

# Credit scoring development in the light of the new Brazilian General Data Protection Law

Robinson A. A. de Oliveira Junior

Universidade de São Paulo, Brazil  
robinson@usp.br

**Abstract.** With the advent of the new Brazilian General Data Protection Law (LGPD) which determines the right to the explanation of automated decisions, the use of non-interpretable models for human beings, known as black boxes, for the purposes of credit risk assessment may remain unfeasible. Thus, three different methods commonly applied to credit scoring – logistic regression, decision tree, and support vector machine (SVM) – were adjusted to an anonymized sample of a consumer credit portfolio from a credit union. Their results were compared and the adequacy of the explanation achieved for each classifier was assessed. Particularly for the SVM, which generated a black box model, a local interpretation method – the SHapley Additive exPlanation (SHAP) – was incorporated, enabling this machine learning classifier to fulfill the requirements imposed by the new LGPD, in equivalence to the inherent comprehensibility of the white box models.

CCS Concepts: • **Computing methodologies** → **Supervised learning by classification**.

Keywords: Classification models, Credit scoring, Interpretability, LGPD, Machine learning

## 1. INTRODUÇÃO

Modelos de *credit scoring* foram uma das primeiras ferramentas de gestão de risco desenvolvidas e também uma das primeiras a admitir dados comportamentais [Thomas et al. 2002]. Vitais à expansão do crédito ao consumidor ao longo das últimas décadas, sua relevância é incontestável, bem como a importância e o protagonismo presentes da inteligência artificial.

Entretanto, os meios utilizados para se prever um evento que possa impactar vidas humanas de forma tão profunda, a exemplo da atribuição de um perfil capaz de regular o acesso individual a bens, serviços ou recursos financeiros, não são incontestes. E a Lei Geral de Proteção de Dados (LGPD), Lei nº 13.709/18<sup>1</sup>, inspirada no Regulamento Geral sobre a Proteção de Dados (*General Data Protection Regulation* (GDPR)) da União Europeia [Erickson 2019], assegura aos titulares dos dados direitos de informação sobre a lógica de processos de tratamento de dados e de revisão de decisões tomadas unicamente com base no tratamento automatizado de dados pessoais que afetem seus interesses.

Para modelos de *credit scoring* baseados em técnicas estatísticas, a exemplo da análise de discriminantes lineares e da regressão logística, a predição pode ser interpretada a partir de seus coeficientes [Hand and Henley 1997]. Em contraste, para alguns modelos baseados em técnicas de aprendizado de máquina específicas, como redes neurais artificiais, máquinas de vetores de suporte e redes neurais profundas, não é possível identificar como são determinadas as suas predições, sendo frequentemente denominados caixas-pretas [Molnar 2020], algo incompatível com a LGPD.

---

<sup>1</sup>Dispõe sobre a proteção dos dados pessoais e regula o tratamento e a transferência desses dados, incluindo os virtuais.

Dentre as técnicas mais recentes para se incorporar interpretabilidade aos modelos caixas-pretas, independentemente do modelo preditivo previamente induzido (*model-agnostic*), há a *SHapley Additive exPlanation* (SHAP)<sup>2</sup>, uma técnica estado da arte, fundamentada na teoria dos jogos cooperativos, capaz de determinar a contribuição marginal média ponderada de um atributo qualquer ao valor de saída do modelo [Molnar 2020].

Diante do exposto, este artigo, que é uma síntese de uma pesquisa [Oliveira-Junior 2021], objetiva aplicar sobre uma amostra anonimizada de uma carteira de crédito ao consumidor de uma cooperativa de crédito, alguns dos métodos de classificação aplicados para *credit scoring* mais difundidos<sup>3</sup> – (i) um probabilístico, de regressão logística, (ii) um baseado em procura, de árvore de decisão, e (iii) um baseado em otimização, de máquina de vetores de suporte<sup>4</sup>, com especial interesse pelo último, por se tratar de um método caixa-preta – comparando os resultados obtidos e avaliando a adequabilidade da explicação obtida para cada classificador, seja ela aparente ou alcançada a partir do emprego do método de interpretação local, à luz da nova LGPD.

