

Detecção de fraude em transações de cartão de crédito via *Machine Learning*

Aluno: Maurício Najjar da Silveira
Orientador: Adriano Polpo de Campos

Universidade Federal de São Carlos

22 de novembro de 2018

Sumário

- 1 Introdução
- 2 Métodos
- 3 Resultados
- 4 Comparação dos Modelos
- 5 Conclusão

Motivação

- Segundo PwC, por volta de 49% das organizações já lidaram com crimes econômicos nos últimos 2 anos;
- Comitê de Basiléia - Lavagem de Dinheiro;
- Vazamento de Bancos de Dados.

Conjunto de Dados

Coletado pela empresa Worldline em parceria com o Machine Learning Group (ULB).

- 285 mil transações de cartão de crédito em e-commerce;
- 2 dias de setembro de 2013 pela Europa;
- Variáveis: 28 componentes principais (ortogonais), Valor, Tempo e Classe;
- Altamente desbalanceado, 0,172% de fraude.

Objetivo

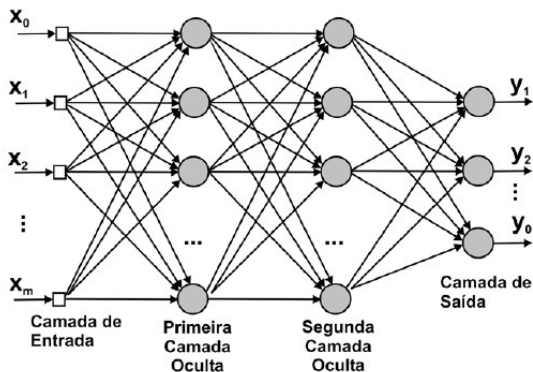
Construir modelos com foco em previsão para detecção em novas transações.

Apresentar alternativas de modelos via *Machine Learning* para detecção de fraude em transações de cartão de crédito.

Machine Learning

- Samuel [1959]
- Aprendizado Supervisionado: Predição
- Aprendizado Não-Supervisionado: Estrutura

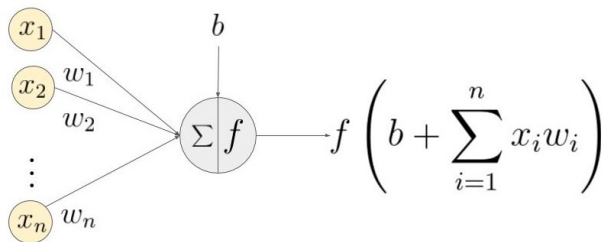
Rede Neural Artificial



Estrutura customizável em:

- # Camadas;
- # Neurônios por Camada;
- Funções de Ativação;
- Funções de Perda.

Rede Neural Artificial

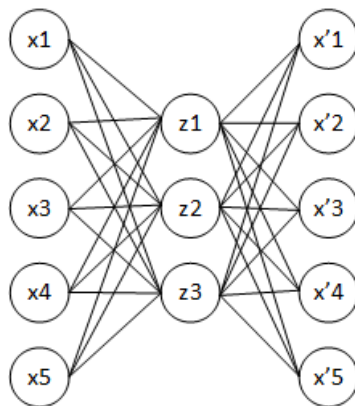


Funções de ativação possíveis:

- Sigmóide/Logito: $\frac{e^x}{e^x + 1}$
- Tangente Hiperbólica: $\frac{1 - e^{-2x}}{1 + e^{-2x}}$
- ReLU: $\max(0, x)$
- etc...

AutoEncoder

Objetivo: Copiar a camada de entrada na camada de saída.



Encoder: $\mathbf{h} = \mathbf{f}(\mathbf{x})$

Decoder: $\mathbf{r} = \mathbf{g}(\mathbf{h})$

AutoEncoder Preditivo

Em Schreyer et al. [2017], fazem uso de AutoEncoder para detectar anomalias em dados de contabilidade.

Treino do modelo: Dados não-anômalos

Se a predição do AutoEncoder divergir muito, é um indicativo de anomalia.

Regressão Logística

Para Y binária, com distribuição Bernoulli de esperança π . Logo

$$Y_i = \pi_i + \epsilon_i, \quad i = 1, \dots, n$$

onde

$$\pi_i = \frac{\exp\{\eta_i\}}{1 + \exp\{\eta_i\}}, \quad i = 1, \dots, n$$

$$\eta_i = \beta_0 + \beta_1 X_{i1} \cdots + \beta_p X_{ip}, \quad i = 1, \dots, n$$

Reescrevendo na forma comum da função de ligação:

- Logito: $\eta_i = \log \frac{\pi_i}{1 - \pi_i}$
- C-Log-Log: $\eta_i = \log(-\log(1 - \pi_i))$
- etc...

