorrentes, pode-se classificá-los de acordo com o seu AIC, com aqueles tendo os menores valores de AIC sendo os melhores (BURNHAM & ANDERSON, 2002).

3. Resultados

O primeiro passo é entender os dados os quais estão sendo trabalhados, por isso foi plotado o Churn no período de tempo de janeiro de 2012 até dezembro de 2015. De acordo com a Figura 1, pode-se ver que os dados tendem a ser estacionários, porém, ainda serão realizados testes que comprovem essa suposição.

Figura 1: Informação visual dos dados ao longo do tempo.



O Quadro 1 apresenta as estatísticas da série. Nele, são observados a quantidade de valores disponíveis (66), média (0,308), desvio-padrão (0,009), assimetria (negativa), curtose (leptocúrtica) e teste de distribuição Shapiro-Wilk (que não rejeita a hipótese nula de normalidade dos dados, com p-valor igual a 0,37).

Quadro 1: Estatísticas da série temporal.

Estatística	Valor	p-valor
Tamanho amostral	66	-
Média	0,308	-
Desvio-padrão	0,009	-
Assimetria	-0,167	-
Curtose	2,569	-
Teste de Shapiro-Wilk	0,980	0,370

3.1 Validação dos dados estacionários

Após verificação visual que os dados tendem a ser estacionários, aplica-se o teste Dickey-Fuller para comprovação, conforme o Quadro 2.

Quadro 2: Resultados do teste de Dickey-Fuller.

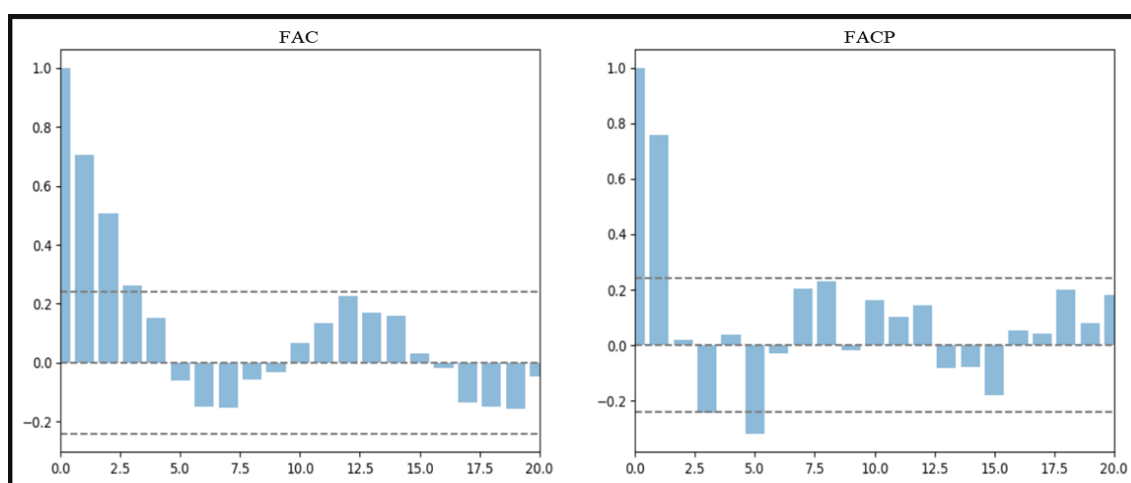
Estatística de Teste	p-valor	Lags Usados	Número de Observações	Valor Crítico (5%)
-3,2915	0,0152	4	61	-3,5424

Além de verificar que a média e o desvio padrão são constantes, constata-se também que o p-valor possui valor menor que o nível de significância adotado (5%). Deste modo, pode-se inferir estatisticamente que os dados são estacionários a um nível de 95% de confiança.

3.2 Aplicação da FAC e FACP

Nessa etapa são definidos os valores p , d e q do modelo ARIMA através das funções FAC e FACP. Após rodarmos a consulta, obteve-se o resultado conforme a Figura 3.

Figura 3: Resultados dos modelos FAC e FACP.



O intuito da FAC e da FACP é definir os valores dos parâmetros que irão ser trabalhados, porém, este estudo será feito de outra forma: a utilização dos valores encontrados serão definidos para o limite máximo de combinações. Sendo assim, serão geradas todas as combinações possíveis $(p, d, q) \times (P, D, Q)$ com menor valor ou igual a dois para todos estes parâmetros.