No Brasil, publicações acerca da interpretabilidade são relativamente recentes, destacando-se a dissertação “A semiotic view on the interpretability of machine learning models” [Santo 2018], em que o autor sugere uma sistemática para lidar com o desafio da interpretabilidade. Em 2020, segundo Oliveira-Junior (2021), a busca conjunta pelas palavras-chave associadas à pesquisa na plataforma “Google Acadêmico”<sup>5</sup> não retornou resultados. Em virtude da maior compatibilidade com o tema proposto e área de aplicação, pode ainda ser citada a dissertação “A utilização de algoritmos de aprendizado de máquina em problemas de classificação” [Batista 2018].

Assim, este artigo visa contribuir com os estudos sobre interpretabilidade ao apresentar o uso da SHAP incorporado ao processo de modelagem de *credit scoring* a fim de permitir a utilização de modelos caixas-pretas no processo de formação de perfis de crédito, em conformidade com os requisitos da nova LGPD.

## 2. METODOLOGIA

Foi extraída uma amostra com 7.761 decisões de crédito (observações) ocorridas entre os anos de 2016 e 2018, e apoiada em 59 atributos – relacionados a dados cadastrais (13) e comportamentais dos cooperados (11), a dados das operações (32) e de performance (3) – oriunda de uma cooperativa de crédito e limitada ao público pessoa física. O conceito de mau (mau pagador) empregado foi de 60 dias de atraso em uma janela de 12 meses, de modo a tornar exequível a classificação<sup>6</sup>.

Para a análise exploratória, realizada em Python, foram verificadas a estrutura do conjunto de dados, as medidas básicas de tendência central e de dispersão, o tipo dos atributos e outras características, tais como sua não-nulidade, valores faltantes, cardinalidade e eventuais redundâncias. Tendo em vista o conceito de mau escolhido, as 7.761 observações iniciais se dividiam em 7.663 bons pagadores, dita classe negativa (“0”), e 98 maus pagadores, dita classe positiva (“1”).

Uma parcela dos atributos originais permitiu a geração de novos atributos (40), os quais, em sua maioria, correspondiam a sumarização de informações: por exemplo, o valor mensal para percentual

<sup>2</sup>Embora a *Local Interpretable Model-agnostic Explanations* (LIME) seja um método de interpretação local pioneiro, que induz um modelo substituto capaz de criar uma aproximação à predição que se deseja explicar a partir de propriedades da instância em questão [Ribeiro et al. 2016], possibilitando que os coeficientes deste sejam utilizados para se obter a explicação pretendida, a técnica foi preterida por não oferecer precisão equivalente a SHAP [Molnar 2020].

<sup>3</sup>A seleção dos métodos teve como base uma revisão sistemática de métodos de classificação aplicados para *credit scoring* [Louzada et al. 2016]: <https://doi.org/10.1016/j.sorms.2016.10.001>

<sup>4</sup>O método de redes neurais artificiais foi preterido em detrimento da máquina de vetores de suporte, pois se relatou observar com mais frequência que o primeiro perdia desempenho preditivo diante de um conjunto de dados desbalanceado [Louzada et al. 2016].

<sup>5</sup><https://scholar.google.com.br/?hl=pt>

<sup>6</sup>Dado o envolvimento do cooperado com a cooperativa, atrasos superiores a 90 dias são pouco usuais.

de utilização do cheque especial produziu quatro novos atributos, apresentando a média de utilização nos (i) 3, (ii) 6 e (iii) 12 meses que antecederam a decisão de crédito, além do (iv) percentual máximo de utilização no período.