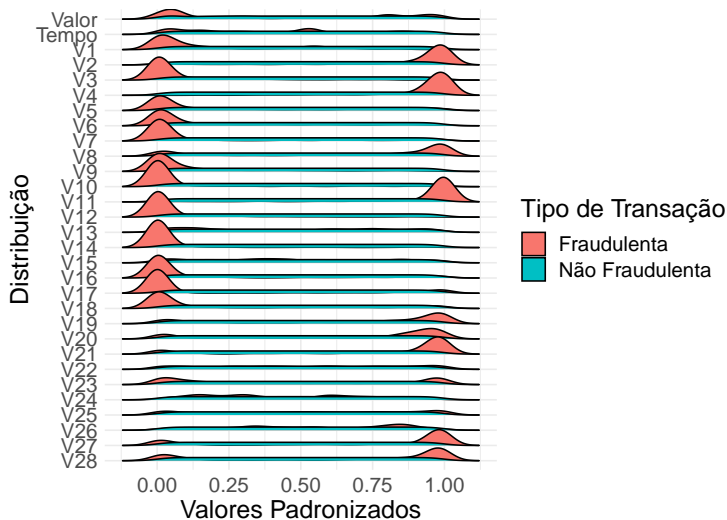
Penalização Lasso

Para seleção de variáveis em um modelo, existe a penalização Lasso.

$$\hat{\beta}_{L_1, \lambda} = \arg \min_{\beta} \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{i,j} \right)^2 + \lambda \sum_{j=1}^p |\beta_j|$$

A estimação é feita numericamente buscando por validação cruzada minimizar a função de risco.

Análise Descritiva

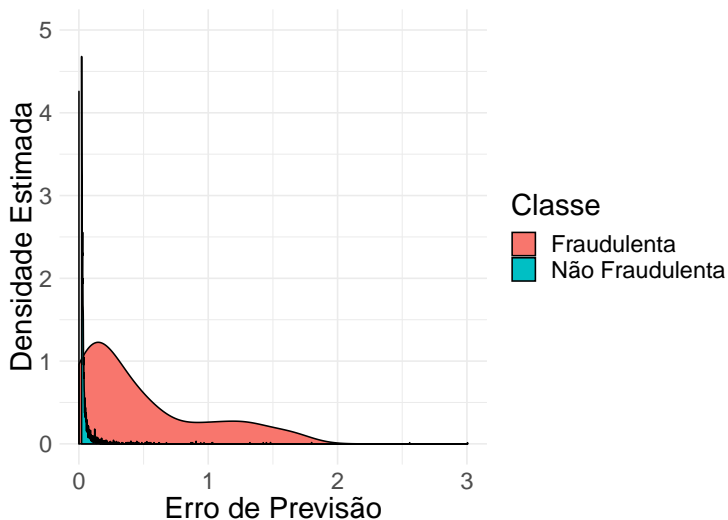


Modelo 1 - AutoEncoder Preditivo

Camada	# Neurônios	# Parâmetros
Entrada	30	-
Camada Oculta 1	15	465
Camada Oculta 2	10	160
Camada Oculta 3	15	165
Saída	30	480
Total	-	1270

$$Classe_i = \begin{cases} \text{Não Fraude,} & \text{se } Erro_i \leq k \\ \text{Fraude,} & \text{CC} \end{cases}$$

