MATHEUS PRESTES ZANI

ANÁLISE COMPARATIVA ENTRE AS ABORDAGENS DE REGRESSÃO LOGÍSTICA E *RANDOM FOREST* EM UM MODELO DE DETECÇÃO DE FRAUDE PARA TRANSAÇÕES *E-COMMERCE* DE CARTÃO DE CRÉDITO

SÃO PAULO
2019

MATHEUS PRESTES ZANI

ANÁLISE COMPARATIVA ENTRE AS ABORDAGENS DE REGRESSÃO LOGÍSTICA E *RANDOM FOREST* EM UM MODELO DE DETECÇÃO DE FRAUDE PARA TRANSAÇÕES *E-COMMERCE* DE CARTÃO DE CRÉDITO

Trabalho de formatura apresentado à Escola Politécnica da Universidade de São Paulo para obtenção do diploma de Engenheiro de Produção

SÃO PAULO
2019

MATHEUS PRESTES ZANI

ANÁLISE COMPARATIVA ENTRE AS ABORDAGENS DE REGRESSÃO LOGÍSTICA E *RANDOM FOREST* EM UM MODELO DE DETECÇÃO DE FRAUDE PARA TRANSAÇÕES *E-COMMERCE* DE CARTÃO DE CRÉDITO

Trabalho de formatura apresentado à Escola Politécnica da Universidade de São Paulo para obtenção do diploma de Engenheiro de Produção

Orientador: Prof. Dr. Mauro de Mesquita Spinola

SÃO PAULO

2019

Catalogação-na-publicação

Zani, Matheus
ANÁLISE COMPARATIVA ENTRE AS ABORDAGENS DE REGRESSÃO LOGÍSTICA E RANDOM FOREST EM UM MODELO DE DETECÇÃO DE FRAUDE PARA TRANSAÇÕES E-COMMERCE DE CARTÃO DE CRÉDITO / M. Zani -- São Paulo, 2019.
130 p.

Trabalho de Formatura - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Produção.

1.APRENDIZADO DE MÁQUINA 2.CARTÃO DE CRÉDITO 3.*E-COMMERCE* 4.FRAUDE 5.TRANSAÇÕES I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Produção II.t.

À minha família e aos colegas acadêmicos e profissionais

AGRADECIMENTOS

Ao professor Mauro, não só pela orientação em trabalhos acadêmicos, mas por todo apoio e por toda a sabedoria transmitida ao longo dos últimos anos em suas disciplinas.

A toda minha família, por todo apoio emocional e conselhos, não somente nos anos de faculdade, mas sempre.

Aos meus amigos da vida pessoal e profissional, por todos os momentos de felicidade e apoio incondicional, além de todos os ensinamentos.

“Imagination is more important than knowledge.
Knowledge is limited. Imagination encircles the world.”
Albert Einstein (1879-1955)

RESUMO

Atualmente, um dos grandes problemas enfrentados no mercado de meios de pagamento está relacionado com as altas perdas financeiras associadas a transações *online*. Os principais fatores que contribuem com a perda financeira em transações desse tipo estão relacionadas a transações declinadas (recusadas) e com transações fraudulentas. Sabe-se que não há uma maneira totalmente assertiva de se reconhecer se uma transação é fraudulenta ou não, ou se uma transação foi corretamente declinada ou não. Entretanto, existem técnicas que permitem, com alto nível de acurácia, classificar transações como fraude ou não-fraude. Assim, o objetivo do presente trabalho foi desenvolver um método que permita classificar, com maior acurácia, quais transações são fraudulentas e quais não são e, assim, justificar a recusa de transações que sejam identificadas como fraudes em *e-commerce* – ou seja, transações não legítimas em ambiente *online*. As transações erroneamente declinadas por fraude (falsos positivos) e erroneamente aceitas sendo que eram fraude (falsos negativos), pelos métodos atualmente utilizados, são os motivos que contribuem para as altas perdas relativas a declínio transacional e a fraude. Assim, o método tem como objetivo classificar, com alta assertividade, qual transação é fraudulenta e qual não é, de modo a prevenir falsos positivos e falsos negativos. Tal análise classificatória é feita com base em variáveis presentes na *string* de transação de cartão de crédito, a qual contém campos como horário da transação, valor, local, moeda, número do cartão etc. Para tanto, abordagens que utilizam muito bem variáveis como essas são os métodos de aprendizado de máquina. Com essas técnicas, consegue-se aprender a prever padrões com base em dados históricos. No presente trabalho, focou-se nos métodos de aprendizado supervisionado, comparando-se a abordagem de regressão logística à de *random forest*. Procurou-se, então, verificar qual das duas técnicas apresenta melhor desempenho em um *dataset* de natureza desbalanceada, como as de transações de cartões de crédito, de modo a servir como base para um modelo de detecção de fraudes. Como resultado, o método *random forest* foi o que teve melhor desempenho em termos das métricas apresentadas – com uma acurácia de 99,94% (1,72% maior que o método de regressão logística), AUC-ROC de 93,67% e AUC-PR de 84,02%.

Palavras-chave: fraude, cartão de crédito, transações, aprendizado de máquina, *machine learning*, regressão logística, *random forest*

ABSTRACT

One of the major problems facing the payment industry nowadays is related to the high financial losses associated with online credit card transactions. The main factors that contribute to the financial loss in such transactions are losses related to declined (refused) transactions and fraudulent transactions. It is known that there is no fully assertive way of recognizing whether a transaction is fraudulent or not, or whether a transaction has been properly declined or not. However, there are techniques that allow, with a certain level of accuracy, to classify transactions as fraud or non-fraud. Therefore, the objective of the present work was to develop a method to classify, with greater accuracy, which transactions are fraudulent, and which are not, and thus justify the refusal of transactions that are fraudulent. The transactions erroneously declined due to fraud (false positives) and frauds erroneously accepted (false negatives), by the methods currently used, are the reasons that contribute to the high losses related to transactional decline and fraud. Hence, the method to be developed needs to classify, with an acceptable assertiveness, which transactions are fraudulent, and which are not, and so have low false positives and false negatives. Such classification analysis is based on variables present in the credit card transaction string, which contains fields such as the time the transaction was made, value, location, currency, card number (*PAN*) etc. To do so, approaches that use very well variables such as these are machine learning methods. With machine learning techniques, it is possible for the algorithms to learn how to predict patterns based on historical data. The present study focused on the supervised learning methods, and the aim was to compare the logistic regression approach with the random forest and verify which of the two techniques best performs in an unbalanced dataset, such as the credit card transactions, in order to serve as the basis for a fraud detection model.

Keywords: fraud, credit card, transactions, machine learning, logistic regression, random forest

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

ABECS	Associação Brasileira das Empresas de Cartões de Créditos e Serviços
ABRINQ	Associação Brasileira dos Fabricantes de Brinquedos
AFD	*Automated Fuel Dispenser*
ARQC	*Authorization Request Cryptogram*
ATM	*Automated Teller Machine*
AUC	*Area Under Curve*
AVS	*Address Verification System*
BACEN	Banco Central do Brasil
BP	*Basis Point*
BIN	*Bank Identification Number*
CAM	*Card Authentication Method*
CAT	*Cardholder Activated Terminal*
CDA	*Combined-Dynamic Data Authentication*
CNP	Cartão Não Presente
CP	Cartão Presente
CVC	Código de Verificação do Cartão
CVM	*Cardholder Verification Method*
DDA	*Dynamic Data Authentication*
DRE	Demonstração do Resultado do Exercício
FN (*FN*)	Falso Negativo (*False Negative*)
FP (*FP*)	Falso Positivo (*False Positive)*
ICA	*Interbank Card Association*
INSS	Instituto Nacional do Seguro Social
IP	*Internet Protocol*
K-NN	*K-Nearest Neighbors*
MO/TO	*Mail Order / Telephone Order*
ONG	Organização Não Governamental
PAN	*Primary Account Number*

P&L	*Profit and Loss Statement*
PCA	*Principal Component Analysis*
PdV	Ponto de Venda
PIN	*Personal Identification Number*
PKE	*PAN-Key Enter*
PoS	*Point of Sale*
PR	*Precision-Recall*
PVV	*PIN Verification Value*
ROC	*Receiver Operating Characteristic*
SDA	*Static Data Authentication*
ULB	*Université Libre de Bruxelles*
USB	*Universal Serial Bus*
USD	*United States Dollar*
VN (*TN*)	Verdadeiro Negativo (*True Negative*)
VP (*TP*)	Verdadeiro Positivo (*True Positive)*

SUMÁRIO

1. INTRODUÇÃO

O presente trabalho de formatura tem seu conteúdo focado na detecção de transações de cartão de crédito fraudulentas, com ênfase nas transações realizadas especificamente em ambiente online, em que o portador do cartão não se encontra fisicamente no estabelecimento – ambiente também conhecido, em inglês, como *e-commerce*. Para tanto, são apresentadas a empresa e a área de atuação do autor como estagiário no segmento de meios de pagamento, assim como definido o problema, sua relevância, o objetivo do trabalho e sua estrutura.

1.1. Empresa e área de atuação profissional do autor como estagiário

Durante o desenvolvimento deste trabalho de formatura, o autor desempenhou seu programa de estágio em uma empresa global de tecnologia voltada ao mercado de meios de pagamento. Tal empresa opera uma massiva rede de processamento de pagamentos e apresenta, como objetivo principal, garantir a rapidez, fluidez, simplicidade e segurança das transações, por meio de sistemas inteligentes. Sendo assim, visa-se conectar todos os *stakeholders* atuais desse mercado, como consumidores, instituições financeiras, estabelecimentos, governos e outros negócios globais, viabilizando o uso de formas eletrônicas de pagamento ao mesmo tempo em que os critérios de rapidez, fluidez, simplicidade e segurança das transações são atendidos. Para tanto, são desenvolvidas tecnologias que visam desde a melhoria de eficiência da operação de pagamento via cartões de crédito, débito e pré-pagos, até robustos produtos de prevenção e de detecção de fraude.

Competem à área de atuação do autor as atividades de *Customer Fraud Management*. Em termos gerais, esta área é responsável por gerar e analisar indicadores de autorização de transações e de fraude para diversos clientes, sendo estes *stakeholders* do mercado de meios de pagamento, principalmente entidades emissoras de cartões de crédito (bancos tradicionais, digitais e outras instituições financeiras). Assim, por meio desses números e dos resultados provenientes de sua análise, é possível identificar fraquezas e pontos estratégicos para atuação junto a essas entidades e, por fim, fornecer recomendações de melhoria, seja do ponto de vista operacional como também tático e estratégico, dependendo da situação. Reuniões periódicas são agendadas com os clientes de forma a engajar um acompanhamento sistemático dos índices de autorização e fraude dessas instituições.

1.2. Definição do problema

Um dos grandes problemas enfrentados pela área de *Customer Fraud Management* nos dias atuais, não só na empresa supracitada como também em outras empresas do mercado de meios de pagamento, está relacionado com as altas perdas financeiras associadas a transações de cartão não presente.

Antes de discutir esse tipo de transação, é importante citar que existem diversas categorias de transações por cartão. Uma dessas categorias é o canal da transação, a qual diz respeito ao modo como a transação foi feita do ponto de vista da inserção dos dados do cartão. Tendo em vista tal categoria, a mesma pode ser dividida nas seguintes classificações:

a) Cartão Presente (CP)

Significa que o cartão estava presente no estabelecimento comercial no ato da inserção de suas informações. Pode ser dividido em:

i) CP-PoS (por máquina de cartão, i.e., *Point of Sale*);
ii) CP-CAT-AFD (Automated Fuel Dispenser);
iii) CP-CAT-ATM (Automated Teller Machine);
iv) CP-CAT-Other (Outros);
v) CP-PKE-Fallback (PAN Key Entry Fallback).

b) Cartão Não Presente (CNP)

Significa que o cartão não estava presente no estabelecimento comercial no ato da inserção de suas informações. Pode ser dividido em:

i) CNP-E-Commerce (Comércio eletrônico);
ii) CNP-MO/TO (Mail Order / Telephone Order);
iii) CNP-Other (Outros);
iv) CNP-Recurring (número do cartão é armazenado para transações recorrentes).

Tendo em vista essas classificações, é possível agrupá-las e apenas referenciar o canal da transação em dois macrogrupos: CP (Cartão Presente) e Cartão Não Presente (CNP).

As grandes perdas associadas a transações do canal de Cartão Não Presente estão concentradas, principalmente, em transações a *e-commerce*, ou seja, transações em que as informações do cartão, seja de crédito, débito ou pré-pago,

além dos dados do portador do cartão, são inseridos em uma loja ou ambiente *online*. Assim, o portador do cartão realiza a transação remotamente, não estando fisicamente no estabelecimento comercial.

As perdas podem ser extratificadas em dois outros tipos principais de perdas, sendo eles:

a) Perdas por transações declinadas (não aprovadas);

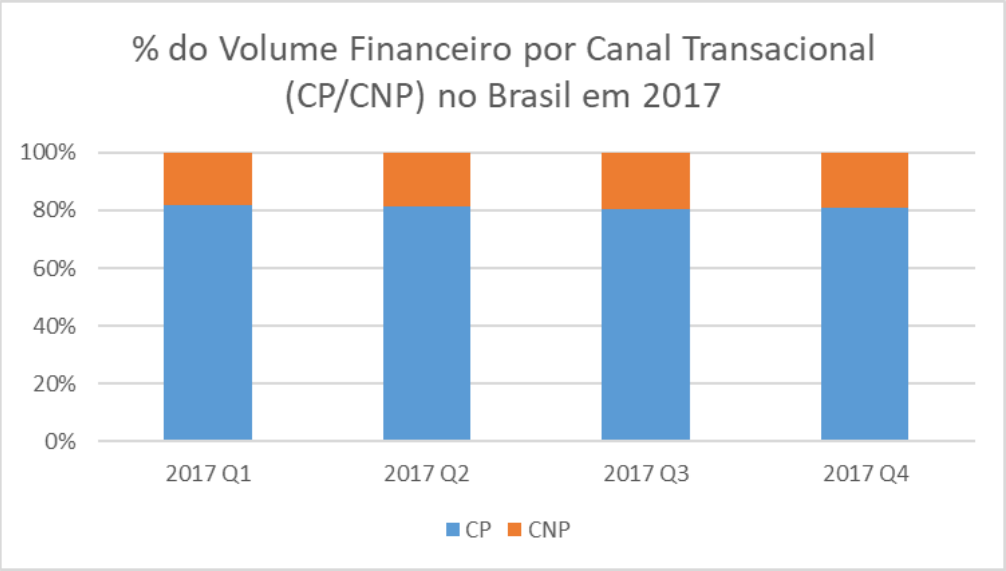
b) Perdas por transações fraudulentas.