Após visualização das distribuições, histogramas, *boxplots* e contagens – também observadas sob a perspectiva de cada classe à variável resposta, e a partir dos dados normalizados – agrupamento de categorias, discretização de atributos, identificação e tratamento de *outliers*, exame da correlação linear entre as variáveis independentes (previsoras) e a variável dependente (resposta), e estimação do *Weight of Evidence* (WoE)<sup>7</sup> e *Information Value* (IV)<sup>8</sup>, foram descartados 76 atributos, restando apenas 23 – 11 originais e 12 novos – mais a variável resposta (“TARGET”), conforme tabela I a seguir, sendo preservadas 5.149 observações, com 5.062 bons pagadores e 87 maus pagadores.

IDADE	ESTADOCIVIL	NUMERODEPENDENTES	ESCOLARIDADE	SUBREGIAO
MUNICIPIO	PA	PRODUTO	RENDA	PATRIMONIO
IAP_DISC	CC_ATIVO_FLAG	LIQUIDADOS_DISC	REPACTUOU_FLAG	DEV_ALINEAS_11E12_FLAG
SALDO_MD_12M	PERC_MD_CHQ_ESP_12M	PERC_MAX_CHQ_ESP	NRO_MAX_DIAS_AD	NRO_MAX_MESES_AD
SALDO_MD_VENCIDOS_12M	SALDO_MAX_VENCIDOS	NRO_MAX_MESES_ACIMA_60	TARGET	

Tabela I. Atributos restantes.

Verde: atributos originais; Amarelo: atributos novos.

Fonte: Elaborada pelo autor.

A análise exploratória no conjunto de dados obtido apresentou uma realidade bastante desafiadora para a modelagem de *credit scoring*. Dentre as razões, podem ser destacadas:

- ao solicitar que não houvesse repetições em diferentes pontos de observação, substituindo solicitantes “repetidos” por “novos”, o volume médio de decisões de crédito mensais foi de apenas 216 ( $7.761 \div 36$ );
- baixíssima frequência à classe positiva (maus pagadores), culminando em um conjunto de dados extremamente desbalanceado (ao final, apenas 1,69% dos exemplos –  $87 \div 5.149$ );
- embora as operações rodem em sistemas transacionais, inúmeras características de grande relevância à modelagem de *credit scoring* não eram registradas à ocasião da tomada de decisão – ex. dados de birôs e do Sistema de Informações de Créditos (SCR) do Banco Central do Brasil;
- uma correlação linear insignificante entre os atributos preditores e a variável resposta.

### 3. DESENVOLVIMENTO DA MODELAGEM DE *CREDIT SCORING*

O conjunto de dados foi separado em treino e teste usando a função “train\_test\_split” do pacote “sklearn”, com 20% para o conjunto de teste, e mantendo a proporção observada às classes da variável resposta no conjunto de dados.

Aos atributos numéricos foi aplicado o “StandardScaler” para padronização dos dados, também do “sklearn”, removendo a média e reescalando a variância para uma unidade. Os atributos categóricos foram transformados usando a classe referente ao *Weight of Evidence* do pacote “Category Encoders”.

Antes de lidar com o desbalanceamento do conjunto – é esperado que diferenças observadas nas frequências das classes afetem significativamente a previsibilidade dos modelos – foi estabelecido um classificador de referência ou *baseline*: uma árvore de decisão sobre a amostra desbalanceada utilizando *grid search* para otimização de hiperparâmetros, ambos do pacote “sklearn”, e validação cruzada com 3 *folds*<sup>9</sup>, após divisão estratificada dos conjuntos de treino e teste.

<sup>7</sup>Aponta o poder preditivo de cada categoria de uma variável independente em relação à variável dependente.

<sup>8</sup>Valor único que aponta todo o poder preditivo de uma variável, bastante útil à seleção de atributos.

<sup>9</sup>Dividindo o conjunto de treinamento em  $k$  subconjuntos de mesmo tamanho e mutuamente exclusivos, onde  $k = 3$ , e utilizando então um subconjunto para teste e os  $k - 1$  restantes para estimação dos parâmetros.

