

19 a 21 de outubro Ponta Grossa - PR - Brasil

ORDENAÇÃO DE ALGORITMOS PARA MODELAGEM PREDITIVA DE CHURN: ANALISANDO O PROBLEMA A PARTIR DOS MÉTODOS SAPEVO-M E VIKOR

ORDERING ALGORITHMS FOR CUSTOMER CHURN PREDICTIVE MODELING: ANALYZING THE PROBLEM USING SAPEVO-M AND VIKOR METHODS

ÁREA TEMÁTICA: ADMINISTRAÇÃO DA INFORMAÇÃO

Antonio Carlos da Silva Júnior, Universidade Federal do Paraná, Brasil, acsjunior@ufpr.br

Isaque David Pereira de Almeida, Marinha do Brasil, Brasil, isaque.mestrado@gmail.com

Marcos dos Santos, Instituto Militar de Engenharia, Brasil, marcosdossantos_doutorado_uff@yahoo.com.br

Resumo

Modelagem preditiva para classificação de churn (abandono de cliente) é uma prática comum em empresas de diversos setores, no entanto, embora seja um tema vastamente explorado, escolher o classificador adequado pode ser uma tarefa árdua, dadas as particularidades de cada empresa e a variedade de algoritmos disponíveis. Portanto, a proposta deste artigo é a escolha de um algoritmo para modelagem preditiva de churn em uma startup brasileira. O processo decisório fez uso do método SAPEVO-M para obter os pesos dos critérios e definir uma medida de interpretabilidade para os algoritmos avaliados, e o método VIKOR foi utilizado para avaliar as alternativas. Após a aplicação dos métodos, os classificadores treinados com os algoritmos SVM (kernel radial) e Regressão Logística foram considerados os mais adequados para o modelo de negócio da empresa em questão.

Palavras-chave: SAPEVO-M; VIKOR; Tomada de decisão; Modelagem preditiva; Churn.

Abstract

Predictive modeling for customer churn classification is a common practice among companies from different sectors, however even though it is a widely explored subject, choosing a suitable classifier may be a difficult task, given the particularities of each company and the variety of available algorithms. Therefore, the purpose of this article is to choice of an algorithm for customer churn predictive modeling in a Brazilian startup. The decision-making process used the SAPEVO-M method to obtain the weights of criteria and to define an interpretability measure for the evaluated algorithms, and the VIKOR method was used to evaluate the alternatives. After applying the methods, the classifiers trained with SVM (radial kernel) and Logistic Regression algorithms were considered the most suitable for the business model of the company in question.

Keywords: SAPEVO-M; VIKOR; Tomada de decisão; Modelagem preditiva; Churn.

1. INTRODUÇÃO

Investir em estratégias para retenção de clientes é uma prática comum entre empresas dos mais variados setores. De acordo com Hennig-Thurau (2004), o relacionamento duradouro com clientes é fundamental para a sobrevivência econômica de uma empresa de serviços, uma

vez que o custo para adquirir novos clientes costuma ser mais alto, dependendo da competitividade do mercado. Para Pfeifer (2005), mesmo não havendo um acordo com relação à razão dos custos para atração e retenção de clientes, há um consenso geral quanto ao investimento de atenção e dinheiro na retenção de clientes.

Identificar antecipadamente quais clientes podem abandonar a empresa pode ser fundamental para elevar a taxa de retenção e, para a obtenção de tal informação, modelos preditivos são comumente utilizados. Segundo Kuhn e Johnson (2013), modelagem preditiva é o processo de criação de um modelo matemático para tentar prever um resultado, podendo esta técnica ser aplicada em diversas áreas.

A predição de abandono de clientes, também conhecida como predição de churn, tem sido um assunto vastamente explorado nos últimos anos, no entanto, diante das particularidades de cada empresa e da quantidade de métodos existentes, escolher um algoritmo para resolver o problema em questão pode ser uma tarefa bastante demorada, já que as comparações normalmente são realizadas de forma manual através de métricas de avaliação.