Posto isso, a organização concentra muitos esforços em tentar entender como desenvolver estratégias que minimizem a taxa de declínio das transações em *e-commerce* (sendo o mesmo que aumentar a taxa de aprovação). Ademais, também vem estudando abordagens para a redução dos índices de fraude associados a transações não só de *e-commerce*, como também ao canal de Cartão Não Presente em geral.

1.3. Relevância do trabalho

Ao se analisar as transações de cartões de crédito do ponto de vista de canais transacionais, identifica-se uma grande disparidade das transações de cartão presente frente às transações de cartão não presente, em termos de volume (em USD). Nos trimestres de 2017 e 2018, aproximadamente 80% das transações se mantiveram no canal de Cartão Presente e 20% no canal de Cartão Não Presente, conforme figura 1 abaixo.

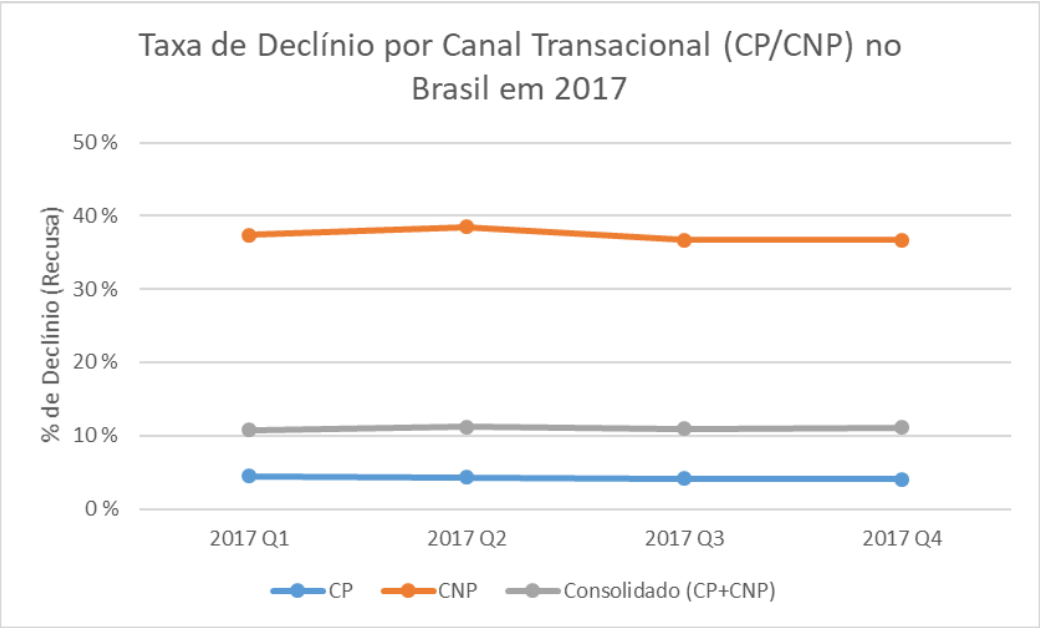
Figura 1 - Porcentagem de Volume Transacional (USD) trimestral por canal transacional no Brasil



Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Ainda assim, mesmo representando apenas 20% do volume total transacionado nos trimestres de 2017 e 2018, as transações no canal de Cartão Não Presente apresentam as maiores taxas de declínio (recusa) de transação, com média de 36% do início de 2017 até o final de 2018 frente a uma média de 4% para o canal de Cartão Presente nesse mesmo período e 11% para o canal consolidado (CP+CNP):

Figura 2 - Taxas de Declínio (%) trimestrais por canal transacional no Brasil



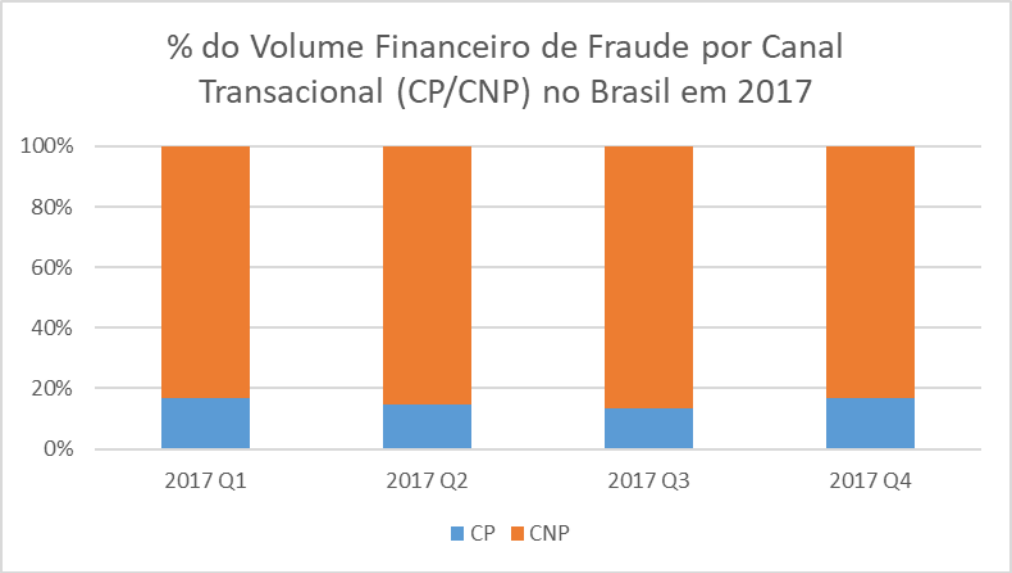
Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Como a maior parte das transações se concentra em CP, então faz sentido que a taxa de canal consolidado (CP+CNP) fique mais próxima da taxa de CP.

A taxa de declínio é calculada pela divisão da quantidade de transações declinadas (recusadas pelo terminal por algum motivo) pela quantidade de transações processadas, num dado período.

Ademais, além de ser a classe transacional com a maior taxa de declínio, também apresenta as maiores perdas com fraude. Aproximadamente 83% do volume de fraude está concentrada no canal de Cartão Não Presente e 17% no canal de Cartão Presente:

Figura 3 - Porcentagem de volume financeiro de fraude (USD) trimestral por canal transacional no Brasil



Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Sendo assim, o problema é de extrema relevância, uma vez que, mesmo representando menor parte no volume transacionado, as transações do canal de Cartão Não Presente ainda assim apresentam maiores perdas por declínio e por fraude frente às transações do canal de Cartão Presente.

1.4. Objetivo do trabalho

Com base na definição do problema e na sua relevância acima apresentados, pode-se afirmar que os principais elementos que constituem a perda financeira com transações do canal de Cartão Não Presente são as perdas com transações declinadas e com transações fraudulentas. Indo um pouco mais a fundo no problema, reconhece-se que não há uma maneira assertiva de se reconhecer se uma transação é fraudulenta ou não, ou se uma transação foi corretamente declinada ou não.

Dessa forma, o objetivo do presente trabalho é desenvolver um método que permita classificar, com maior acurácia, quais transações de cartão de crédito são fraudulentas e quais não são.

As transações erroneamente declinadas por fraude (falsos positivos) e erroneamente aceitas sendo que eram fraude (falsos negativos), pelos métodos atualmente utilizados, são os motivos que contribuem para as altas perdas relativas a declínio transacional e à fraude. Assim, o método a ser desenvolvido necessita

classificar, com assertividade, qual transação é fraudulenta e qual não é, de modo a prevenir falsos positivos e falsos negativos. Deseja-se que tal análise classificatória seja feita com base em variáveis presentes na *string* transacional, isto é, na cadeia de texto que é gerada em todas as transações e que contém campos como o horário em que a transação foi feita, valor, local, moeda, número do cartão etc.

Para tanto, uma abordagem que se utilizaria muito bem de variáveis como essas seria o *machine learning*. Com técnicas de aprendizado de máquina, um algoritmo pode aprender com ele mesmo e com base em dados históricos que servem de *input* a ele. Assim, de forma a melhorar a assertividade ao decidir se uma transação é fraudulenta ou não e se deve ser declinada, um método baseado em *machine learning* se utilizaria do carregamento de um *dataset* com inúmeras variáveis presentes na *string* transacional e, por conseguinte, aprenderia (ou "treinaria") com esses dados, indentificando padrões e sendo capaz de julgar a legitimidade das transações a serem posteriormente analisadas quando tal método estiver em fase de funcionamento (pós-"treinamento"). A base de aprendizado dos métodos de *machine learning* são os métodos estatísticos e os mais reconhecidos são os de regressão, *clustering* e classificação.

Existe uma grande quantidade de métodos de *machine learning*. No presente trabalho, foca-se nos métodos de aprendizado supervisionado, sendo esta categoria de métodos explicada na revisão bibliográfica. Dessa categoria, as abordagens mais conhecidas são as de redes neurais, árvores de decisão, *random forest* e regressão logística.

Neste trabalho, o objetivo é comparar dois desses métodos – *random forest* e regressão logística – com base em critérios e indicadores pré-definidos como acurácia e AUC, que são mais bem abordados na metodologia. Após a comparação, determinou-se o melhor método entre os dois e o mesmo foi considerado como base de um modelo a ser utilizado para detecção de fraude, com base em diversas variáveis da *string* transacional (chamadas, no modelo, de *features*) e em outras variáveis calculadas utilizando as obtidas na *string* transacional.

Com a aplicação de tal método em um *dataset* de exemplo, objetiva-se também a obtenção de números e indicadores como falsos positivos, falsos negativos e outros a serem explorados na revisão bibliográfica.

1.5. Estrutura do trabalho

Com a finalidade de atingir o objetivo definido, o trabalho possui a seguinte estrutura:

a) Primeiro capítulo – Introdução: diz respeito à contextualização do âmbito em que está situado o problema em questão, à abordagem do problema e sua pertinência e relevância, além do que foi feito para prosseguir com a solução deste;
b) Segundo capítulo – O mercado de meios de pagamento: introduz o mercado de meios de pagamento ao leitor, provendo um panorama e detalhes acerca dos termos e definições que permeiam este mercado. Após a leitura deste capítulo, o leitor ficará mais familiarizado com o mercado em questão e estará apto a se aprofundar mais nos detalhes do problema;
c) Terceiro capítulo – Transações CNP e o cenário *e-commerce*: aprofunda-se mais na parte do mercado de meios de pagamento em que o problema se encontra, ou seja, são fornecidos mais detalhes acerca do cenário brasileiro de *e-commerce* e também acerca da relevância do canal CNP frente ao canal CP. É um aprofundamento do que foi exibido no capítulo introdutório;
d) Quarto capítulo – Revisão bibliográfica: fornece as explicações necessárias para o entendimento do *machine learning*, suas categorizações e técnicas, além de detalhes acerca do funcionamento dessas técnicas. Ademais, aborda os indicadores de performance mais importantes no que diz respeito à detecção de fraude;
e) Quinto capítulo – Metodologia: aborda tanto os passos que foram necessários para a determinação do problema, sua estratificação e breve identificação da causa dos estratos desse problema, como também os passos necessários para se atingir o objetivo determinado no capítulo introdutório, ou seja, o que deve ser seguido em termos da solução do problema em questão, desde a obtenção e tratamento dos dados a serem utilizados pelos dois métodos testados até os etapas dos algoritmos de *machine learning* em si;
f) Sexto capítulo – Aplicação da Metodologia: é o capítulo que exibe a aplicação de todo aparato explorado na etapa de metodologia;
g) Sétimo capítulo – Resultados: este capítulo mostra e explana todos os resultados obtidos mediante a aplicação da metodologia;
h) Oitavo capítulo – Conclusão: sintetiza todas as etapas do trabalho, passando pela obtenção dos dados, tratamento, metodologia e aplicação da mesma, além de

conter uma breve análise das respostas obtidas por meio da aplicação metodologia apresentada e suas principais contribuições, como também ideias e recomendações para otimizações contínuas e/ou futuras do modelo, sobretudo considerando possíveis inclusões de novas variáveis ao modelo e outras modificações que sejam provavelmente possíveis com o avanço da tecnologia e com o estudo de novas abordagens;

i) Bibliografia: apresentar a bibliografia utilizada para consultas e/ou citações no presente trabalho.

2. O MERCADO DE MEIOS DE PAGAMENTO

Nesta seção do trabalho, é explorado o mercado de meios de pagamento – explorando sua história e apresentando suas características. Explica-se, com um maior nível de detalhe, as terminologias envolvendo os cartões de crédito, as características dos mesmos, assim como estratégias de gestão de autorizações e de fraude, além de conceitos específicos utilizados nesse mercado.

2.1. O mercado de meios de pagamento

Antes de adentrar aos dados históricos e estatísticas do mercado de pagamento, se faz fundamental apresentar a história do mercado em si e do próprio cartão de crédito, como também apresentar todos os *players* participantes e influenciadores do segmento e suas relações e responsabilidades.

2.1.1. A história do cartão de crédito

Em consulta a materiais internos utilizado em treinamentos fornecidos pela empresa em que o autor estagiou, foi possível obter informações acerca da história do cartão de crédito.