A acurácia obtida foi de 97,38%, porém a medida não possui qualquer utilidade por força do desbalanceamento observado, tendo concentrado as predições à classe majoritária, ou seja, classificando as observações no conjunto de teste como bons pagadores. Para a decisão de crédito é importante que o modelo apresente alta revocação<sup>10</sup>, pois o custo de confundir um mau pagador com um bom pagador costuma ser elevado. Ao mesmo tempo, não se pode prescindir da precisão<sup>11</sup>, uma vez que classificar bons pagadores como maus pagadores pode restringir negócios, e o crédito é a maior fonte de receita das instituições financeiras.

Por sua vez, a medida  $F$  corresponde à média harmônica da precisão e da revocação. Como não podemos desacolher ambas, nos caberia apenas ajustá-la a fim de redistribuir pesos, conferindo maior importância à revocação: tal medida é conhecida como  $F_2$  e é expressa por  $[(1 + 2^2) \times \text{precisão} \times \text{revocação}] \div [2^2 \times (\text{precisão} + \text{revocação})]$ . Para o classificador de referência, precisão, revocação e medida  $F_2$  atingiram 18,75%, 17,65% e 17,86%, respectivamente.

Na sequência foi desenvolvido um *pipeline* para implementar as técnicas de balanceamento<sup>12</sup> – tão somente no conjunto de treino – e ajustar os modelos de regressão logística, de árvore de decisão e de máquina de vetores de suporte, disponíveis no pacote “sklearn”, em conjunto com a utilização de *grid search* para otimização de hiperparâmetros e validação cruzada com 3 *folds*, concluindo sua execução com a exibição da matriz de confusão e medidas de avaliação<sup>13</sup>. Abaixo, na tabela II, são exibidos os melhores modelos obtidos para cada método, incluindo técnica de balanceamento e hiperparâmetros:

Classificador	Técnica	Hiperparâmetros	VN	FN	VP	FP	Precisão	Revocação	$F_2$	Acurácia	Acurácia Balanceada	Especificidade	ROC AUC
DT	N/A	criterion: 'gini', max_depth: '7', splitter: 'best'	1000	14	3	13	0.1875	0.1765	0.1786	0.9738	0.5818	0.9872	0.6903
LR	ROS	C: '100.0'	849	5	12	164	0.0682	0.7059	0.2459	0.8359	0.772	0.8381	0.8134
DT	RUS	criterion: 'gini', max_depth: '4', splitter: 'random'	753	6	11	260	0.0406	0.6471	0.1622	0.7417	0.6952	0.7433	0.7151
SVM	CW	C: '10', class_weight: 'balanced', gamma: '0.0001', kernel: 'rbf'	809	7	10	204	0.0467	0.5882	0.1773	0.7951	0.6934	0.7986	0.7538

Tabela II. Melhores modelos obtidos para cada técnica previamente definida.  
Classificador: DT = árvore de decisão; LR = regressão logística; SVM = máquina de vetores de suporte.  
Técnica: ROS = *random over sampling*; RUS = *random under sampling*; CW = *class weighting*.  
Fonte: Dados da pesquisa.

Para o modelo de melhor performance, o de regressão logística, em relação ao classificador de referência, foi possível alcançar um incremento de aproximadamente 300% à revocação, 37,7% à medida  $F_2$ , 32,7% à acurácia balanceada e 17,8% à área sob a curva ROC.

## 4. INTERPRETABILIDADE DOS MÉTODOS

Após a obtenção dos modelos verificou-se a adequabilidade da explicação alcançada para cada classificador.

### 4.1 Regressão Logística

Para a regressão logística, um método que resulta em um modelo caixa-branca, a relação entre as entradas e as saídas são distinguíveis a partir de seus coeficientes. Usando o pacote “Eli5” (“Explain Like I’m 5”) para visualizar como o modelo funciona, vemos na figura 1, à esquerda, seus coeficientes

<sup>10</sup>Mostra o quão frequente o classificador detecta instâncias pertencentes à classe minoritária:  $VP \div (VP + FN)$

<sup>11</sup>Mostra o quão frequente o classificador está correto ao predizer uma instância da classe minoritária:  $VP \div (VP + FP)$

<sup>12</sup>De subamostragem com *Random Under Sampling* e *Cluster Centroids*, de sobreamostragem com *Random Over Sampling* e *Synthetic Minority Over-sampling Technique* (SMOTE) e combinada com SMOTEEN, usando o pacote “imblearn”, e de ponderação de classes diretamente no classificador (“class\_weight”).