Modelo 1 - AutoEncoder Preditivo

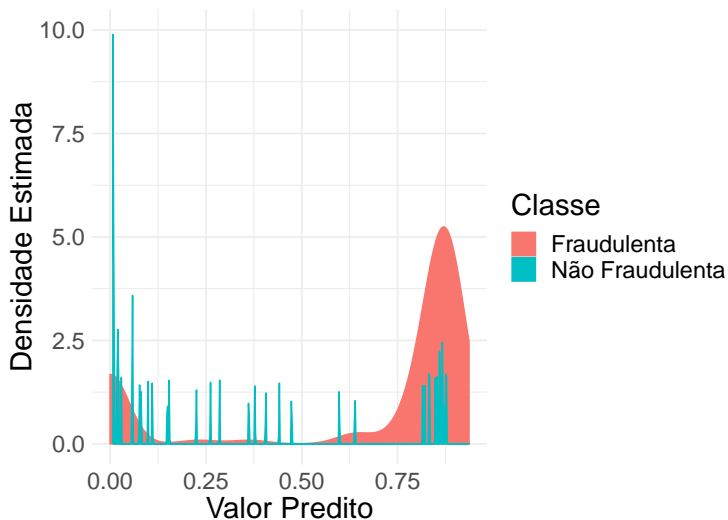


Modelo 2 - AutoEncoder com resposta binária

Camada	# Neurônios	# Parâmetros
Entrada	30	-
Camada Oculta 1	15	465
Camada Oculta 2	10	160
Camada Oculta 3	15	165
Camada Oculta 4	30	480
Saída	1	31
Total	-	1301

$$Classe_i = \begin{cases} \text{Não Fraude,} & \text{se } \hat{Y}_i \leq k \\ \text{Fraude,} & \text{CC} \end{cases}$$

Modelo 2 - AutoEncoder com resposta binária

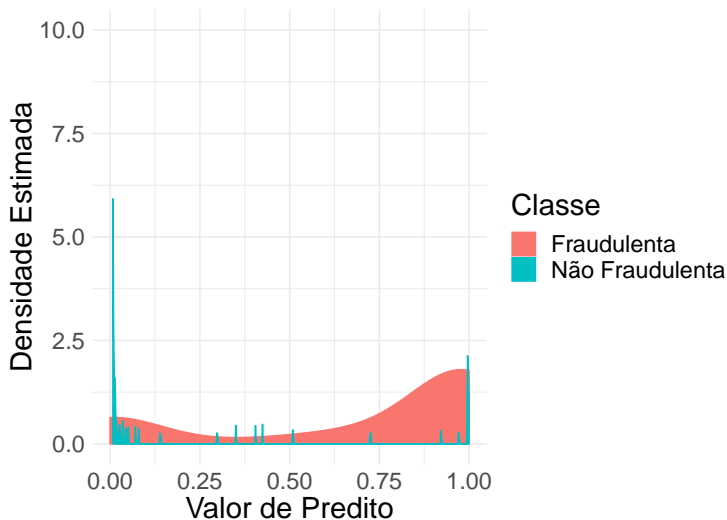


Modelo 3 - RNA padrão com resposta binária

Camada	# Neurônios	# Parâmetros
Entrada	30	-
Camada Oculta 1	15	465
Camada Oculta 2	15	240
Camada Oculta 3	15	240
Saída	1	16
Total	-	961

$$Classe_i = \begin{cases} \text{Não Fraude,} & \text{se } \hat{Y}_i \leq k \\ \text{Fraude,} & \text{CC} \end{cases}$$

Modelo 3 - RNA padrão com resposta binária



Modelos Logísticos

- Modelo 4: Logístico Completo ligação Logito
- Modelo 5: Logístico Completo ligação C-Log-Log
- Modelo 6: Lasso-Logístico

Modelos de Regressão Linear

Sendo a variável resposta na forma:

$$Y_i^* = \begin{cases} (0, 01) * Valor_i, & \text{se transação } i \text{ não for fraudulenta} \\ (-1) * Valor_i, & CC \end{cases}$$

Regredindo esta variável resposta nas covariáveis, a classificação se constrói como:

$$Classe_i = \begin{cases} \text{Fraude,} & \text{se } \hat{Y}_i^* \leq k \\ \text{Não Fraude,} & CC \end{cases}$$