Já no início de 1900, as empresas de petróleo e lojas de departamento americanas ofereciam aos seus clientes cartões de cortesia ou placas de cobrança que permitiam que comprassem agora e pagassem mais tarde. Esses primeiros cartões só podiam ser usados nas estações de serviço ou lojas que os emitiam, e os clientes precisavam pagar seus saldos de conta no final de cada mês.

A Bloomingdale's criou o primeiro cartão de crédito da conta rotativa em 1938, e em 1950, um empresário de Nova York, que nunca mais quis ser pego sem dinheiro no final de um jantar de negócios, ajudou a criar o primeiro cartão de viagem e entretenimento. O episódio exemplificador desse surgimento do primeiro cartão ocorreu em um restaurante na cidade de Nova Iorque, quando o executivo Frank MacNamara, na presença de outros executivos financeiros, percebeu que havia esquecido seu dinheiro e seu talão de cheques ao pagar a conta. Após o incidente, Frank teve a ideia de criar um cartão baseado em confiança, funcionando como uma garantia de pagamento posterior e sendo portado por cliente cujos donos de estabelecimentos acreditavam serem confiáveis e quitarem suas dívidas em dia. Após a criação deste modelo primordial de cartão, no mesmo ano, Frank lançou o papel-cartão Diners Club International, aceito em apenas 27 restaurantes e usado, na época,

somente por pessoas importantes (aproximadamente 200 pessoas). Como passar dos anos, o cartão começou a atrair mais adeptos e ser aceito em vários estabelecimentos.

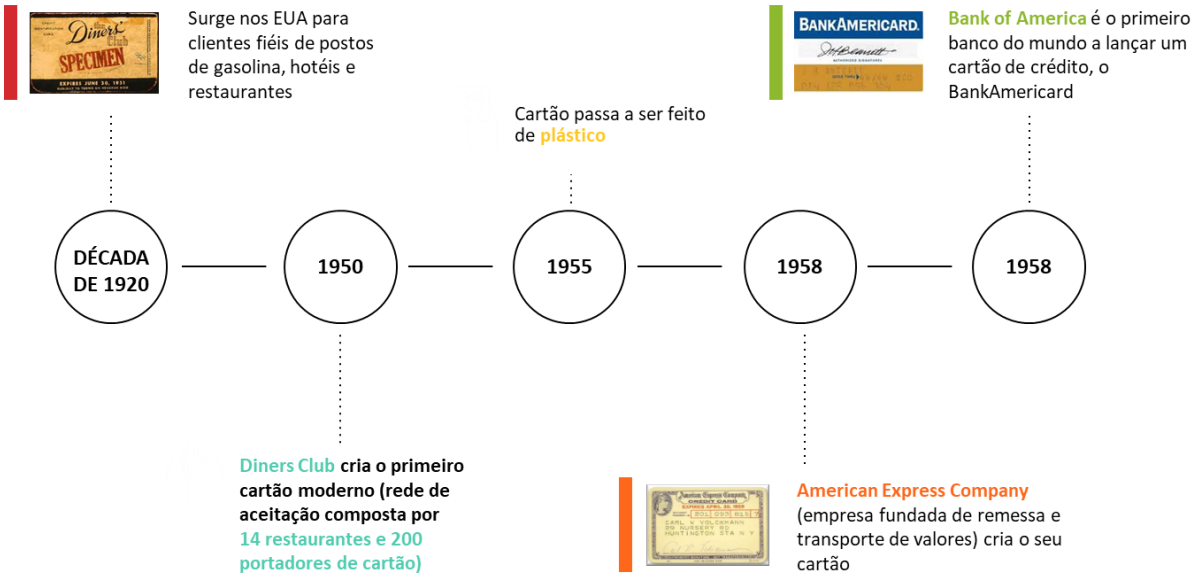
Em 1955, passou a ser confeccionado com plástico. Com o aumento do acesso financeiro e oportunidade, os americanos começaram a apreciar a conveniência de ter um cartão que poderia ser usado para fazer compras em várias lojas, em vez de ter uma conta separada com cada comerciante.

Durante a década de 1950, os bancos tornaram-se ativos no negócio de crédito ao consumidor. O Franklin National Bank em Long Island emitiu o primeiro cartão de crédito bancário em 1951 e, em meados da década, vários pequenos bancos seguiram a liderança de Franklin, embora com sucesso limitado.

Enquanto isso, o maior banco da Califórnia, o Bank of America, sediado em São Francisco, observou e aprendeu com os erros de outros e, em 1958, introduziu um cartão de crédito próprio, o BankAmericard. No início, os comerciantes hesitaram em aceitar o BankAmericard, mas quando o Bank of America colocou os cartões nas mãos de 60.000 californianos, as dificuldades chegaram a um impasse. Em apenas um ano, o número de comerciantes que aceitaram o cartão aumentou de 800 para 25000.

O negócio de cartão de crédito decolou e, na primavera de 1965, o Bank of America decidiu licenciar o BankAmericard para bancos de todo os Estados Unidos.

Figura 4 - Linha do tempo da história do cartão de crédito (Parte 1)

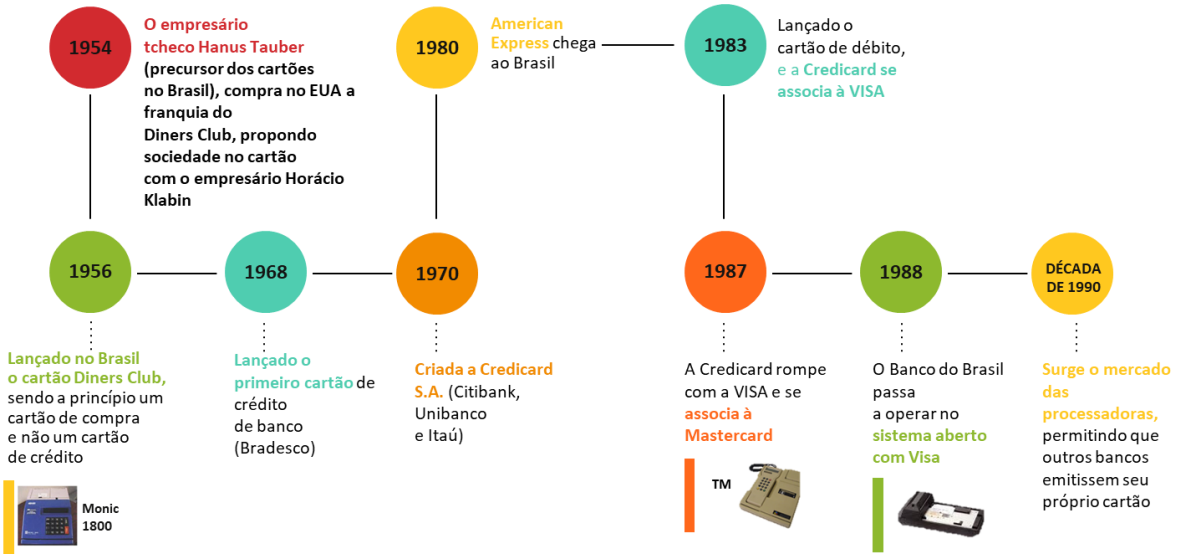


Fonte: elaboração do autor com dados fornecidos pela empresa onde estagiou

No que concerne o mercado de meios de pagamento brasileiro, o cartão de crédito chegou ao Brasil em 1956, graças ao empresário tcheco Habus Tauber, que adquiriu a franquia do Diners Club e propôs sociedade no cartão com o empresário Horácio Klabin. O cartão funcionava como um cartão de compra e não de crédito. O primeiro cartão de crédito no Brasil, por sua vez, foi o cartão do Bradesco, em 1968. Já em 1970, é criada a Credicard como uma união do Citibank, Unibanco e Itaú. Em 1980, a American Express chega ao Brasil e, em 1983, é lançado o cartão de débito concomitantemente com a associação entre Credicard e VISA. Em 1987, contudo, houve o rompimento entre VISA e Credicard, e esta se associa à Mastercard. Em 1988, o Banco do Brasil passa a operar no sistema aberto com a VISA.

Entrando na década de 1990, surge o mercado das peocessadoras, permitindo que outros bancos emitissem seus próprios cartões. Em 1991, ocorre a abertura para o cartão internacional e, em 1994, o plano real acentua o crescimento do produto e surgem novos tipos de cartão, como INSS, Afinidade, Loja, Ticket Combustível, Alimentação, Transporte etc.

Figura 5 - Linha do tempo da história do cartão de crédito (Parte 2)



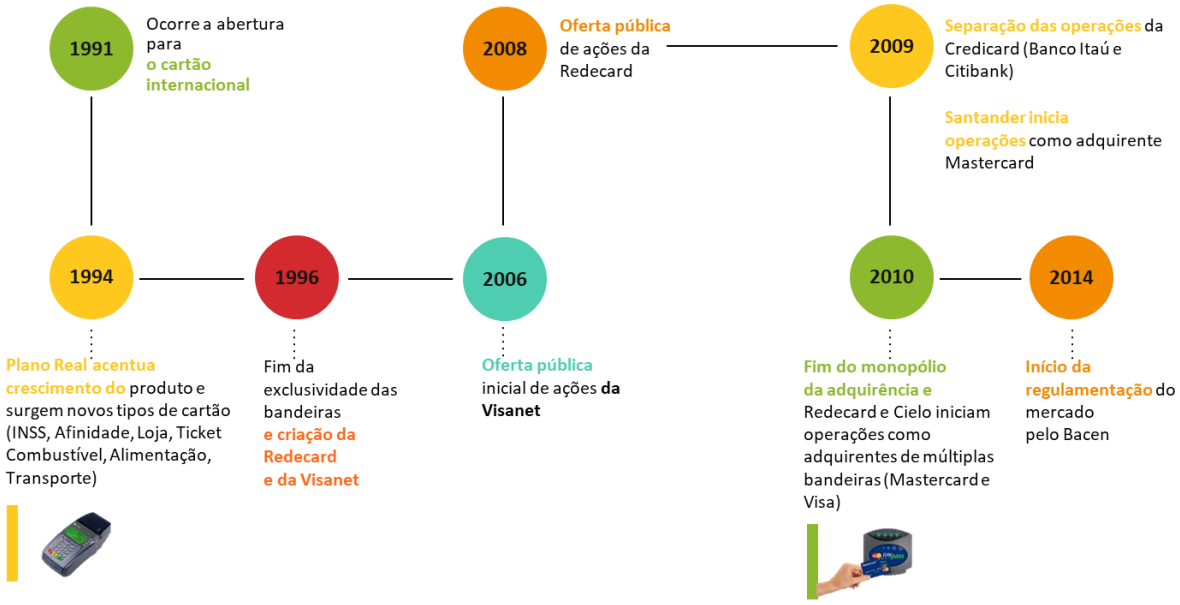
Fonte: elaboração do autor com dados fornecidos pela empresa onde estagiou

Em 1996, ocorreu o fim da exclusividade das bandeiras e a consequente criação da Redecard e da Visanet. Já em 2006, ocorreu a oferta pública inicial de ações da Visanet e, em 2008, da Redecard. Em 2008, houve a separação das operações da Credicard (Banco Itaú e Citibank) e, em 2010, o fim do monopólio de

adquirência de transações, com a Redecard e Cielo iniciando suas operações como adquirentes de transações de cartões de múltiplas bandeiras (Mastercard e VISA).

Por fim, em 2014, o BACEN iniciou a regulamentação do mercado de meios de pagamento.

Figura 6 - Linha do tempo da história do cartão de crédito (Parte 3)



Fonte: elaboração do autor com dados fornecidos pela empresa onde estagiou

2.1.2. Modelos de negócio e stakeholders

Ao se tratar do mercado de meios de pagamento, existem três modelos de negócio em vigência no mundo. Antes de introduzir os modelos, contudo, é interessante permear os principais *stakeholders* presentes nesse mercado:

- *Acquirer*: é a entidade adquirente (também conhecida como credenciadora), ou seja, a instituição que fornece toda a rede e infraesturtura de captura de transação para um estabelecimento comercial. Fisicamente, a captura da transação ocorre por meio de PoS (*Point of Sale*, ou PdV – Ponto de Venda), ou seja, a máquina de passar/inserir/encostar o cartão, dependendo do seu tipo;
- *Bandeira*: fornece as “regras do jogo” a todos os *stakeholders* que operam em sua rede de pagamentos, além de ter sua marca estampada nos cartões emitidos e que transacionam na rede;
- *Cardholder*: é o portador do cartão;

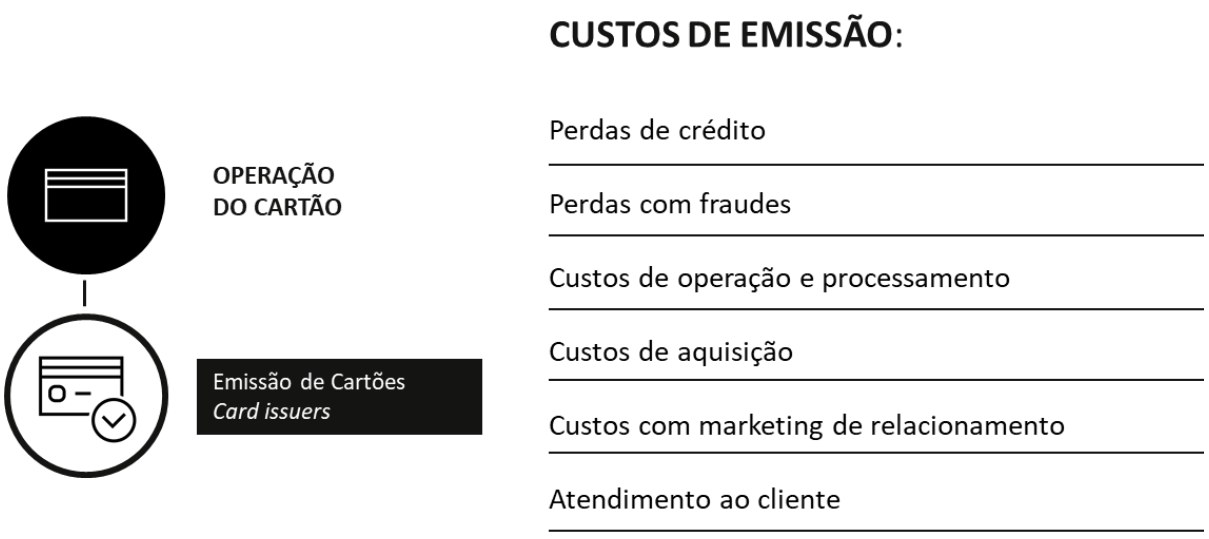
- *Issuer*: é a entidade que emite o cartão. Normalmente são bancos ou instituições financeiras;
- *Merchant*: é o estabelecimento comercial, o qual vende o produto ou serviço adquirido pelo portador do cartão mediante a sua utilização em uma máquina provida pelo adquirente;
- *Service Providers*: são instituições que colaboram com os outros *stakeholders* ao fornecer serviços que melhoram a operação deles. Exemplos: *Payment Facilitators* (Paypal), Processadoras (Conductor) etc.