<sup>13</sup> $F_2$ , acurácia, acurácia balanceada, especificidade e área sob a curva ROC.

(incluindo o intercepto, “BIAS”), com destaque para o atributo DEV\_ALINEAS\_11E12\_FLAG, que aparenta ser o atributo mais importante para o classificador; à direita, ao selecionar um exemplo “mau pagador” (y=1), identificamos a contribuição de cada atributo (para modelos lineares, a contribuição é igual ao peso × valor do atributo) para a pontuação final ou probabilidade:

Fig. 1. Esquerda: Coeficientes do modelo. Direita: Exemplo predito como mau pagador.

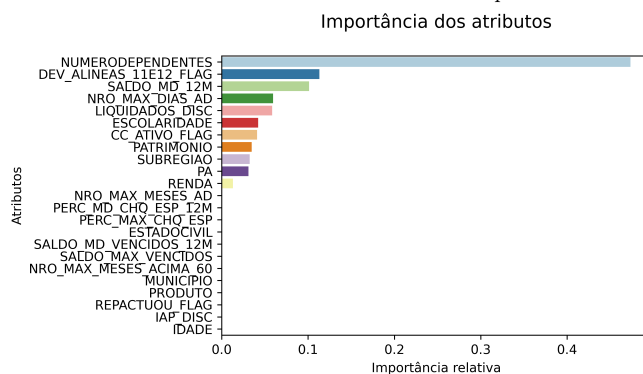
Weight?	Feature	Contribution?	Feature	Value
+0.994	DEV_ALINEAS_11E12_FLAG	+1.395	NUMERODEPENDENTES	1.671
+0.854	LIQUIDADOS_DISC	+0.807	MUNICIPIO	-1.247
+0.835	NUMERODEPENDENTES	+0.702	PERC_MAX_CHQ_ESP	1.022
+0.751	NRO_MAX_MESES_AD	+0.485	IDADE	-1.506
+0.687	PERC_MAX_CHQ_ESP	+0.462	LIQUIDADOS_DISC	0.541
+0.420	CC_ATIVO_FLAG	+0.329	ESCOLARIDADE	1.566
+0.389	PATRIMONIO	+0.293	PA	-0.805
+0.210	ESCOLARIDADE	+0.268	PRODUTO	2.118
+0.199	ESTADOCIVIL	+0.220	SALDO_MD_12M	-0.372
+0.127	PRODUTO	+0.132	IAP_DISC	-0.173
+0.110	SALDO_MAX_VENCIDOS	+0.011	NRO_MAX_DIAS_AD	-0.181
+0.084	SALDO_MD_VENCIDOS_12M	+0.000	NRO_MAX_MESES_ACIMA_60	-0.064
+0.058	RENTA	-0.001	CC_ATIVO_FLAG	-0.001
-0.004	NRO_MAX_MESES_ACIMA_60	-0.007	SALDO_MAX_VENCIDOS	-0.065
-0.035	SUBREGIAO	-0.010	SALDO_MD_VENCIDOS_12M	-0.122
-0.060	NRO_MAX_DIAS_AD	-0.028	RENTA	-0.487
-0.090	REPACTUOU_FLAG	-0.037	PERC_MD_CHQ_ESP_12M	0.043
-0.322	IDADE	-0.052	REPACTUOU_FLAG	0.580
-0.364	PA	-0.053	SUBREGIAO	1.505
-0.593	SALDO_MD_12M	-0.070	ESTADOCIVIL	-0.351
-0.647	MUNICIPIO	-0.152	PATRIMONIO	-0.391
-0.762	IAP_DISC	-0.355	NRO_MAX_MESES_AD	-0.473
-0.857	PERC_MD_CHQ_ESP_12M	-0.369	DEV_ALINEAS_11E12_FLAG	-0.371
-0.931	<BIAS>	-0.931	<BIAS>	1.000

Fonte: Dados da pesquisa.