Segundo Hossin e Sulaiman (2015), as métricas de avaliação são o ponto chave para a escolha do algoritmo ideal e, de acordo com Kumar (2012), não existe um algoritmo que seja ideal em todos os cenários, portanto, entende-se que a escolha do melhor algoritmo é um ponto crítico durante o desenvolvimento da solução, uma vez que uma escolha inadequada poderá acarretar resultados indesejados.

Com base nos dados de uma startup brasileira que comercializa uma plataforma digital para vendas on-line, os autores deste artigo propõem um estudo para definir, entre os algoritmos com melhor desempenho, qual o modelo de classificação binária ideal para realizar a predição de churn.

A melhor forma para a escolha de um modelo preditivo de churn, para este trabalho ao ver dos autores, foi a utilização dos métodos VIKOR e SAPEVO-M, uma vez que, para o apoio à tomada de decisão, foram levantadas algumas alternativas e critérios, tendo a necessidade de utilização de dados qualitativos e quantitativos.

2. ESTRUTURAÇÃO DO PROBLEMA

Simon (1973) define um problema mal estruturado como um problema cuja composição, em algum aspecto, carece de definição. Seguindo um pensamento parecido, Kitchner (1983) relaciona um problema mal estruturado com a existência de evidências e opiniões conflitantes que podem levá-lo a múltiplas soluções.

Os métodos de estruturação de problemas (PSM – *Problem Structuring Methods*) é uma das etapas do processo de tomada de decisão que tem como objetivo organizar assuntos, questões e/ou dilemas para os quais se buscam proposições de decisões, inicialmente; e não resolvê-los a priori. Uma das principais características dos PSM a participação dos envolvidos (stakeholders) no processo de investigação da situação-problema e na definição das ações a serem realizadas para melhorá-la. A filosofia por trás dos PSM assume que cada stakeholder tem uma percepção particular da situação-problema, e essa percepção varia de acordo com os interesses, visões de mundo e experiências de vida de cada um (Bandeira, Mattos, Neyra Belderrain, Correia, & Kleba, 2018).

Segundo Bandeira et al. (2018), existe um propósito para a utilização desses métodos que é diminuir ao máximo as incertezas, os conflitos e a complexidade. De uma forma geral o objetivo do PSM é ordenar todo um contexto, mesmo que complexo, com o propósito de apresentar o problema àqueles que irão decidir. Para Mingers e Rosenhead (2004), todo PSM possui uma

forma distinta de apresentar a situação, permitindo o esclarecimento aos participantes, assim como auxiliando no direcionamento de uma possível solução.

De acordo com Mingers e Rosenhead (2004), um dos métodos que compõem o PSM mais utilizados em diversas áreas de aplicação é a Metodologia de Sistemas Suaves (SSM - *Soft Systems Methodology*). Através do redesenho do sistema, cada participante, dentro das suas experiências e visões do mundo, constrói modelos conceituais e ideias, possibilitando as comparações com as percepções do sistema existente, tendo como foco as mudanças desejáveis e viáveis. Segundo Abuabara, Paucar-Caceres e Burrowes-Cromwell (2019), a figura rica é uma ferramenta simplória da SSM voltada para produzir discussões sobre os pontos principais que cada participante obteve do problema. Para Checkland (1999), a figura rica deve ser produzida sem auxílio de recursos computacionais e sem formas estruturadas, possibilitando a real expressão de tudo o quanto fora entendido pelos participantes, a fim de demonstrar o entendimento do problema. Para isso são necessárias duas etapas:

- Pesquisar um problema sem estrutura; e
- Confeccionar uma figura rica, a fim de expressá-lo.

Com a intenção de explorar uma situação problemática mal estruturada, os autores, através da técnica de brainstorm, demonstraram o que compreenderam sem a interposição de ideias contrárias ao que fora apresentado por cada integrante e expressarem o entendimento através dos aspectos principais do problema, foi confeccionada uma figura rica como na Figura 1.