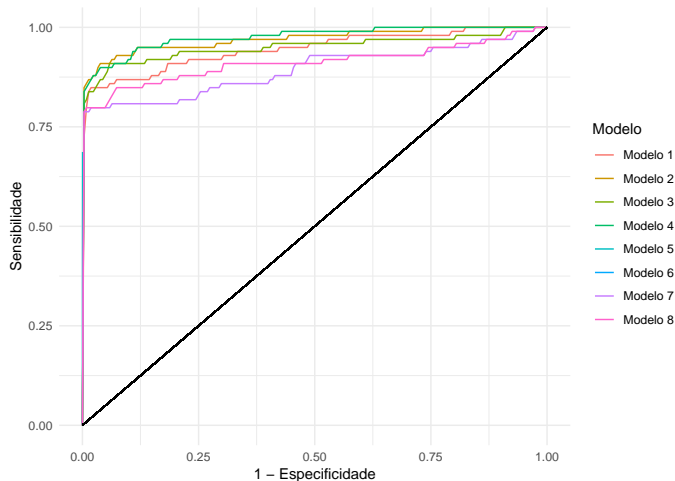
Modelo 7: Regressão Lasso-Linear

Modelo 8: Regressão Linear Completa

Curva ROC

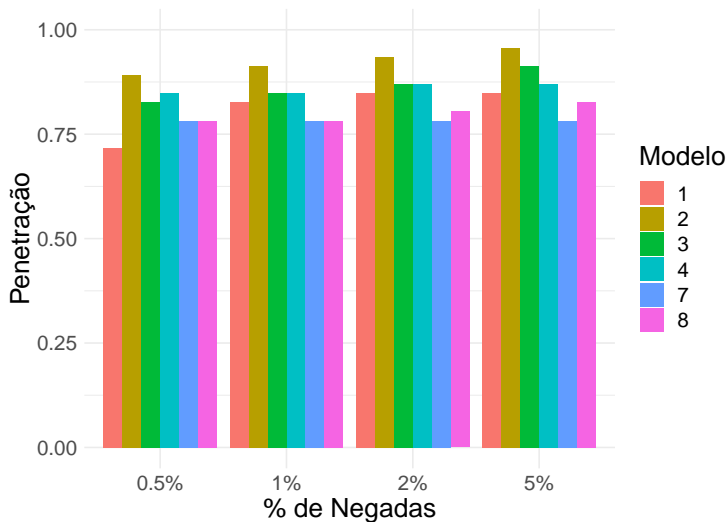
Sensibilidade: $P(\hat{Y} = 1 \mid Y = 1)$

Especificidade: $P(\hat{Y} = 0 \mid Y = 0)$



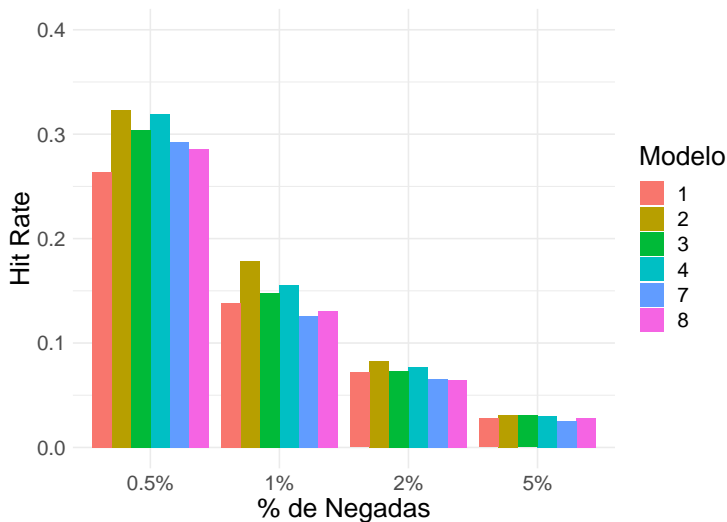
Penetração

$$P(\hat{Y} = 1 \mid Y = 1)$$

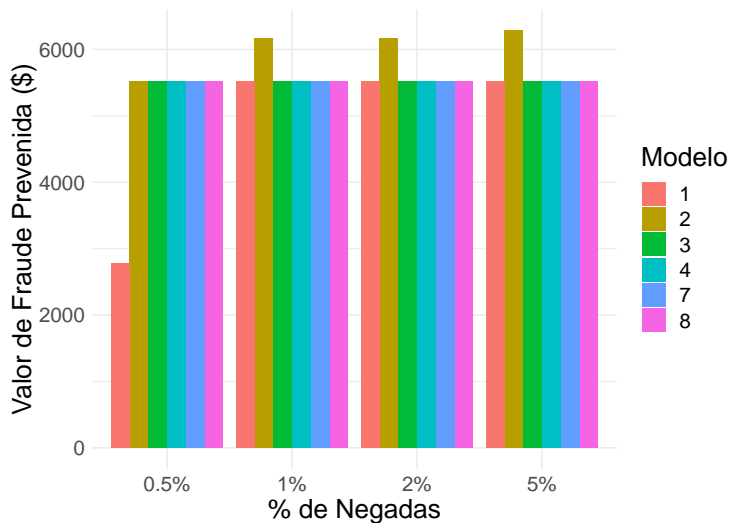


Hit Rate

$$P(Y = 1 \mid \hat{Y} = 1)$$



Valor Prevenido de Fraude



Desempenho Modelos

- No modelo Lasso-Logístico, não houve convergência para o valor de λ ;
- Logística com ligação C-Log-Log não gerou scores muito diferente;
- Variável transformada para regressão linear gerou bons indicadores;
- Modelos via RNA com resposta binária só ficaram significativamente bons após o uso da função de perda por Cross Entropy.