#### 4.2 Árvore de decisão

Para a árvore de decisão, um método que também resulta em um modelo caixa-branca, é possível verificar após ajuste do modelo (figura 2) a importância relativa de cada atributo ao fazer uma previsão, a partir da propriedade “feature\_importances\_” do classificador:

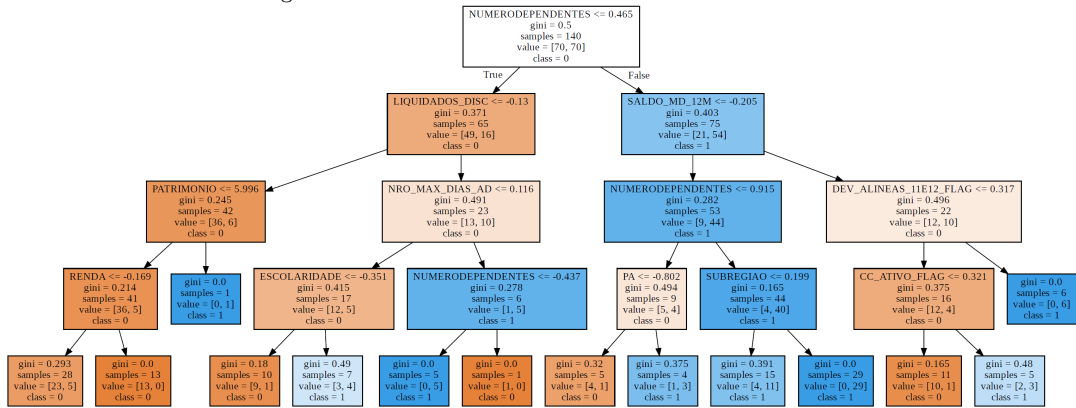
Fig. 2. Árvore de decisão: atributos mais relevantes para o classificador.



Fonte: Dados da pesquisa.

Adicionalmente, é também possível percorrer os ramos da árvore e verificar seus testes condicionais, permitindo explicar o modo como foi realizada a previsão, conforme exibido na figura 3:

Fig. 3. Árvore de decisão: nós com testes condicionais.



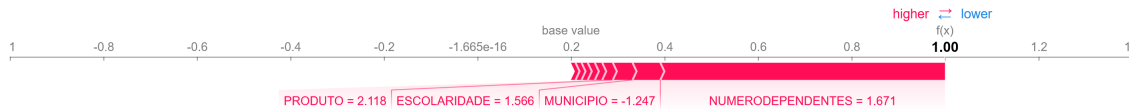
Fonte: Dados da pesquisa.

### 4.3 Máquina de vetores de suporte (*Support Vector Machine (SVM)*)

Em relação à máquina de vetores de suporte, um método que resulta em um modelo caixa-preta, não é possível identificar como o modelo determina as suas saídas. Portanto, para interpretá-lo usaremos a SHAP, instanciando um “explicador” genérico (*model-agnostic*) a fim de explicar as predições do modelo a partir dos valores SHAP aproximados<sup>14</sup>, conquanto suficientes para racionalizar a decisão.

A seguir, na figura 4, uma decisão de crédito predita pelo classificador como “mau pagador”:

Fig. 4. Máquina de vetores de suporte: visualização gráfica da contribuição de cada atributo à saída do modelo.



Fonte: Dados da pesquisa.