A figura rica exhibe algumas opções de algoritmos para análise preditiva que podem ser utilizados para classificar a base de clientes, a fim de identificar quais deles podem ou não abandonar o serviço contratado, em outras palavras, para realizar a predição do churn. A figura também exhibe algumas métricas para comparação dos modelos preditivos, que inclusive podem ser influenciadas pelas características dos dados, tais como o desbalanceamento da variável resposta e a existência de valores atípicos (outliers). Outro ponto representado é o custo computacional de cada algoritmo.

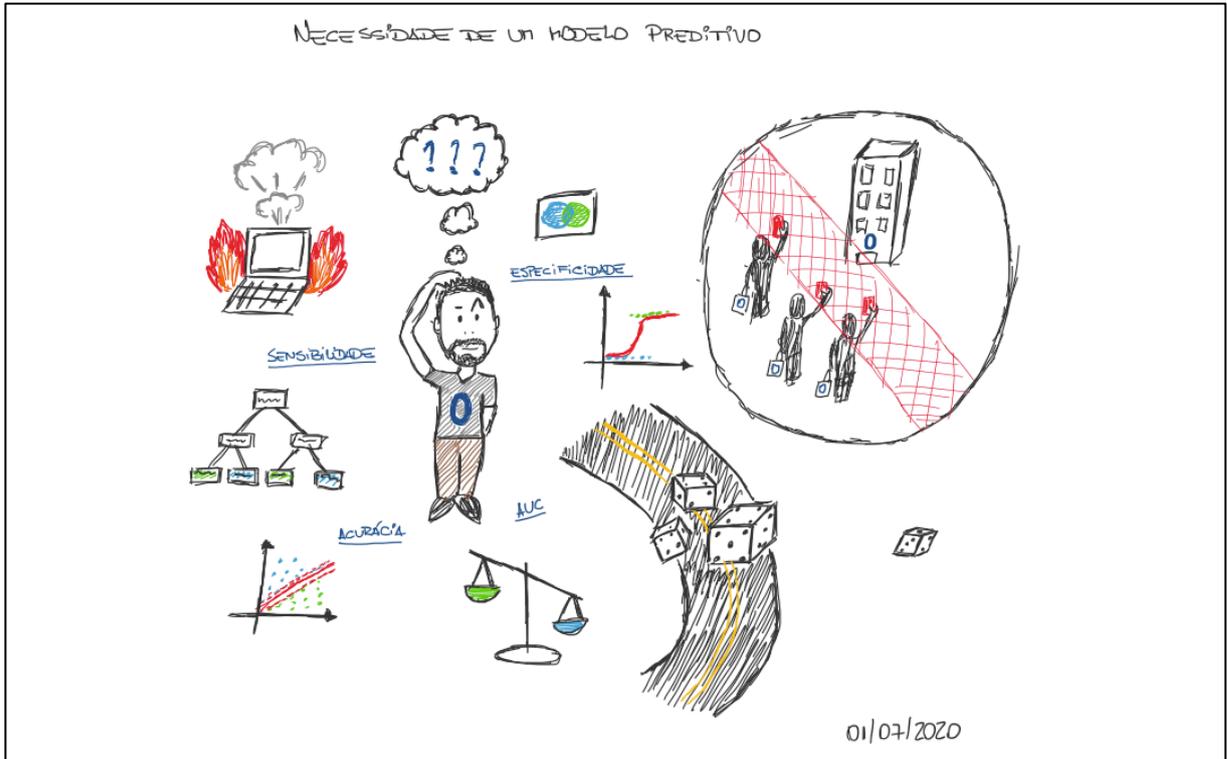


Figura 1 – Figura rica, intitulada “NECESSIDADE DE UM MODELO PREDITIVO”

Após análise da figura rica foram estabelecidos quatro critérios: Sensibilidade, Especificidade, Eficiência e Interpretabilidade. A Sensibilidade e a Especificidade foram escolhidas por serem métricas que possibilitam aos decisores analisar individualmente as classificações positivas (detecção do churn) e negativas, além de serem comumente utilizadas para avaliação de modelos de classificação binária. Já a Eficiência e a Interpretabilidade foram selecionadas com base na opinião de profissionais com vasta experiência em modelagem preditiva e estratégias de retenção de clientes.