Tempo de Processamento

- Lasso-Logístico levou 28 minutos;
- AutoEncoder Preditivo levou 6 minutos;
- Demais modelos não levaram mais de 1 minuto para rodar.

Detecção de Fraude

Em valor prevenido de fraude, não houve uma grande diferença entre os modelos, implicando que a maioria das fraudes de maior valor foram 'fáceis' de identificar por qualquer modelo.

Para efeito de escolha, o modelo 2 performou melhor em praticamente todos os cenários.

Sobre o Conjunto de Dados

- Quantidade suficiente de observações para geração dos modelos
- Covariáveis independentes e não interpretáveis
- Falsa independência entre transações

Referências I

JJ Allaire and François Chollet. *keras: R Interface to 'Keras'*. URL <https://keras.rstudio.com>. R package version 2.1.6.9000.

Jason Brownlee. Supervised and unsupervised machine learning algorithms, 2016. <https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/> [Acessado: 21-06-2018].

Andrea Dal Pozzolo, Olivier Caelen, Reid A Johnson, and Gianluca Bontempi. Calibrating probability with undersampling for unbalanced classification. In *Computational Intelligence, 2015 IEEE Symposium Series on*, pages 159–166. IEEE, 2015.

Daniel Falbel. Predicting fraud with autoencoders and keras, 2018. <https://tensorflow.rstudio.com/blog/keras-fraud-autoencoder.html> [Acessado: 22-06-2018].

Referências II

Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.

<http://www.deeplearningbook.org>.

Machine Learning Group. Machine learning group, 2018.

<http://mlg.ulb.ac.be/> [Acessado: 07-06-2018].

Rafael Izbicki and Tiago Mendonça dos Santos. *Machine Learning sob a ótica estatística*. Notas de Aula, 2018.

Didier Lavion. Pulling fraud out of the shadows, 2018.

<https://www.pwc.com/gx/en/forensics/global-economic-crime-and-fraud-survey-2018.pdf>
[Acessado: 21-06-2018].

Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.

Referências III

John Neter, Michael H Kutner, Christopher J Nachtsheim, and William Wasserman. *Applied linear statistical models*, volume 4. Irwin Chicago, 1996.

Basel Committee on Banking Supervision. Sound management of risks related to money laundering and financing of terrorism, 2017. <https://www.bis.org/bcbs/publ/d405.pdf>[Acessado: 14-11-2018].

Lazy Programmer. A tutorial on autoencoders for deep learning - lazy programmer, 2015. <https://www.learnopencv.com/understanding-activation-functions-in-deep-learning/>[Acessado: 22-06-2018].

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017. URL <https://www.R-project.org/>.

Referências IV

Arthur L Samuel. Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, 3 (3):210–229, 1959.

Jurgen Schmidhuber. *Deep Learning in Neural Networks: An Overview*. Technical Report IDSIA, 2014.

Marco Schreyer, Timur Sattarov, Damian Borth, Andreas Dengel, and Bernd Reimer. Detection of anomalies in large scale accounting data using deep autoencoder networks. *arXiv preprint arXiv:1709.05254*, 2017.

Aditya Sharma. Understanding activation functions in deep learning learn opencv, 2017. <https://www.learnopencv.com/understanding-activation-functions-in-deep-learning/> [Acessado: 21-06-2018].

Referências V

- Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.
- Machine Learning Group ULB. Predicting fraud with autoencoders and keras, 2018.
<https://www.kaggle.com/mlg-ulb/creditcardfraud>
[Acessado: 22-06-2018].
- Strother H Walker and David B Duncan. Estimation of the probability of an event as a function of several independent variables. *Biometrika*, 54(1-2):167–179, 1967.
- Wikipedia. Data analysis techniques for fraud detection - wikipedia, 2018a. https://en.wikipedia.org/wiki/Data_analysis_techniques_for_fraud_detection [Acessado: 21-06-2018].

Referências VI

Wikipedia. Machine learning - wikipedia, 2018b. https://en.wikipedia.org/wiki/Machine_learning[Acessado: 21-06-2018].