3.3 Aplicação do modelo ARIMA testando sazonalidade

Nessa etapa se calcula o Critério de Informação de Akaike (AIC) para todos os modelos ARIMA de combinações com valor máximo de $(p, d, q) \times (P, D, Q)$ igual a dois. Dado o cálculo para todas as combinações, e considerando o menor valor de AIC para a

escolha, após ordenados todos os modelos, foi selecionado o ARIMA (1,0,0) x (0,0,0). Deste modo, o melhor resultado obtido foi o AR(1), ou seja, Modelo Autorregressivo com parâmetro $p = 1$.

Rodou-se este modelo para os dados de treino, e foi verificado que o mesmo é significativo ao nível de 5%, como pode ser visto no Quadro 3.

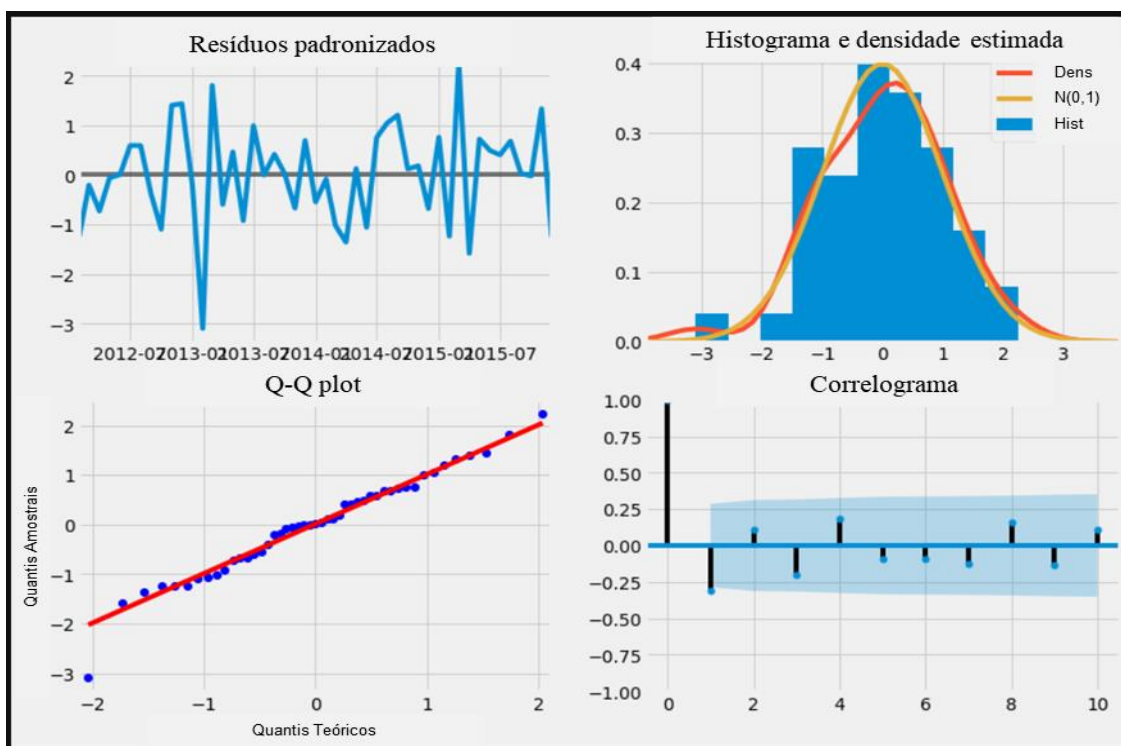
Quadro 3: Resultados do modelo ARIMA (1,0,0) x (0,0,0).

	Coefficiente	Desvio Padrão	Valor z	p-valor	[0,025	0,975]
AR(1)	0,9988	0,003	358,618	0,000	0,933	1,004
σ^2	$3,3e^{-5}$	$6,2e^{-6}$	5,343	0,000	$2,1e^{-5}$	$4,53e^{-5}$

3.4 Diagnóstico do modelo ARIMA escolhido

Após a escolha, foi realizado a parte de diagnóstico do modelo selecionado, como pode ser observado na Figura 4.

Figura 4: Conjunto de diagnósticos do modelo ARIMA (1,0,0) x (0,0,0).



As principais conclusões foram:

- I. Gráfico de análise dos resíduos padronizados: a média de erros está bem próxima de uma constante (zero);

- II. Gráfico Q-Q plot: a distribuição do modelo está bem próxima da distribuição dos dados;
- III. Histograma e densidade estimada: os erros estão seguindo uma distribuição normal;
- IV. Correlograma: pode-se verificar uma baixa autocorrelação, que é um bom resultado para o modelo o qual está sendo aplicado (modelo autorregressivo).

O Quadro 4 revela os testes de autocorrelação Box-Pierce e Ljung-Box (p-valor = $2,2e^{-16}$) e o teste de heterocedasticidade Breusch-Pagan (p-valor = 0,287). Com esses resultados, rejeita-se a hipótese de que os resíduos são identicamente e independentemente distribuídos e não se rejeita a hipótese de homoscedasticidade dos erros para os resíduos do modelo.