Experimentos individuais foram realizados com diversos algoritmos através do software R, com o intuito de obter suas versões mais adequadas para o negócio e, após a análise dos critérios e eliminação dos modelos que claramente apresentaram péssimos resultados, restaram as seguintes alternativas: Árvore de Decisão (CART), KNN, LDA, Random Forest, Regressão Logística e SVM (kernel radial).

2.1 Apresentação dos critérios

Os critérios escolhidos para a escolha do classificador binário foram:

a) Sensibilidade: A matriz de confusão (quadro 1) é construída através do cruzamento dos valores preditos pelo modelo e os valores observados, produzindo em sua diagonal principal o número de detecções corretas e, na diagonal secundária, o número de detecções incorretas.

	OBSERVADO (+)	OBSERVADO (-)
Predito (+)	Verdadeiro positivo (VN)	Falso positivo (FP)
Predito (-)	Falso negativo (FN)	Verdadeiro negativo (VN)

Quadro 1 – Matriz de confusão

A sensibilidade pode ser definida como a taxa de detecção correta entre os clientes que abandonaram a empresa. Segundo Verbeke, Dejaeger, Martens, Hur e Baesens (2012), é a proporção dos exemplos positivos preditos como positivo pelo modelo, e é calculada através da equação 1:

$$\text{Sensibilidade} = \frac{VP}{VP+FN} \quad (1)$$

b) Especificidade: Pode ser definida como a taxa de detecção correta entre os clientes que não abandonaram a empresa. De acordo com Verbeke et al. (2012) é a proporção dos exemplos negativos classificados como negativo pelo modelo e é calculada através da equação 2:

$$\text{Especificidade} = \frac{VN}{VN+FP} \quad (2)$$

c) Eficiência: É o tempo necessário para realizar o treino do algoritmo. Diante da existência de fatores não controlados capazes de influenciar o tempo da execução do algoritmo, a eficiência foi calculada extraindo a média aritmética do tempo, em segundos, de cinquenta execuções.

d) Interpretabilidade: É a capacidade do algoritmo de fornecer uma resposta interpretável. De acordo com Martens, Vanthienen, Verbeke e Baesens (2011) e Miller (2019), a interpretabilidade é o quão bem o usuário é capaz de compreender a lógica por trás da predição do modelo e segundo Verbeke et al. (2012), um classificador interpretável torna possível a extração de informações importantes do modelo para apoiar o desenvolvimento de estratégias e campanhas de retenção dos clientes.

2.2 Apresentação das alternativas

Os algoritmos escolhidos para compor as alternativas da análise multicritério foram:

- a) baseado em árvore:** Árvore de Decisão (CART);
- b) linear / análise discriminante:** Regressão Logística e LDA;
- c) não lineares / ensemble:** KNN, SVM (kernel radial) e Random Forest.

3. REFERENCIAL TEÓRICO DOS MÉTODOS

Foram utilizados dois métodos, SAPEVO-M e VIKOR, como uma ferramenta híbrida para apoio à tomada de decisão.

3.1 O método SAPEVO-M

Teixeira, Santos e Gomes (2019) apresentaram o método SAPEVO-M (*Simple Aggregation of Preferences Expressed by Ordinal Vectors – Multi Decision Makers*) que representa um aprimoramento da versão original do SAPEVO, gerado inicialmente por Gomes, Mury e Gomes (1997). Esta transformação possibilitou a aplicação de múltiplos decisores.

O sistema SapevoWeb foi desenvolvido a partir de uma parceria entre o corpo técnico do Centro de Análise de Sistemas Navais (CASNAV), um grupo de pesquisa do Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal Fluminense (UFF) e um grupo de pesquisa do Programa de Pós-Graduação em Engenharia de Sistemas e Computação do Instituto Militar de Engenharia (IME) (Teixeira et al., 2019). Segundo Teixeira et al. (2019), o código do SapevoWeb foi desenvolvido na linguagem Python, pelo uso do framework Django, e HTML. O programa possibilita a inclusão de um grande número de decisores, critérios e alternativas utilizando de valores qualitativos e pode ser acessado em www.sapevoweb.com.