Os modelos existentes são o de duas, três e o de quatro partes. São chamados assim pois cada parte representa um *stakeholder*, ou parte envolvida com o modelo.

a) Modelo de duas partes

Neste modelo, existem apenas dois *stakeholders*: o estabelecimento (ou *merchant*) e o emissor de cartões (ou *card issuer*) são uma entidade só, e o portador do cartão. Quando se fala de emissão de cartões, existem alguns custos associados a esta atividade. No caso do modelo de duas partes, tais custos estão dispostos a seguir:

Figura 7 - Modelo de duas partes



Fonte: elaboração do autor com dados fornecidos pela empresa onde estagiou

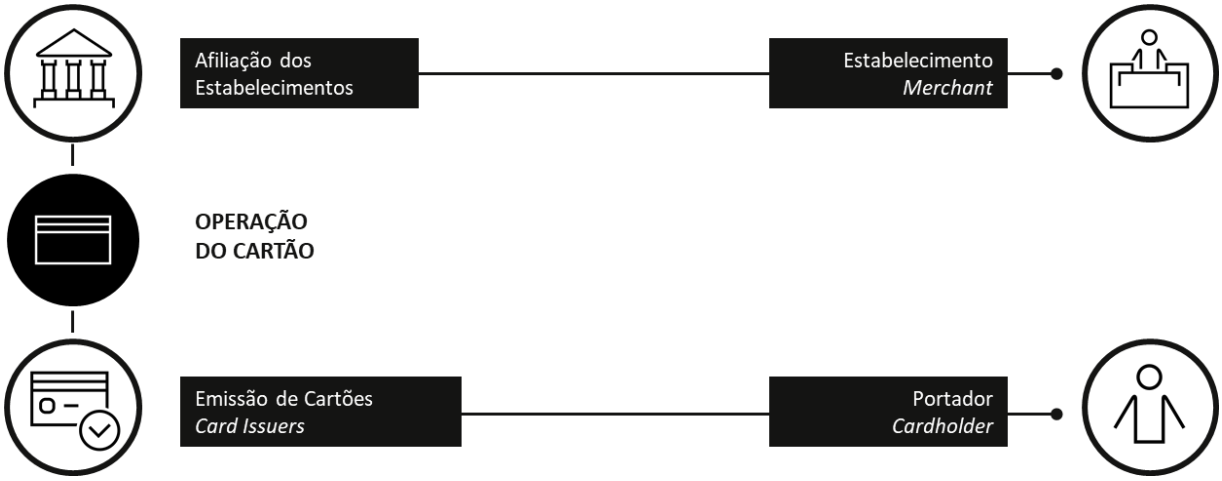
Um exemplo deste modelo é a Macys, uma loja de departamentos norte-americana. Nesse caso, a loja é um estaabelecimento comerical, mas possui um setor

de emissão de seu próprio cartão. Tal setor também provê o seu próprio sistema financeiro. Contudo, não se sabe se o cliente pagará a conta no final do mês, pois não existe garantia de pagamento. Nesse caso, as lojas que seguem esse modelo podem fazer a venda de seus “recebíveis” a um banco e o banco assume o risco e a dívida, cobrando uma taxa de desconto. A garantia de pagamento custa dinheiro e o banco a cobra através desta taxa de desconto.

b) Modelo de três partes

Neste modelo, existem três *stakeholders*: o estabelecimento comercial, o portador e operador. A operação (emissão dos cartões, processamento e adquirência das transações) é feita por uma só entidade, representada por uma associação de estabelecimentos. Tal associação usa a receita que ganha do do estabelecimento comercial associado para pagar parcialmente o custo de emissão dos cartões (*issuing cost*). Ademais, a associação precisa controlar e manejar todos os riscos do sistema. Este modelo tem como exemplo a Amex e a Diners.

Figura 8 - Modelo de três partes



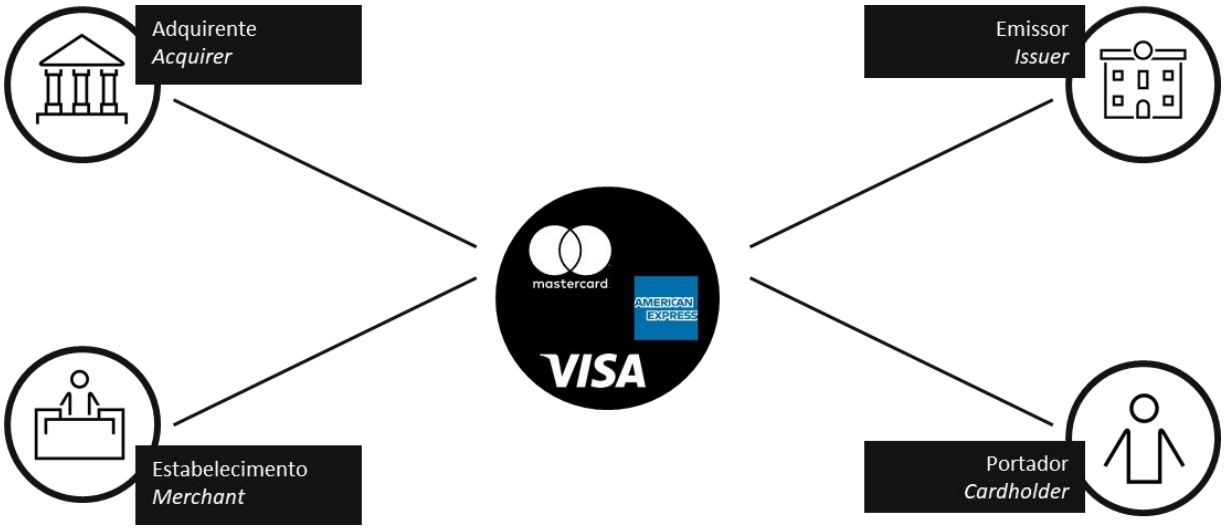
Fonte: elaboração do autor com dados fornecidos pela empresa onde estagiou

c) Modelo de quatro partes

É o modelo de negócio mais usado no mercado de meios de pagamento. Trata-se de um modelo com quatro *stakeholders*: emissor do cartão, portador do cartão, estabelecimento comercial e adquirente (detentora das “maquininhas”, responsável por capturar as transações para os estabelecimentos afiliados). Os quatro

stakeholders são unidos por uma única entidade, a qual dita as “regras do jogo”. Tal entidade é conhecida como bandeira.

Figura 9 - Modelo de quatro partes



Fonte: elaboração do autor com dados fornecidos pela empresa onde estagiou

Neste modelo, os emissores e adquirentes não podem ser uma mesma entidade. Alguns benefícios deste modelo são:

- Ganho de escala, pois pode ser usados por milhares de emissores e adquirentes em todo o mundo;
- Viabiliza maior competição no mercado.

A seguir, mostra-se alguns exemplos de empresas que atuam como bandeiras, como emissores e como adquirentes:

Figura 10 - Participantes do modelo de negócio de quatro partes

Fonte: elaboração do autor

Abaixo, tem-se mais detalhes acerca das funções de cada uma das entidades presentes neste modelo:

Figura 11 - Bandeiras

Fonte: elaboração do autor

Figura 12 - Adquirentes

Fonte: elaboração do autor

Figura 13 - Emissores

Fonte: elaboração do autor

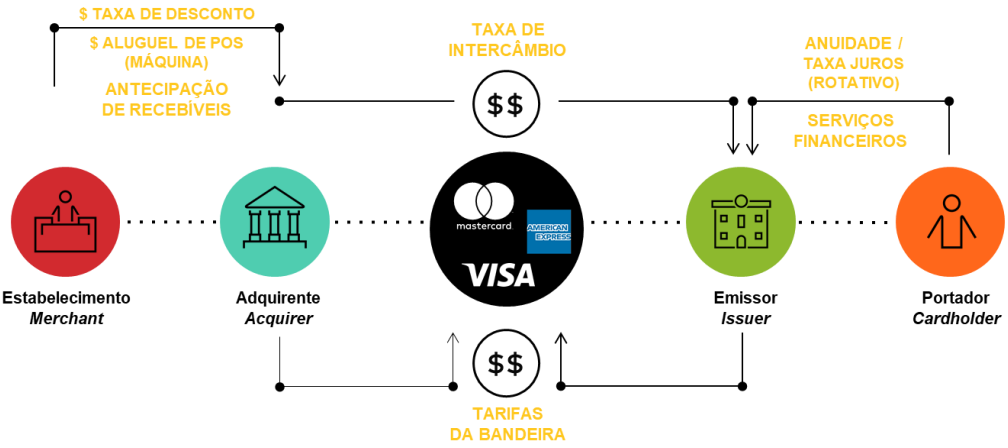
Ademais, há as entidades cooperativas (*service providers*), que não atuam como um *stakeholder* primário (ou essencial) no modelo, mas sim como instituições facilitadoras do mercado de meios de pagamento. Uma dessas entidades é a processadora. Alguns exemplos de processadoras se encontram a seguir:

Figura 14 - Processadoras

Fonte: elaboração do autor

No tocante às movimentações financeiras, no modelo de quatro partes, por cada uma das entidades primárias participantes do modelo, tem-se o seguinte:

Figura 15 - Movimentações financeiras do modelo de negócio de quatro partes



Fonte: elaboração do autor com dados fornecidos pela empresa onde estagiou

Nesse caso, para utilizar a estrutura de adquirência, o estabelecimento deve pagar ao adquirente uma taxa de desconto acrescida do aluguel do POS (*Point-Of-Sale*, ou Ponto-De-Venda, ou seja, a “maquininha” de cartão). Para poder utilizar o cartão de um emissor, o portador deve arcar com taxa de anuidade e de juros,

podendo optar por pagar, também, por serviços financeiros oferecidos pelo banco emissor. O adquirente, por sua vez, deve arcar com a taxa de intercâmbio para poder aceitar o cartão de um determinado emissor em sua infraestrutura de adquirência. Por fim, para o adquirente poder usar a infraestrutura e tecnologias de rede de transação e para o emissor também poder usar essa mesma rede e o nome da bandeira em seus cartões, existem tarifas que devem ser pagas a tal bandeira.

2.1.3. Tipos de cartão

Existem diversos tipos de cartão no mercado de meios de pagamento. Os mais conhecidos são os cartões de crédito, débito, corporativo e pré-pago. A figura a seguir exibe outros tipos de cartão também existentes nesse mercado.

Figura 16 - Tipos de cartões

Fonte: elaboração do autor

Dois tipos menos conhecidos de cartão são o co-branded e o afinidade:

- Co-branded: duas marcas fortes que vendem o cartão. Trata-se de um *business* para ambas as empresas. Exemplo: TAM Fidelidade;
- Afinidade: uma das marcas vende a sua marca para uso do banco. Exemplos: cartão Abrinq, cartão Rocerto Carlos, clubes de futebol, celebridades, ONGs. O termo afinidade diz respeito à afinidade do consumidor com a marca.

2.2. Operações

Ao se tratar de operações, são apresentadas as terminologias relativas ao cartão de crédito e sua estrutura, assim como os processos, conceitos e estratégias relevantes para a autorização de transações de cartão.

2.2.1. Estrutura do cartão e terminologias

Existem terminologias que são fundamentais de serem compreendidas ao se estudar operações envolvendo cartões:

a) ICA: em termos gerais, identifica o banco emissor do cartão. Um emissor pode ter mais de um ICA, representando distintos segmentos.