Quadro 4: Testes de autocorrelação e heterocedasticidade.

Teste	Estatística de teste	p-valor
Box-Pierce	24,256	$2,2e^{-16}$
Ljung-Box	22,800	$2,2e^{-16}$
Breusch-Pagan	1,135	0,287

Para os coeficientes de resíduos do modelo, os dados se revelaram significantes ao nível de 5%, conforme os resultados do Quadro 5.

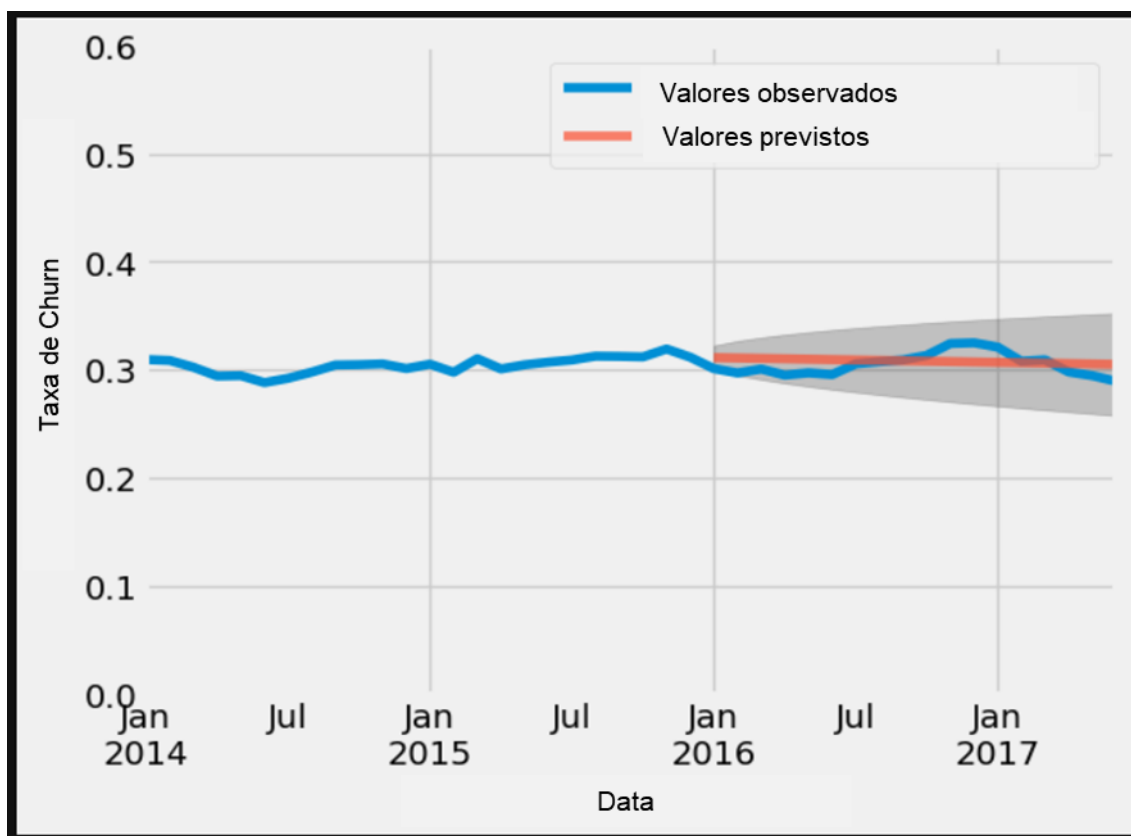
Quadro 5: Resultados para os coeficientes de resíduos do modelo.

	Coefficiente	Desvio Padrão	Valor t	p-valor
Intercepto	0,309	0,0008	352,945	$2,2e^{-16}$
Resíduos	0,959	0,1618	5,923	$3,778e^{-7}$

3.5 Modelo preditivo

Com o modelo escolhido e validado, o próximo passo é gerar 18 valores de *churn* para o período de janeiro de 2016 até junho de 2017. A Figura 5 apresenta os valores previstos e comparados com os valores reais observados, elaboradas um passo a frente pelo método da máxima verossimilhança.

Figura 5: Projeção dos dados um passo a frente para 18 meses.



Para validação dos valores próximos da realidade, o Erro Quadrado Médio (EQM) obtido foi de 0,00013.