3.2 O método VIKOR

O método ViseKriterijumska Optimizacija i Kompromisno Resenje (VIKOR) foi desenvolvido como um método de programação de compromisso, com o objetivo de determinar um ranking com pesos de um conjunto de alternativas (Duckstein & Opricovic, 1980).

1º passo: Através da equação 3, com $i = 1, 2, \dots, m$ e $j = 1, 2, \dots, n$, deve-se normalizar a matriz de decisão.

$$f_{ij}(x) = \frac{x_{ij}}{\sqrt{\sum_{i=1}^m x_{ij}^2}} \quad (3)$$

2º passo: Deve-se determinar o melhor (f_i^*) e o pior valor (f_i^-) apresentado pelas alternativas em cada critério. Se o critério for de lucro, considera-se as equações $f_i^* = \text{Max}_i f_{ij}$ e $f_i^- = \text{Min}_i f_{ij}$, caso contrário, as equações $f_i^* = \text{Min}_i f_{ij}$ e $f_i^- = \text{Max}_i f_{ij}$ devem ser consideradas, com $i = 1, 2, \dots, m$ e $j = 1, 2, \dots, n$.

3º passo: Deve-se calcular o grupo de utilidades (S_i) e o arrependimento individual (R_i), através das equações 4 e 5, respectivamente, onde $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$ e w_j corresponde ao peso de cada critério.

$$S_i = \sum_{j=1}^n w_j \frac{(f_j^* - f_{ij})}{(f_j^* - f_j^-)} \quad (4)$$

$$R_i = \text{Max}_j \left[w_j \frac{(f_j^* - f_{ij})}{(f_j^* - f_j^-)} \right] \quad (5)$$

4º passo: Deve-se calcular o índice de cada alternativa (Q_i) através da equação 6, com $i = 1, 2, \dots, m$ e $j = 1, 2, \dots, n$.

$$Q_i = \gamma \frac{(S_i - S^*)}{(S^- - S^*)} + (1 - \gamma) \frac{(R_i - R^*)}{(R^- - R^*)} \quad (6)$$

Onde $S^* = \text{Min}_i S_i$, $S^- = \text{Max}_i S_i$, $R^* = \text{Min}_i R_i$, $R^- = \text{Max}_i R_i$ e sendo γ o peso de estratégia, correspondendo ao valor de 0,5.

5º passo: Deve-se realizar 3 classificações decrescentes das alternativas, sendo S_i , R_i e Q_i os valores considerados em cada um dos rankings.

6º passo: Inicia-se a verificação das condições de estabilidade C1 e C2.

- C1 (vantagem aceitável): $Q(A_2) - Q(A_1) \geq 1/(m - 1)$, onde A_1 e A_2 são as alternativas com o menor se o segundo menor valor de Q , respectivamente, e m é o número de alternativas; e
- C2 (estabilidade aceitável na tomada de decisão): A_1 deve estar melhor classificada em S e/ou R .

Em seguida, é proposto um conjunto de soluções de compromisso com base na análise das condições de estabilidade.

- Solução I: Se C1 não for satisfeita, são selecionadas as alternativas A_1, A_2, \dots, A_M , sendo A_M determinada por $Q(A_M) - Q(A_1) < 1/(m - 1)$;
- Solução II: Se C2 não for satisfeita, são selecionadas as alternativas A_1 e A_2 ; e
- Solução III: Se C1 e C2 forem satisfeitas, é selecionada a alternativa com o menor valor de Q .

4. APLICAÇÃO DO MÉTODO SAPEVO-M

Na primeira etapa do problema foi aplicado o método SAPEVO-M para obtenção dos pesos dos critérios e, em seguida, para se obter os valores do critério Interpretabilidade, por meio do software SapevoWeb, Teixeira, Santos e Gomes (2018).