Figura 17 - Terminologia ICA

Fonte: elaboração do autor

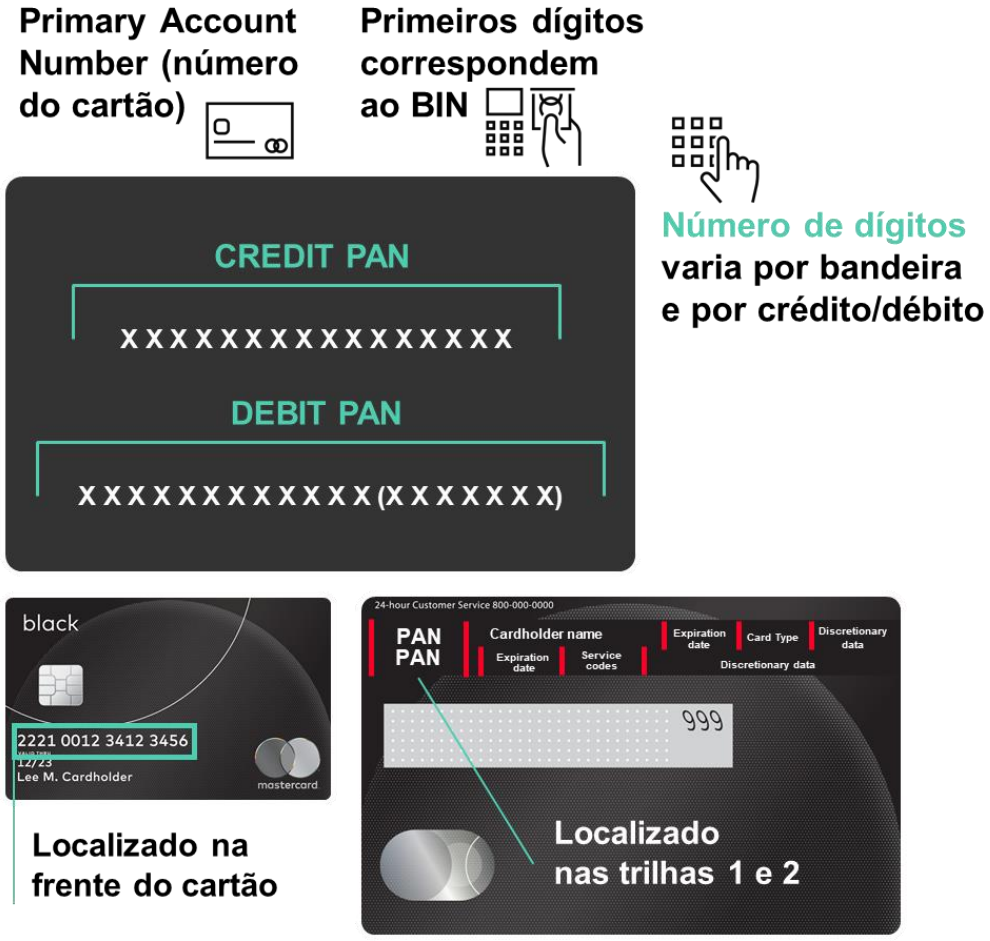
b) BIN: identifica qual tipo de produto do banco emissor o cartão representa, sendo conectado a um ICA do banco.

Figura 18 - Terminologia BIN

Fonte: elaboração do autor

c) PAN: é o número do cartão

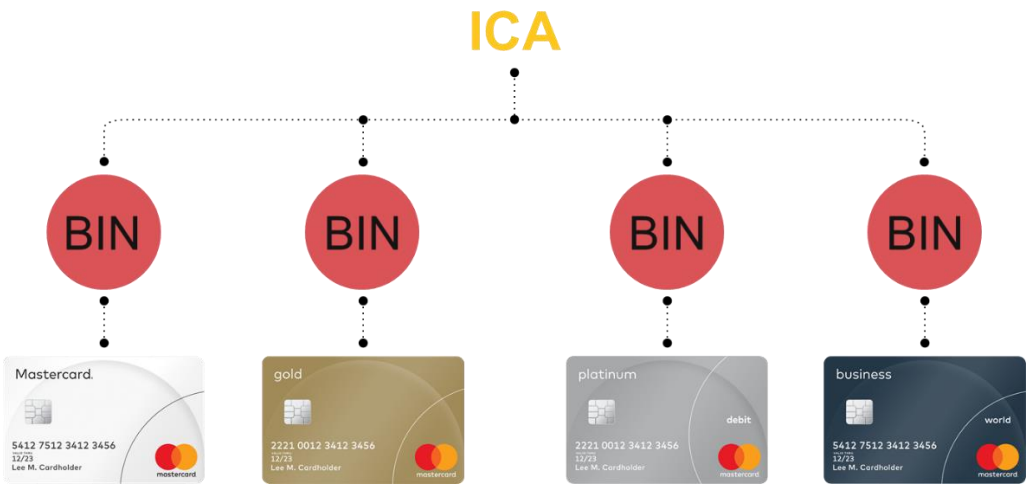
Figura 19 - Terminologia PAN



Fonte: elaboração do autor

Tais terminologias estão relacionadas conforme esquema exemplificado a seguir para a bandeira Mastercard e alguns de seus diferentes produtos. Cada PAN de cada produto (cartão) está atrelado a um BIN, e o conjuto de BINs está atrelado a um ICA:

Figura 20 - Esquematização da relação entre as terminologias ICA, BIN e PAN



Fonte: elaboração do autor

2.2.2. Processos de autorização

A autorização de uma transação tem seu início em um ponto de interação. Em transações de cartão presente, esses pontos são representados pelos PoS (*Point of Sale*, ou PdV – Ponto de Venda), em que as informações do cartão e do portador são captadas.

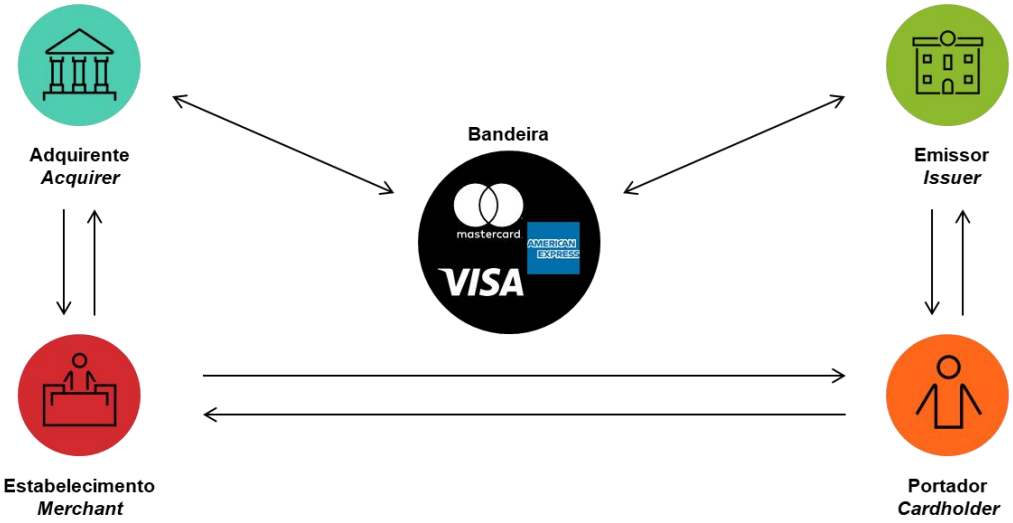
Figura 21 - Pontos de Venda de cartão

Fonte: elaboração do autor com dados

Já em transações de cartão não presente, as informações são captadas pelo próprio *gateway* associado ao *website* do estabelecimento comercial. Após o ponto de interação captar as informações do cartão, entra-se no fluxo da transação.

> *"O portador passa o cartão no terminal de pagamento do lojista. O lojista envia a transação para a credenciadora (adquirente), empresa responsável pelo credenciamento das lojas. Em alguns casos, há a participação do facilitador de pagamento, que faz a ponte entre o lojista e a credenciadora. A transação sai da credenciadora, passa pela rede da bandeira (que confere o BIN) e chega ao emissor do cartão. Após verificar o limite de crédito do portador (no caso de cartão de crédito) ou se há saldo em sua conta-corrente (no caso de cartão de débito), o emissor autoriza a transação. A transação retorna pela rede da bandeira até a credenciadora, que envia a resposta para o lojista. O lojista, por sua vez, conclui a compra." (ABECS, 2018, p. 18-19).*

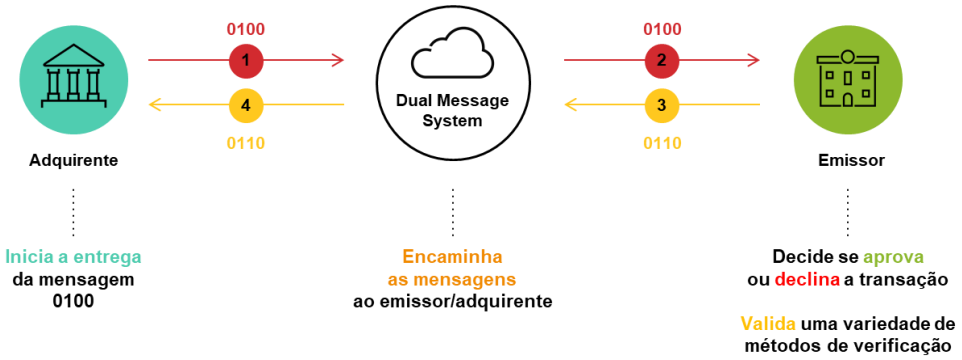
Figura 22 - Relações entre as partes no processo de autorização



Fonte: elaboração do autor

Para autorizar uma transação, primeiramente, o emissor analisa os dados constantes na mensagem de autorização (também conhecida como mensagem 0100) e toma a decisão de aprovar ou não, baseado nas suas políticas de autorização e nas da bandeira. Em seguida, o emissor avalia o saldo disponível na conta do portador, possíveis bloqueios, dados da conta, possibilidade de fraude etc. Após esse procedimento, o emissor fornece a resposta de autorização: aprovar, declinar, referir ou capturar cartão.

Figura 23 - Fluxo de autorização entre adquirente e emissor

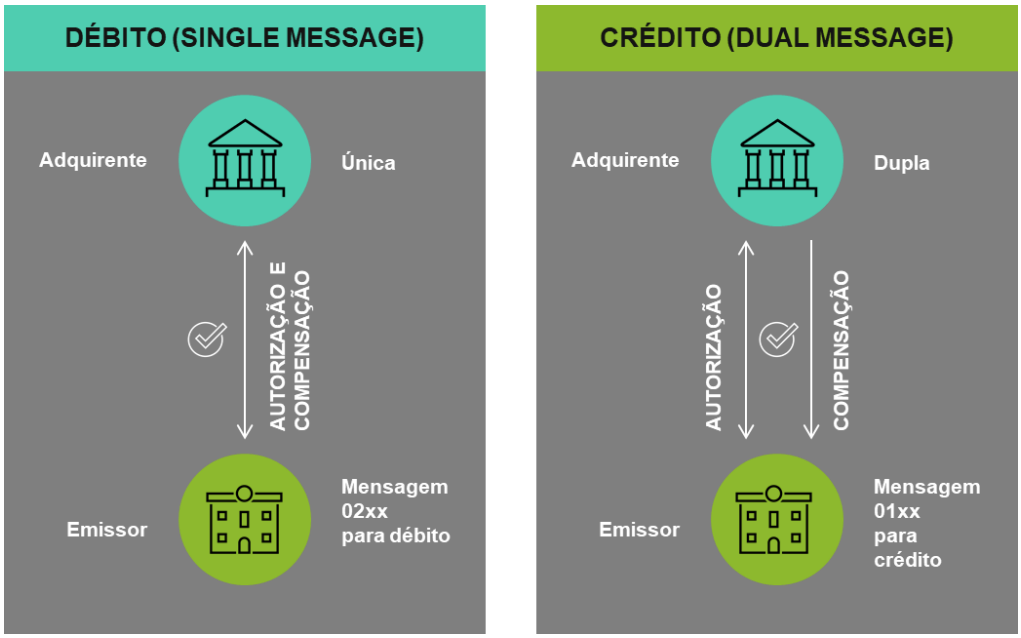


Fonte: elaboração do autor

Assim, a resposta de autorização é transmitida pelo emissor ao adquirente. Após o adquirente receber tal resposta, ele deve confirmar a autorização por meio de uma contra-resposta chamada de compensação.

Basicamente, é o processo pelo qual o adquirente apresenta, ao emissor, as transações, efetuadas em seus meios de captura. Para transações de débito, a autorização e a compensação viajam em uma única mensagem. Já para transações realizadas por cartão de crédito, há uma mensagem de autorização e uma mensagem para compensação, sendo este modelo conhecido como *dual message*.

Figura 24 - Mensagem simples de débito e mensagem dupla de crédito



Fonte: elaboração do autor

2.2.3. Canais e corredores de transação

Conforme dito na definição do problema, subcapítulo 1.2, existem diversas maneiras e critérios de categorização de transações no mercado de meios de pagamento.

Uma das segmentações possíveis de serem feitas é a de canal de transação. Uma transação pode ser iniciada com o cartão fisicamente em contato com o ponto de interação, com as informações do cartão e do portador sendo captadas no local do estabelecimento comercial. Dizemos que este tipo de transação pertence a um canal chamado de Cartão Presente (CP). Por outro lado, também existe o canal de Cartão Não Presente (CNP), em que o portador e o cartão não precisam estar presentes na localização física do estabelecimento, sendo os dados captados remotamente via internet.

As categorias são as seguintes:

a) Cartão Presente (CP)
   - i) CP-PoS (por máquina de cartão, i.e., *Point of Sale*);
   - ii) CP-CAT-AFD (Automated Fuel Dispenser);
   - iii) CP-CAT-ATM (Automated Teller Machine);
   - iv) CP-CAT-Other (Outros);
   - v) CP-PKE-Fallback (PAN Key Entry Fallback).

b) Cartão Não Presente (CNP)
   - i) CNP-E-Commerce (Comércio eletrônico);
   - ii) CNP-MO/TO (Mail Order / Telephone Order);
   - iii) CNP-Other (Outros);
   - iv) CNP-Recurring (número do cartão é armazenado para transações recorrentes).

Ademais, também existe uma segmentação chamada de corredor de transação, a qual diz respeito se o estabelecimento comercial, do qual determinado produto ou serviço está sendo adquirido, se encontra dentro ou fora do país em que se encontra o banco emissor do cartão.