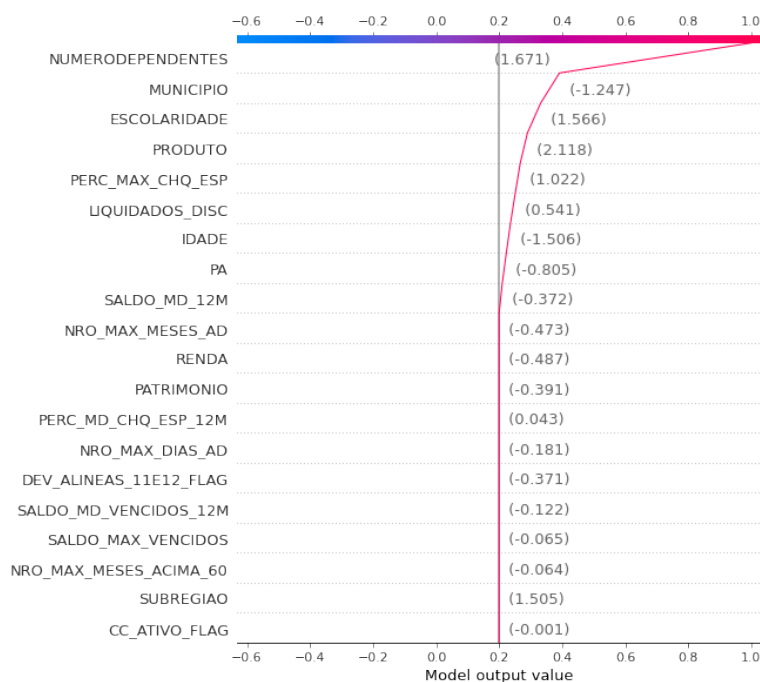
- $f(x)$  corresponde à predição – “1” para mau pagador;
- valor base (*base value*) representa o valor médio predito para o conjunto de dados, equivalente ao valor obtido se não conhecêssemos qualquer atributo à saída em análise;
- os atributos que aumentam a predição média são exibidos em vermelho, com o tamanho da faixa representando a magnitude do efeito do atributo, enquanto aqueles que reduzem o valor predito são exibidos em azul (não presentes neste exemplo na figura 4):
  - o atributo “NUMERODEPENDENTES” teve grande impacto à classificação, sendo o valor absoluto observado para o exemplo (–1.671) maior que o valor médio absoluto observado para o atributo (–0.547), referência à obtenção da contribuição marginal, o que “empurrou” a predição em direção a “1”;
  - o mesmo ocorreu, por exemplo, com os atributos “MUNICIPIO”, “ESCOLARIDADE” e “PRODUTO”.

Tomando o mesmo exemplo da figura 4, os valores SHAP<sup>15</sup> antes representados graficamente pelas faixas em vermelho podem agora ser identificados individualmente junto aos seus valores aproximados, indicando o quanto cada atributo contribuiu para a predição “mau pagador” realizada pelo modelo caixa-preta (figura 5):

<sup>14</sup>Não há um “explicador” otimizado para máquinas de vetores de suporte.

<sup>15</sup>Contribuição marginal média ponderada de um atributo qualquer ao valor de saída do modelo.

Fig. 5. Máquina de vetores de suporte: visualização gráfica alternativa da contribuição de cada atributo à saída do modelo.



Fonte: Dados da pesquisa.

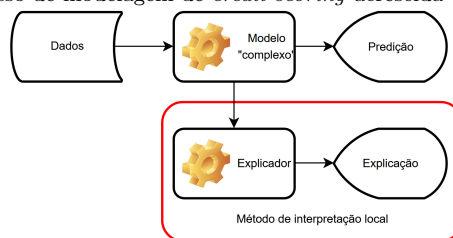
## 5. CONCLUSÃO

Sob a perspectiva da precisão, dentre os métodos escolhidos, a regressão logística apresentou o melhor desempenho. Por se tratar de um método caixa-branca, sua interpretação é aparente e não se requer qualquer modificação ao processo habitual de modelagem de *credit scoring*, diferentemente de quando empregados métodos caixas-pretas. Testemunhou-se que a melhor solução pode residir em uma abordagem simples, todavia eficaz.

À avaliação da interpretabilidade, foco da análise experimental, através da implementação de um método de interpretação local foi possível compreender e explicar as predições do método caixa-preta, ao se identificar os atributos de maior relevância para a predição, juntamente com a estimativa de sua contribuição, revelando como cada atributo afetou a predição realizada e justificando adequadamente, e com precisão, a lógica do processo de tratamento de dados.