4.1 Inclusão dos decisores, alternativas e critérios

Tanto para obtenção dos pesos dos critérios como para estabelecer os valores do critério Interpretabilidade, a aplicação do SAPEVO-M foi constituída em três partes. A primeira foi a inclusão dos decisores (figura 2).

Projeto Modelo de Churn

Cadastrar Decisor

Nome:

Decisores

Decisor 1

Decisor 2

Decisor 3

CASNAV IME UFF

Figura 2 – Tela de inclusão dos decisores na plataforma SapevoWeb

Na sequência a inclusão dos algoritmos como alternativa (figura 3).

Projeto Modelo de Churn

Cadastrar Alternativas

Nome:

Alternativas

Alternativa 1

Alternativa 2

Alternativa 3

CASNAV IME UFF

Figura 3 – Tela de inclusão das alternativas na plataforma SapevoWeb

E, por fim, a inclusão dos critérios (figura 4). A aplicação para obtenção dos valores da interpretabilidade considerou, na plataforma, cada algoritmo como critério.

Projeto Modelo de Churn

Cadastrar Critério

Nome:

Critérios

- Critério 1
- Critério 2
- Critério 3

CASNAV IME UFF

Figura 4 – Tela de inclusão dos critérios na plataforma SapevoWeb

4.2 Avaliação dos critérios

Uma vez realizada a inclusão dos decisores, alternativas e critérios, coube aos especialistas e stakeholders, cadastrados como decisores, a avaliação comparativa de cada critério (figura 5).

Projeto Modelo de Churn

Avaliar Critérios

Qual é o decisor? Decisor 1

Critério 1 VS Critério 2

Absolutamente Pior
 Muito Pior
 Pior
 Equivalente
 Melhor
 Muito melhor
 Absolutamente melhor

Critério 1 VS Critério 3

Absolutamente Pior
 Muito Pior
 Pior
 Equivalente
 Melhor
 Muito melhor
 Absolutamente melhor

Critério 1 VS Critério 4

Absolutamente Pior
 Muito Pior
 Pior
 Equivalente
 Melhor
 Muito melhor
 Absolutamente melhor

Critério 2 VS Critério 3

Absolutamente Pior
 Muito Pior
 Pior
 Equivalente
 Melhor
 Muito melhor
 Absolutamente melhor

Critério 2 VS Critério 4

Absolutamente Pior
 Muito Pior
 Pior
 Equivalente
 Melhor
 Muito melhor
 Absolutamente melhor

Critério 3 VS Critério 4

Absolutamente Pior
 Muito Pior
 Pior
 Equivalente
 Melhor
 Muito melhor
 Absolutamente melhor

Avaliar





Figura 5 – Tela de avaliação dos critérios na plataforma SapevoWeb

4.3 Obtenção dos pesos dos critérios e da medida de interpretabilidade

Aplicado o método SAPEVO-M, conforme os itens 4.1 e 4.2, a relação dos pesos dos critérios e os valores do critério Interpretabilidade foram estabelecidos, conforme as tabelas 2 e 3, respectivamente.

CRITÉRIO	PESO
Sensibilidade	5,6667
Especificidade	1,1821
Eficiência	1,7617
Interpretabilidade	4,4389

Tabela 1 – Relação dos pesos dos critérios

ALTERNATIVA	INTERPRETABILIDADE
Árvore de Decisão (CART)	3,0000
KNN	0,5375
LDA	1,9829
Random Forest	0,0106
Regressão Logística	2,7507
SVM (kernel radial)	1,0076

Tabela 2 – Medida de interpretabilidade dos algoritmos

Segundo Teixeira et al. (2019), a aplicação do software SapevoWeb traz como resultado uma ordenação, em valores, das alternativas e/ou critérios fazendo com que haja uma informação adicional, auxiliando assim os tomadores de decisão. Como Exemplo, sabe-se que, em relação aos pesos, o critério sensibilidade (5,6667) foi considerado aproximadamente cinco vezes mais importante que o critério especificidade (1,1821).