Se um portador, com um cartão emitido no Brasil, por exemplo, realiza uma compra em um estabelecimento localizado no Brasil, diz-se que a transação é doméstica. Por outro lado, se o protador, com esse mesmo cartão, realiza uma compra

em um comércio eletrônico chileno, por exemplo, diz-se que o corredor dessa transação é *crossborder*. Tal corredor pode se subdividir em duas categorias:

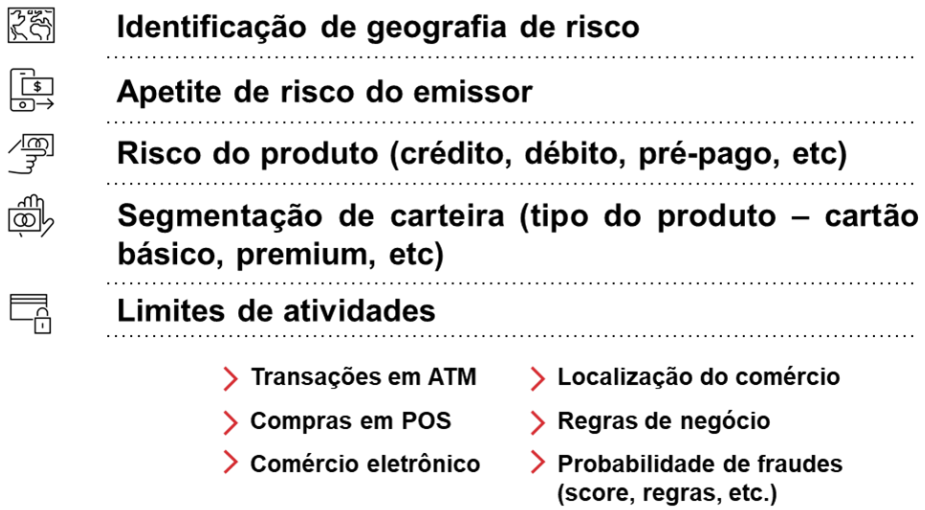
a) *Crossborder Inter*: portador se encontra em um determinado país, seja o país em que o cartão foi emitido ou algum outro, e realiza uma compra em um estabelecimento localizado em um país diferente deste em que o portador se encontra e que não seja o próprio país em que o cartão foi emitido. Exemplos:
    i. O portador, com um cartão emitido no Brasil, está no Chile e utiliza tal cartão para realizar uma compra em um comércio localizado no Canadá;
    ii. O portador, com um cartão emitido no Brasil, está no Brasil e utiliza tal cartão para realizar uma compra em um comércio localizado na Espanha.

b) *Crossborder Intra*: portador se encontra em um país estrangeiro, diferente do país em que o cartão foi emitido, e realiza uma compra em um estabelecimento neste mesmo país estrangeiro. Exemplos:
    i. O portador, com um cartão emitido no Brasil, está no Chile e utiliza tal cartão para realizar uma compra em um comércio localizado no Chile;
    ii. O portador, com um cartão emitido em Portugal, está na Suíça e utiliza tal cartão para realizar uma compra em um comércio localizado na Suíça.

2.2.4. Estratégias de autorização

Para decidir se uma transação deve ser autorizada ou não, existem distintas estratégias a serem consideradas. Tais estratégias variam dependendo de alguns fatores, como tipo de produto ou as atividades para as quais o cartão será destinado e seus limites. Alguns desses fatores estão dispostos a seguir:

Figura 25 - Estratégias de autorização



Fonte: elaboração do autor

2.3. Gestão de fraude

Este subcapítulo aborda a gestão de fraude, primeiramente definindo o que é fraude e *chargeback* e depois explorando as características operacionais e custos da fraude, assim como categorias, modalidades e estratégias de prevenção e detecção de transações fraudulentas.

2.3.1. Definição de fraude e *chargeback*

De acordo com o dicionário *Houaiss* da Língua Portuguesa, o conceito de fraude está relacionado a “qualquer ato ardiloso, enganoso, de má-fé, com o intuito de lesar ou ludibriar outrem, ou de não cumprir determinado dever; logro”. No mercado de meios de pagamento, especificamente ao se tratar do mercado de cartões, o conceito de fraude se faz visível, de maneira prática, quando o portador legítimo do cartão contata o emissor, informando o não reconhecimento de pelo menos uma das transações presentes em sua fatura. Tal processo é conhecido como geração de *chargeback*.

De acordo com Lau (2006) a definição de fraude está relacionada à distorção intencional da verdade ou de um fato, que busca em geral a obtenção de lucro ilícito.

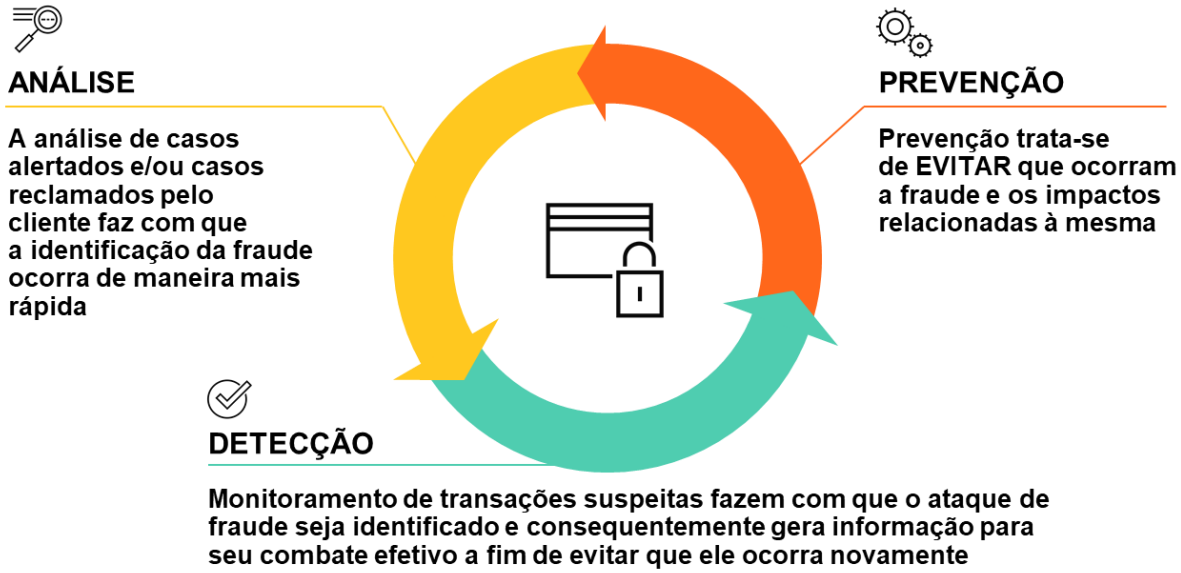
2.3.2. Características operacionais e custos

Em gestão de fraude, é comum de se deparar com um conceito chamado de ciclo de fraude. Basicamente, uma organização que começa suas atividades com

poucas noções de prevenção de fraude pode, certamente, começar investindo recursos em análise (e monitoramento) de casos alertados ou reclamados pelos clientes e em ferramentas capazes de detectar, corretamente e a um nível rico de detalhes, tais casos de fraude. Somente assim poderá construir um sistema robusto de prevenção à fraude.

Em outras palavras, os sistemas de identificação e detecção alimentam o sistema de prevenção com informações para o combate efetivo a tipos específicos de fraude. Desse modo, o sistema de prevenção estará apto a evitar a ocorrência de tipos de fraude iguais ou parecidos a casos que já ocorreram.

Figura 26 - Prevenção, análise e detecção



Fonte: elaboração do autor

Ademais, ao se tratar de cenários fraudulentos, existem diversas perdas associadas a eles, não apenas financeiras e operacionais. A figura abaixo fornece sumariza tais perdas:

Figura 27 - Perdas relacionadas a fraudes

Fonte: elaboração do autor

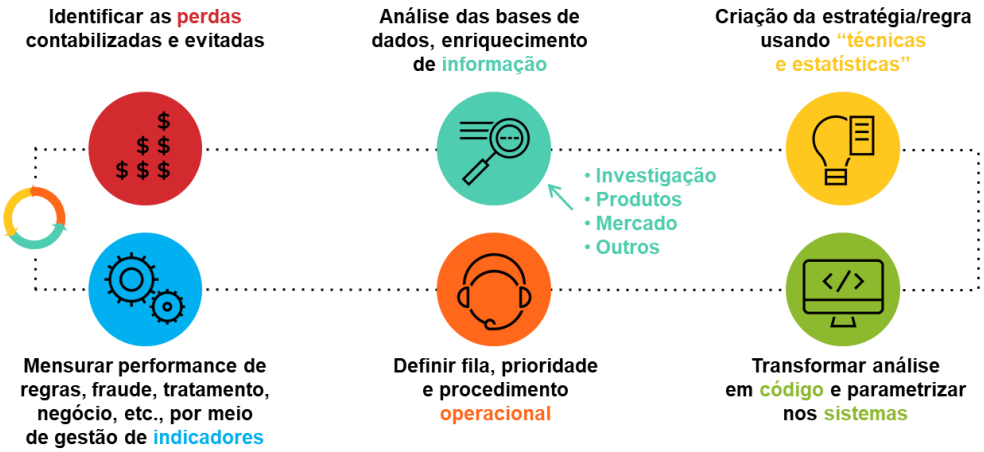
Existem diversas áreas envolvidas na gestão de fraudes de uma empresa de emissão de cartões. No caso da empresa de estágio do autor, as principais áreas são a área de negócios, atuando em um nível mais estratégico; a área estratégia de fraudes, atuando em um nível tático e; por fim, a área de operações de fraude, em um nível operacional. O que é buscado por cada uma dessas áreas se encontra na figura que segue:

Figura 28 - Times envolvidos com a gestão de fraudes

Fonte: elaboração do autor

Após a ocorrência de um caso de fraude, as equipes citadas anteriormente, principalmente as de estratégia e operações de fraude, se unem para atuar no chamado ciclo de fraude. As atividades desempenhadas no mesmo possuem um papel muito importante na gestão de fraude. O ciclo pode ser observado na figura abaixo:

Figura 29 - Ciclo de fraude



Fonte: elaboração do autor

2.3.3. Categorias de fraude

A fraude pode ser dividida em categorias e em modalidades. Como categorias, pode-se dividir as fraudes em Fraude Efetiva, Auto Fraude e Fraude Amiga, conforme a ilustração a seguir:

Figura 30 - Categorias de fraude

Fonte: elaboração do autor

2.3.4. Modalidades de fraude

Já ao se tratar da divisão de fraudes por modalidade, existem cinco tipos principais de modalidade:

- Falsificação do cartão: também conhecida, em inglês, por *skimming*, está relacionada à falsificação ou clonagem de um cartão. Ou seja, o fraudador monta um cartão com as mesmas informações do cartão original do emissor, obtidas por algum processo de cópia em terminais ATM ou máquinas *PoS* violadas pelo criminoso. Geralmente, as informações são obtidas por meio da trilha magnética do cartão verdadeiro;

- Perda ou roubo do cartão: o fraudador simplesmente toma posse do cartão do portador legítimo, seja por roubo ou encontrando o cartão perdido em algum lugar por descuido do portador, e o utiliza para a realização de transações;

- Extravio: ocorre quando os cartões com suas senhas são roubados durante o processo em que o envelope do emissor (contendo o cartão e a senha) é enviado para o endereço informado pelo portador legítimo do cartão. Os riscos de extravio estão relacionados às empresas que efetuam as entregas, como Correios e outras transportadoras;

- *Fraud application*: o fraudador faz uso indevido de informações do portador legítimo para criação de conta e emissão de cartão em um banco. Tais informações podem ser obtidas por meio de diferentes abordagens, como vírus de computador (cavalos de tróia, *keyloggers*, *backdoors*, além de outros *spywares*), por *phishing* e até por engenharia social. Pode acontecer, também, do fraudador ser um conhecido do proprietário legítimo das informações pessoais;

- Invasão de conta: o fraudador consegue invadir a conta bancária do usuário legítimo e, por conseguinte, alterar o endereço registrado no banco e pedir para que um novo cartão seja entregue ao novo endereço, de forma que o crimonoso o possa utilizar para futuras transações fraudulentas. Do mesmo jeito que no *fraud application*, o fraudador pode se utilizar de diversas abordagens para a invasão,

como as citadas anteriormente (vírus de computador, *phishing*, engenharia social etc);

- *E-commerce / Mail Order – Telephone Order (MO/TO)*: é uma das modalidades mais difíceis de serem detectadas. Nesse caso, por estarem relacionadas a transações de cartão não presente, o fraudador apenas precisa conseguir as informações obtidas no cartão (número, nome do portador, data de validade e o código de segurança), não precisando estar no estabelecimento comercial. Assim, é muito difícil realizar a autenticação do portador. Na maioria das vezes, é o comércio que arca com as perdas de fraude dessa modalidade.