5. Considerações finais

O objetivo desse estudo foi analisar estatisticamente o percentual de clientes que deixam de comprar ao longo de 12 meses (taxa de *Churn*), dentro de uma empresa específica do ramo de *e-commerce*, podendo dessa forma fazer projeções mais assertivas de venda e custo de retenção de clientes.

A metodologia estatística de séries temporais utilizada foi a modelagem ARIMA, que envolveu as análises e os testes utilizados, elaborando todo o ciclo iterativo - especificação, identificação, estimação e diagnóstico (verificação). O software utilizado foi o JupyterLab, que roda as informações e testa os modelos de forma online.

Por fim, ao realizar as previsões através da base de treino e comparar com os dados de teste, foi possível concluir que o método AR (1) obteve o melhor resultado para realizar a projeção mensal. Sendo assim, pode-se inferir que este resultado se aproxima de um modelo naive, que leva em consideração apenas o último período da série, ou seja, o

mesmo pressupõe que o futuro repetirá o passado, logo a previsão é correspondente ao último valor observado.

Como proposta para futuros trabalhos, sugere-se a utilização das técnicas de alisamento exponencial e aprendizado de máquina para previsão das séries temporais.

Referências

- AKAIKE, H. **Information theory and an extension of the maximum likelihood principle**. Proc. 2a Inter. Symposium on Information Theory, 267-281, Budapest, 1974.
- BARROS, M. **Processo Estocástico**. São Paulo: Publit, 2005.
- BOX, G.E.P.; JENKINS, G.M. & REINSEL G.C. **Time Series Analysis: Forecasting and Control**. 3 Ed. Prentice Hall, 1994.
- BURNHAM, K. P.; ANDERSON, D. R. **Model Selection and Multimodel Inference**. 2002.
- DICKEY, D.A. & W.A. FULLER Distribution of the Estimators for Autoregressive **Time Series with Unit Root**. **Journal of the American Statistical Association**, 74, 427- 431, 1979.
- DOWNING, D; CLARK, J. **Estatística Aplicada**. 2 ed. São Paulo: Saraiva, 2006.
- FRANCO, G. C.; REISEN, V. A.; BARROS, P. A. **Testes de raízes unitárias através de estimadores semiparamétricos do parâmetro de memória longa**. O Impacto da Pesquisa Operacional nas Novas Tendências Multidisciplinares. XXXVI SBPO. São João Del-Rei, 2004.
- JUPYTER. **JupyterLab Documentation**. 2020. Disponível em: <https://jupyterlab.readthedocs.io/en/stable/#>.
- KUNST, R.M. **Applied Time Series Analysis - Part I**. University of Vienna an Institute for Advanced Studies Vienna, 2009.
- KURTZ, D. L; CLOW, K. E. **Services marketing**. New York: John Wiley & Sons, 1998.
- MAKRDAKIS, S.; WHEELWRIGHT, S.; HYNDMAN, R.J. **Forecasting methods and applications**. 3 Ed. New York; John Wiley e Sons, 1998.
- MARGARIDO, M. A.; DE MEDEIROS JUNIOR, H. Teste para mais de uma raiz unitária: uso do software SAS® na elaboração de uma rotina para o teste dickey-pantula. **Pesquisa & Debate**. Revista do Programa de Estudos Pós-Graduados em Economia Política, v. 17, n. 1 (29), 2006.
- MORETTIN, P. A.; TOLOI, C. M. **Previsão de séries temporais**. 2 ed. São Paulo: Atual, 1987.
- MORETTIN, P. A.; TOLOI, C. M. **Análise de Séries Temporais**. 2 ed. São Paulo: Atual, 2006.
- WERNER, L.; RIBEIRO, J. L. D. Previsão de demanda: uma aplicação dos modelos Box-Jenkins na área de assistência técnica de computadores pessoais. **Gestão e Produção**. São Carlos, SP. Vol. 10, n. 1 (abr. 2003), p. 47-67, 2003.

USE OF ARIMA MODELS FOR CHURN RATE FORECASTING: CASE STUDY FOR AN E-COMMERCE COMPANY

Abstract

The Churn rate, or simply Churn, calculates the number of users who disconnect from a company's services in a specific period of time. For some sectors, this is a basic metric for assessing business success, as it has a direct impact on revenue. In this work, the Churn of an e-commerce company is projected in the short term based on the history of its data. For that, the time series are used to forecast these data, the Integrated Autorregressive Moving Average Model (ARIMA). The work went through all the stages of the iterative cycle of a data forecasting process, starting with the study and analysis of the database, going through the choice and validation of the model parameters until the data projection was reached. The Dickey-Fuller test showed that the series is stationary, the best model found was AR (1) and the residues follow a normal distribution.

Keywords: *Churn Rate, Time Series, ARIMA, Data Forecast.*