5. APLICAÇÃO DO MÉTODO VIKOR

Na segunda etapa do problema foi utilizado o método VIKOR para obter a classificação dos algoritmos e a verificação das condições de estabilidade. A aplicação foi realizada através do software Microsoft Excel, com os cálculos verificados através do software VIKOR, OnlineOutput (2017).

5.1 Obtenção da matriz de decisão

Aplicados os vetores de pesos (tabela 1), os valores da Interpretabilidade (tabela 2) e os valores obtidos através das simulações de cada algoritmo no software R, foi obtida a matriz de decisão (quadro 2). Em seguida foram determinados os valores de f_i^* e f_i^- , que correspondem ao melhor e pior valor de cada critério, respectivamente, conforme descreve o passo do 1º item 3.2.

CUSTO / LUCRO PESOS	LUCRO	LUCRO	CUSTO	CUSTO
	5,6667	1,1821	1,7617	4,4389
	SENSIBILIDADE	ESPECIFICIDADE	EFICIÊNCIA	INTERPRETABILIDADE
Árvore de Decisão	0,7869	0,8176	0,4345	3,0000
KNN	0,8577	0,7756	0,5717	0,5375
LDA	0,8242	0,7766	0,0872	1,9829
Random Forest	0,9315	0,8849	12,7759	0,0106
Regressão Logística	0,8151	0,8107	0,1159	2,7507
SVM	0,8820	0,8205	5,0794	1,0076

Quadro 2 – Matriz de decisão com os pesos dos critérios

5.2 Obtenção dos rankings de S, R e Q

Após normalizada a matriz de decisão através da equação 3, as alternativas foram classificadas em ordem decrescente com base nos valores de S_i , R_i e Q_i , obtidos por meio das equações 4, 5 e 6, respectivamente. Com os resultados foi constituída a tabela 3.

ALTERNATIVA	S	CLASS. (S)	R	CLASS. (R)	Q	CLASS. (Q)
Árvore de Decisão (CART)	6,4428	4	5,6667	6	0,6710	6
KNN	7,7980	6	3,6565	2	0,6289	5
LDA	6,8865	5	4,2050	3	0,5089	4
Random Forest	6,2006	2	4,4389	4	0,3856	3
Regressão Logística	5,7382	1	4,5616	5	0,2960	2
SVM (kernel radial)	6,2879	3	2,9585	1	0,1334	1

Tabela 3 – Rankings de S, R e Q

5.3 Análise dos resultados

Nesta etapa foi constatado que o algoritmo melhor classificado em Q (SVM) não possui vantagem aceitável com relação ao segundo colocado (Regressão Logística), entretanto o fato do SVM também estar melhor classificado em R , satisfaz a condição de estabilidade na tomada de decisão. Portanto, o fato de somente a segunda condição de estabilidade ter sido satisfeita, direcionou a conclusão da análise para a primeira solução de compromisso, trazendo como resultado a seleção das alternativas SVM e Regressão Logística.

6. CONCLUSÕES

Através do método SAPEVO-M foi possível estabelecer os pesos dos critérios para a escolha do classificador com base na avaliação de especialistas e stakeholders. Outra contribuição do método foi a definição de uma medida de interpretabilidade coerente com relação aos tipos de algoritmos utilizados.

Como resultado da aplicação do método VIKOR, duas alternativas foram selecionadas, sendo uma delas mais equilibrada com relação aos critérios avaliados e outra com maior intensidade no critério considerado o mais importante para os decisores.

Por fim, a combinação dos métodos VIKOR e SAPEVO-M proporcionou um processo decisório coeso, transparente e adaptado às necessidades da empresa, e permitiu aos decisores a escolha do modelo preditivo de churn adequado, fundamentada em critérios sólidos e relevantes.