2.3.5. Medição da fraude: Fraud Basis Points

Atualmente, no mercado de meios de pagamento, o índice de fraude é dado pela relação entre o volume financeiro de perda de fraude e o volume total transacionado (processado e autorizado) em um dado período. No entanto, é importante frisar que a maneira como se mostra o índice de fraude é padronizada para todos os *stakeholders* participantes desse mercado. A exibição do dado, em vez de ser dada por uma mera porcentagem, é feita em *Basis Points*, conceito que diz respeito a uma parte em dez mil. Assim, um Basis Point de fraude equivale a:

$$Fraud\ Basis\ Point\ (período\ t) = \ 10000\ x\ \frac{\$\ Fraude_t}{\$\ Transacionado_t}$$

2.3.6. Exemplos de fraude e dispositivos

De modo a viabilizar a clonagem de cartões ou intercepção e obtenção de dados, tanto do próprio cartão como também do portador deste, é imprescindível a utilização de aparatos específicos. Dentre esses dispositivos, existem os “chupa cabras” para clonagem de cartões, *keylogger* para obtenção de dados digitados em teclados, o *loop* libanês para roubo de cartões em caixas eletrônicos, dispositivo interceptador de dados na linha telefônica etc. Tais dispositivos são exemplificados nas figuras abaixo:

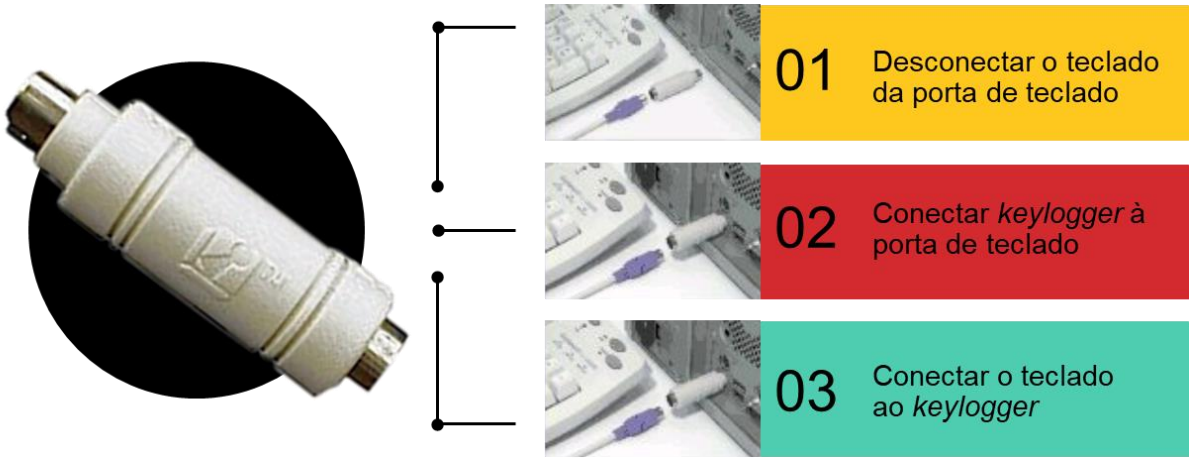
- Dispositivos para falsificação/clonagem de cartões:

Figura 31 - Dispositivos de clonagem de cartões

Fonte: material fornecido pela empresa onde estagiou

- Dispositivo *keylogger* físico para obtenção de informações a serem utilizadas para invasão de conta ou *fraud application*:

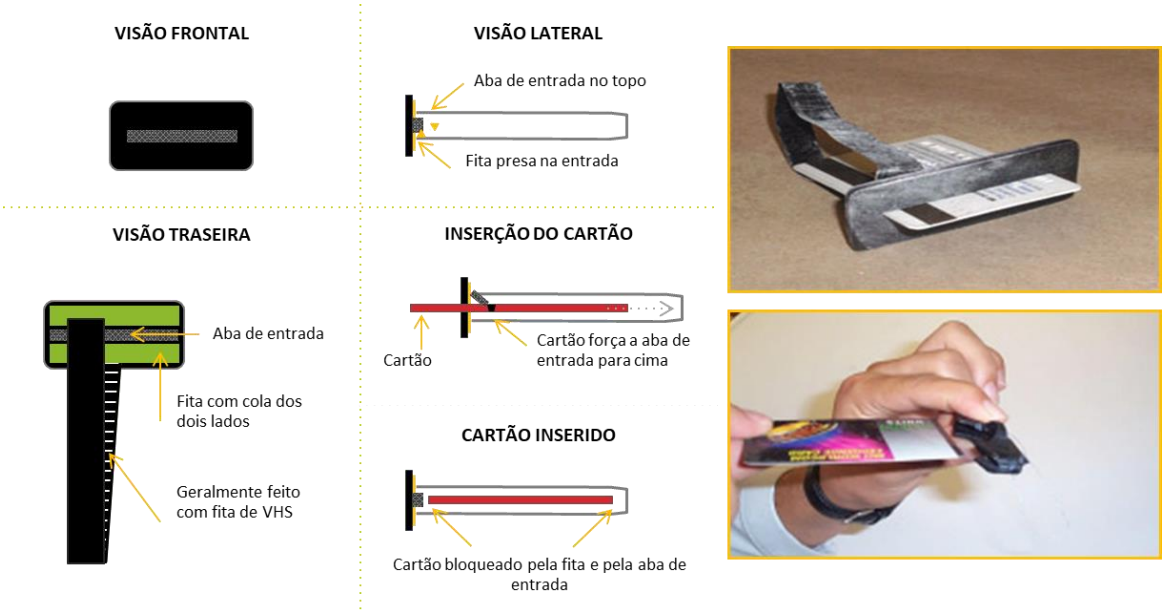
Figura 32 - Dispositivo *keylogger*



Fonte: material fornecido pela empresa onde estagiou

- Dispositivo para roubo de cartões – *Loop* Libanês:

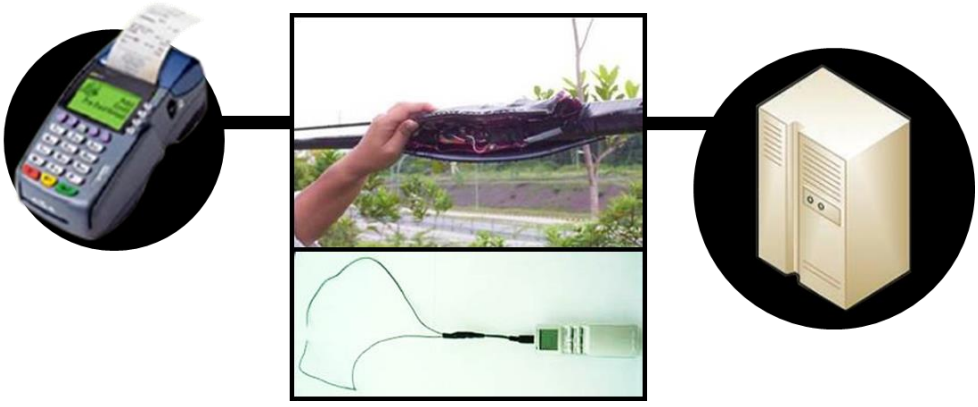
Figura 33 - Dispositivo para roubo de cartões



Fonte: material fornecido pela empresa onde estagiou

- Interceptação na linha telefônica entre o Ponto de Venda e o *host* (servidor) do emissor para obtenção de informações do cartão e/ou do portador a serem utilizadas em invasão de conta ou *fraud application*:

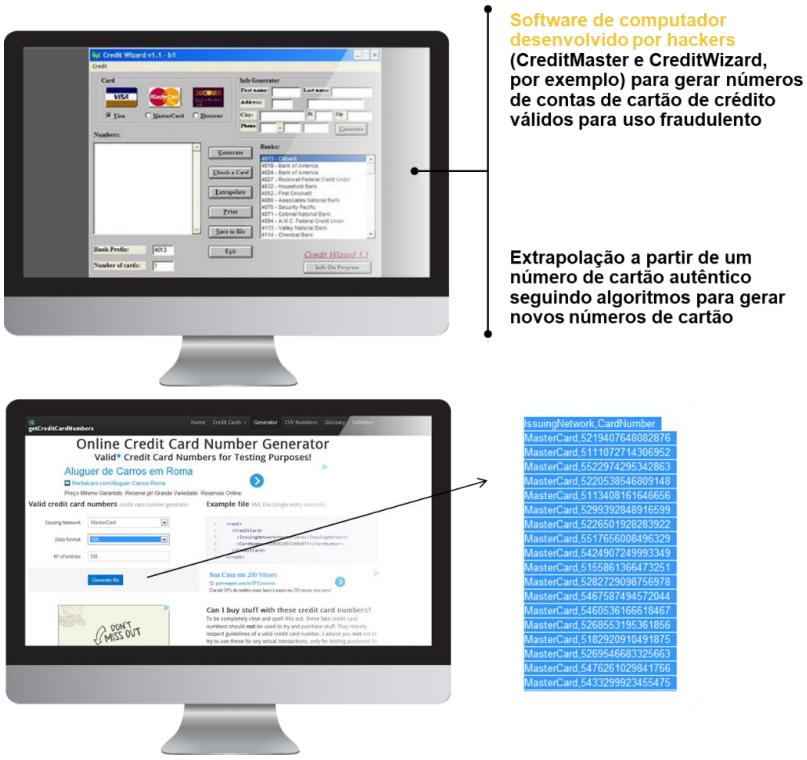
Figura 34 - Intercepção entre Ponto de Venda e servidor do emissor



Fonte: material fornecido pela empresa onde estagiou

- Fraude com *software* de geração de números de cartões e consequente utilização em transações fraudulentas de cartão não presente. Após os números de cartões serem gerados, eles são validados e, posteriormente, testados com diferentes códigos de verificação até funcionarem:

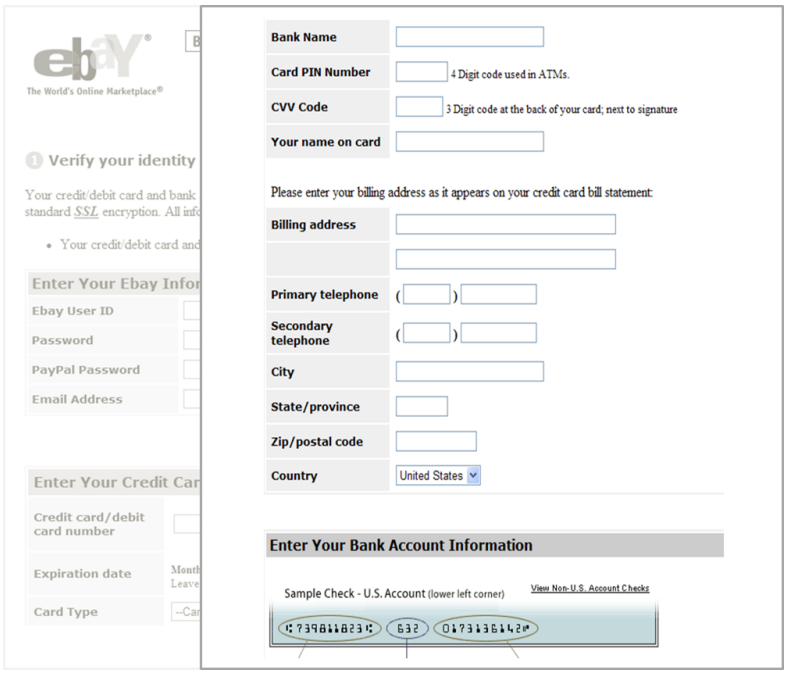
Figura 35 - *Software* para geração de números de cartões em massa



Fonte: elaboração do autor com dados fornecidos pela empresa onde estagiou

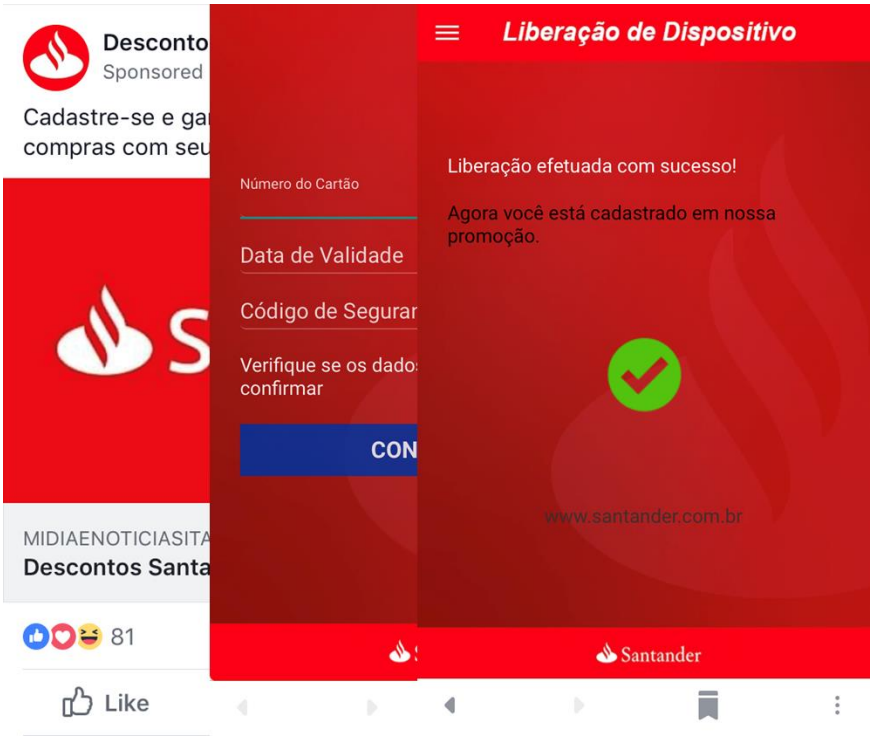
- Exemplos de páginas de *phishing* para obtenção de informações do portador e do seu cartão – casos do Ebay/Paypal e Santander:

Figura 36 - Exemplo de *phishing* no *e-commerce* Ebay



Fonte: elaboração do autor com dados de material da empresa onde estagia

Figura 37 - Exemplo de phishing do banco Santander na rede social Facebook

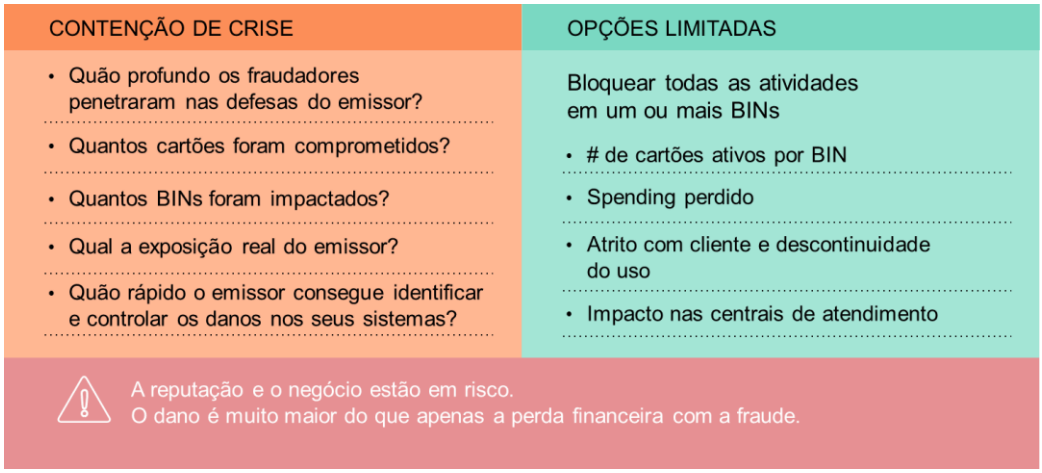


Fonte: elaboração do autor com base em material da rede social Facebook

2.3.7. Ferramentas de detecção de fraude

Ao se tratar de detecção de fraude, deve-se ter em mente que o ataque já ocorreu. Ao contrário de prevenção, a detecção visa a identificar os detalhes e o impacto do ataque de fraude depois dele ter ocorrido e conter tal ataque o mais rápido possível desde o momento de sua ocorrência.