Logo, ao empregar métodos que produzam modelos caixas-pretas, sugere-se agregar ao processo habitual de modelagem de *credit scoring* a explicação/valor de Shapley, como exibido na figura 6:

Fig. 6. Etapa final do processo de modelagem de *credit scoring* acrescida da explicação/valor de Shapley.



Fonte: Elaborada pelo autor.

Ainda que a LGPD não explicita o que poderia ser considerada uma resposta aceitável ao titular dos dados, o Comitê Europeu para a Proteção de Dados (*European Data Protection Board*) um órgão consultivo independente para proteção de dados e privacidade, recomendou esclarecer [EDPB 2018]:

- as categorias de dados utilizadas à formação do perfil comportamental;
- uma síntese do porquê essas categorias são consideradas apropriadas;
- a razão pela qual esse perfil comportamental é relevante ao processo de decisão automatizado; e
- como esse perfil é utilizado para uma decisão relativa ao titular dos dados.

Para além disso, visando prover maior transparência ao tratamento de dados, pode-se apresentar quais dados especificamente foram utilizados, de onde vêm, quais os parâmetros de tratamento, tornando clara e compreensível a sua política e, com o esclarecimento oferecido pela SHAP, como esses dados contribuíram efetivamente à tomada de decisão. Tais explicações serão um dos principais instrumentos de apelação contra o tratamento automático realizado de forma inadequada.

Trabalhos futuros incluem a implementação de *ensembles* e de redes neurais profundas, e a variação da abordagem, desta vez sob a perspectiva de um problema de classificação de uma classe (*One-Class Classification*) – em que se assume como premissa que todo o conjunto de dados conteria apenas instâncias “normais” e que o público pertencente à atual classe positiva representaria instâncias anômalas – e também como uma tarefa de aprendizado incremental (*online learning/data stream learning*) – capaz de alterar dinamicamente sua estrutura sempre que surgirem novos dados – pois é esperado que cenários de estresse financeiro sejam cada vez mais recorrentes, o que exigirá rápida adaptação à nova realidade.

## REFERENCES

- BATISTA, M. R. S. A utilização de algoritmos de aprendizado de máquina em problemas de classificação. Universidade de São Paulo (USP), 2018.
- EDPB. EUROPEAN DATA PROTECTION BOARD. Article 29 Data Protection Working Party. Guidelines on automated individual decision-making and profiling for the purposes of regulation 2016/679. *EDPB*, Feb., 2018.
- ERICKSON, A. Comparative analysis of the EU’s GDPR and Brazil’s LGPD: Enforcement challenges with the LGPD. *Brooklyn Journal of International Law* 44 (2): 859–888, 2019.
- HAND, D. J. AND HENLEY, W. E. Statistical classification methods in consumer credit scoring: A review. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 160 (3): 523–541, 1997.
- LOUZADA, F., ARA, A., AND FERNANDES, G. B. Classification methods applied to credit scoring: Systematic review and overall comparison. *Surveys in Operations Research and Management Science* 21 (2): 117–134, 2016.
- MOLNAR, C. Interpretable machine learning, 2020.
- OLIVEIRA-JUNIOR, R. A. A. Um estudo acerca dos métodos estatísticos clássicos e de inteligência artificial aplicados ao desenvolvimento de credit scoring à luz da nova Lei Geral de Proteção de Dados brasileira. Universidade de São Paulo (USP), 2021.
- RIBEIRO, M. T., SINGH, S., AND GUESTRIN, C. “Why should I trust you?”: Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD ’16. Association for Computing Machinery, New York, pp. 1135–1144, 2016.
- SANTO, J. L. C. F. A semiotic view on the interpretability of machine learning models. Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), 2018.
- THOMAS, L. C., CROOK, J. N., AND EDELMAN, D. B. *Credit scoring and its applications*. Society for Industrial and Applied Mathematics, New York, 2002.