REFERÊNCIAS

- Abuabara, L., Paucar-Caceres, A., & Burrowes-Cromwell, T. (2019). Consumers' values and behaviour in the Brazilian coffee-in-capsules market: promoting circular economy. *International Journal of Production Research*, 57(23), 7269–7288. <https://doi.org/10.1080/00207543.2019.1629664>
- Bandeira, M., Mattos, R., Neyra Belderrain, M., Correia, A., & Kleba, J. (2018). *Modelo de negócio em uma comunidade agrícola: Aplicação de Soft Systems Methodology e Strategic Choice Approach*.
- Checkland, P. (1999). *Soft systems methodology: a 30-year retrospective*. Chichester New York: John Wiley.
- Duckstein, L., & Opricovic, S. (1980). Multiobjective optimization in river basin development. *Water Resources Research*, 16(1), 14–20. <https://doi.org/10.1029/wr016i001p00014>
- Gomes, L. F. A. M., Mury, A. R., & Gomes, C. F. S. (1997). Multicriteria ranking with ordinal data. *Systems Analysis-Modelling-Simulation*, 27(2), 139–146.
- Hennig-Thurau, T. (2004). Customer orientation of service employees. *International Journal of Service Industry Management*, 15(5), 460–478. <https://doi.org/10.1108/09564230410564939>
- Kitchner, K. S. (1983). Cognition, Metacognition, and Epistemic Cognition. *Human Development*, 26(4), 222–232. <https://doi.org/10.1159/000272885>
- Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling*. <https://doi.org/10.1007/978-1-4614-6849-3>
- Kumar, P. (2012). A Benchmark to Select Data Mining Based Classification Algorithms for Business Intelligence and Decision Support Systems. *International Journal of Data Mining & Knowledge Management Process*, 2(5), 25–42. <https://doi.org/10.5121/ijdkp.2012.2503>
- Hossin, M., & Sulaiman, M. N. (2015). A Review on Evaluation Metrics for Data Classification Evaluations. *International Journal of Data Mining & Knowledge Management Process*, 5(2), 1–11. <https://doi.org/10.5121/ijdkp.2015.5201>

- Martens, D., Vanthienen, J., Verbeke, W., & Baesens, B. (2011). Performance of classification models from a user perspective. *Decision Support Systems*, 51(4), 782–793. <https://doi.org/10.1016/j.dss.2011.01.013>
- Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1–38. <https://doi.org/10.1016/j.artint.2018.07.007>
- Mingers, J., & Rosenhead, J. (2004). Problem structuring methods in action. *European Journal of Operational Research*, 152(3), 530–554. [https://doi.org/10.1016/s0377-2217\(03\)00056-0](https://doi.org/10.1016/s0377-2217(03)00056-0)
- OnlineOutput. (2017). *Vikor Software*. Retrieved from <http://www.onlineoutput.com/vikor-software/>
- Pfeifer, P. E. (2005). The optimal ratio of acquisition and retention costs. *Journal of Targeting, Measurement and Analysis for Marketing*, 13(2), 179–188. <https://doi.org/10.1057/palgrave.jt.5740142>
- Simon, H. A. (1973). The structure of ill structured problems. *Artificial Intelligence*, 4(3–4), 181–201. [https://doi.org/10.1016/0004-3702\(73\)90011-8](https://doi.org/10.1016/0004-3702(73)90011-8)
- Teixeira, L. F. H. S. B., Santos, M., & Gomes, C. F. S. (2018). *SapevoWeb Software (v.1)*. Retrieved from <http://www.sapevoweb.com>
- Teixeira, L. F. H. S. B., Santos, M., & Gomes, C. F. S. (2019). *Proposta e implementação em Python do Método Simple Aggregation of Preferences Expressed by Ordinal Vectors - Multi Decision Makers: uma ferramenta web simples e intuitiva para Apoio à Decisão Multicritério*. <https://doi.org/10.5151/spolm2019-168>
- Verbeke, W., Dejaeger, K., Martens, D., Hur, J., & Baesens, B. (2012). New insights into churn prediction in the telecommunication sector: A profit driven data mining approach. *European Journal of Operational Research*, 218(1), 211–229. <https://doi.org/10.1016/j.ejor.2011.09.031>