Figura 38 - Ações após detecção de fraudes



Fonte: elaboração do autor com dados fornecidos pela empresa onde estagiou

Para tanto, existem ferramentas e abordagens de detecção de fraude que podem ser utilizadas pelo emissor de forma a identificar os mais diversos tipos de ataques de fraude.

a) Sistemas baseados em regras:

Segundo Tan et al. (2005), citado por Oliveira (2016, p. 24), os sistemas baseados em regras são algoritmos de aprendizagem supervisionada cujo classificador é constituído por um conjunto de regras, cada uma com o formato “se *(determinada condição é verdadeira)* então *(faça determinada ação)*”. O lado que contém a condição é chamado de *antecedente* da regra e o lado que contém a ação é chamado *consequente* da regra. É comum que haja mais de um atributo envolvido no antecedente. Uma regra *r* cobre um registro *x* se os atributos de *x* satisfizerem as condições expressas no antecedente de *r*. Nesse caso, também é dito que a regra foi disparada. Com a geração de várias regras, cria-se um conjunto de regras que pode ter duas propriedades:

- Conjunto de regras completo: os classificadores baseados em regras contêm essa propriedade quando cada um dos registros submetidos ao classificador dispara pelo menos uma regra;
- Regras mutuamente excludentes: um conjunto de regras é composto por regras mutuamente excludentes quando não houver nenhum registro que é coberto por mais de uma regra. Em outras palavras, cada instância dispara uma, e apenas uma, regra.

Uma dificuldade presente nos casos nos quais as regras não são mutuamente excludentes é que classes diferentes podem ser atribuidas a um mesmo registro. Há duas abordagens para contornar essa dificuldade.

- Ordenação das regras: ao utilizar essa estratégia, cria-se uma ordem dentro do conjunto de regras seguindo algum critério de ordenação. No momento em que um registro é exposto ao conjunto de regras, ele é testado contra as regras nessa ordem de prioridade, a primeira regra que ele disparar atribuirá a classe definida nela e o processo para. Portanto, não há como um registro ser

rotulado como pertencente a duas classes distintas. Quando o critério de ordenação considera alguma métrica individual de qualidade das regras, a ordenação é conhecida como ordenação baseada em regras. Uma vantagem desse esquema é que, pelo menos a priori, o registro será classificado pela melhor – segundo um critério de qualidade previamente escolhido – regra disponível no conjunto. De outra forma, a ordenação pode ser feita baseada em classes, ou seja, primeiramente, o registro é exposto a todas as regras que, se disparadas, atribuem, por exemplo, a classe “transação legítima”, para, posteriormente, o registro ser exposto às regras que o classificam como “transação fraudulenta”;

- Não ordenação das regras (busca de consenso): com essa estratégia, aceita-se que diversas regras classifiquem uma instância, sendo que o disparo de uma regra funciona como um voto que, inclusive, pode ser ponderado. A instância recebe a classe que tiver mais votos. Uma vantagem é que essa estratégia pode diminuir os erros de classificação por considerar o parecer de várias regras.

Conforme Beraldi (2014, p. 30), os pontos positivos e negativos da utilização de modelos baseados em regras constituem:

- Pontos positivos:
  i. Atualização dinâmica (usualmente, uma nova regra entra em poucos segundos em ambiente de produção);
  ii. Facilidade de desenvolvimento e implantação;
  iii. Controle;
  iv. Baixo custo e rapidez.

- Pontos negativos:
  i. Requer atualização frequente;
  ii. Grande volume de regras;
  iii. Necessidade de *experts* para o desenvolvimento de regras;
  iv. Reflete um padrão limitado;
  v. Difícil entendimento da relação entre regras e duplicidade de regras.

b) Modelos de pontuação:

Ainda conforme as palavras de Beraldi (2014, p. 31), os modelos de pontuação, os quais são também conhecidos como modelo de *scoring*, utilizam-se de técnicas estatísticas para o retorno de uma pontuação (*score*) para uma determinada transação. Geralmente, quanto maior a pontuação, maior a probabilidade (suspeita) de uma transação ser fraudulenta.

A pontuação pode ser computada para cada transação no banco de dados e utilizada no processo de prevenção para aprovar, negar ou referir uma transação, bem como ser utilizada em sistemas baseados em regras para combinação com outras variáveis (o sistema, assim, passa a ser conhecido como um híbrido de regras e *scoring*). Com o valor da pontuação, os casos com maior pontuação podem ser priorizados no processo de investigação da transação.

Nesse momento, questões de custo são consideradas, dado que é muito caro realizar uma investigação detalhada de todos os casos. Uma investigação deve se concentrar sobre casos com maior suspeita de fraude.

Em geral, modelos de regressão logísitca binária e redes neurais também são utilizados para gerar a pontuação do modelo. Mede-se, em uma escala de 0 a 100 ou de 0 a 1000, a probabilidade de uma transação ser fraudulenta com base em características como hora da transação, ramo de atividade do estabelecimento, valor, entre outras variáveis. Pontos de corte são adotados para adequar a capacidade de tratamento do volume de casos/alertas gerados em filas de trabalho. O autor também destaca pontos positivos e negativos associados aos modelos de pontuação:

- Pontos positivos:
  - i. Utilizado por toda a indústria financeira, dada a sua efetividade na decisão;
  - ii. Ideal para grandes volumes de transações nos quais a decisão precisa ser tomada rapidamente;
  - iii. Abrange perfis de comportamentos individuais.
- Pontos negativos:
  - i. Não acompanha tendências de fraudes recentes;
  - ii. Dependendo da modelagem, pode não refletir características da fraude local;
  - iii. Alto custo e sem controle para mudanças rápidas;

iv. Pode gerar resultados que não se pode explicar.

3. TRANSAÇÕES CNP E O CENÁRIO *E-COMMERCE*

Neste capítulo, são exploradas as transações dentro do canal de cartão não presente, ou seja, transações que ocorrem quando o portador não se encontra fisicamente no estabelecimento, geralmente em uma loja *online*.

3.1. A fraude nas transações de cartão não presente

A fraude está majoritariamente concentrada em transações do canal de cartão não presente. Para esta fraude ser analisada, primeiramente, focou-se no subcanal *e-commerce*, que abriga grande parte dessa fraude, assim como foi comparada tal fraude com a do canal de cartão presente.

3.1.1. Situação atual do *e-commerce* no Brasil

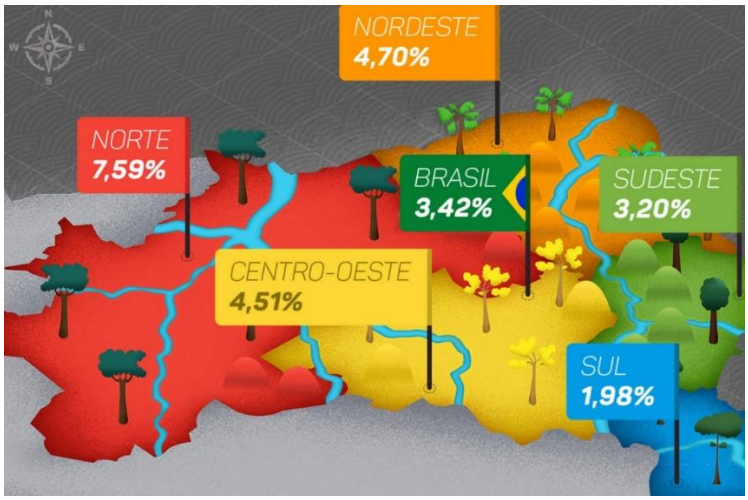
Segundo uma análise da Totvs (2018) da 37ª edição da pesquisa Webshoppers, de 2018, produzida pela Ebit em parceria com a bandeira Elo, o volume de vendas de bens de consumo no *e-commerce* tradicional, em termos financeiros, correspondeu a R$ 47,7 bilhões. O crescimento foi de 7,5% em relação ao ano anterior, 2016, quando o faturamento atingiu R$ 44,4 bilhões. Ademais, ainda segundo a organização, ao se observar o *e-commerce* geral, o qual inclui itens novos, mercadorias usadas e artesanado, a elevação foi de 21,9% de 2017 em relação a 2016, tendo chegado ao faturamento de R$ 73,4 bilhões. Nesse mesmo ano de 2017, os principais mercados de *e-commerce*, organizados por ordem decrescente de faturamento, são:

a) Telefonia e celulares – representando 21,2% do faturamento do comércio eletrônico brasileiro em 2017;
b) Eletrodomésticos – representando 19,3%;
c) Eletrônicos – representando 10%;
d) Informática – representando 8,9%;
e) Casa e decoração – representando 8,4%;
f) Outros – por fim, representando 32,3%.

Ao entrar no cenário de fraude no subcanal CNP *E-Commerce*, de acordo com a pesquisa “Mapa de Fraude” realizada pela Clearsale (2018), a qual analisou

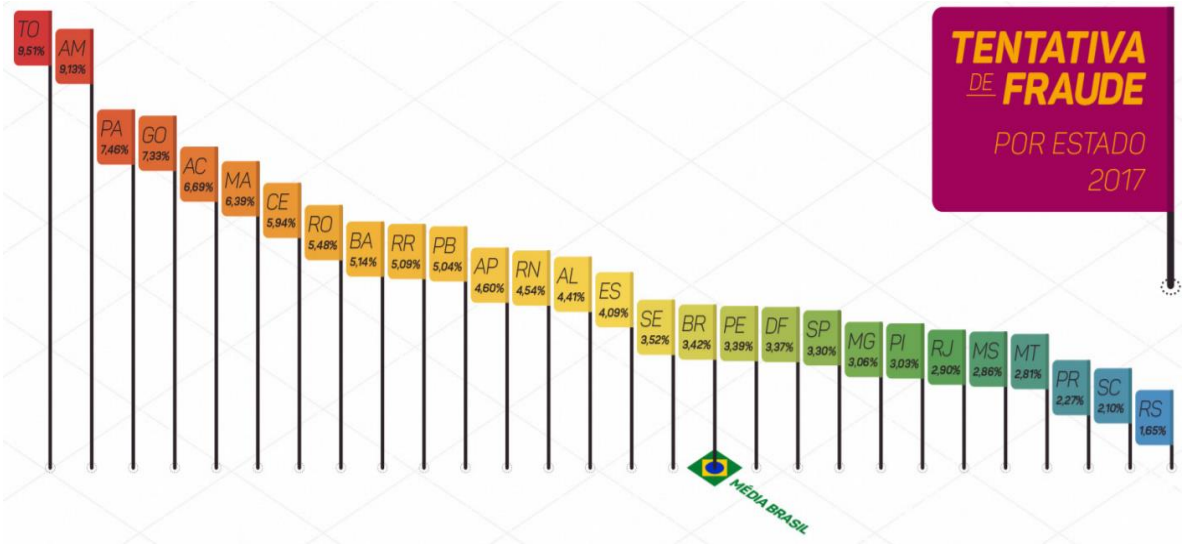
informações de 85% do mercado de e-commerce e mobile com extrapolação de dados – utilizando uma metodologia que considera apenas lojas virtuais que vendem bens físicos – a tentativa de fraude representou 3.42% de todo o volume financeiro transacionado em *e-commerce* no território brasileiro em 2017. A seguir, tem-se as porcentagens por região e por estado, além da própria “média Brasil”:

Figura 39 - Porcentagem de fraude em *e-commerce* por região no Brasil em 2017



Fonte: Clearsale (2018)

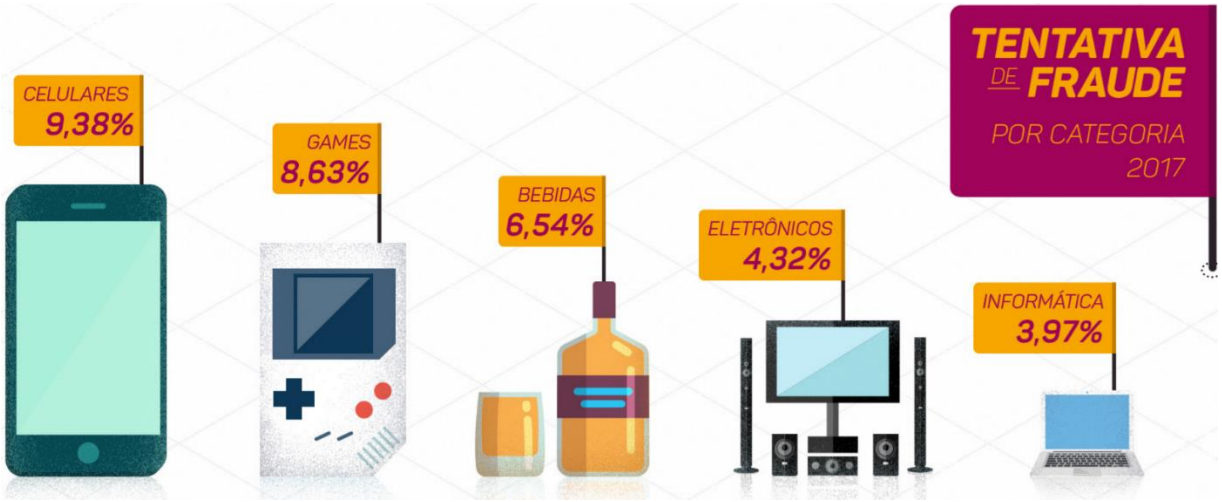
Figura 40 - Tentativa de fraude por estado (%) em *e-commerce* no Brasil em 2017



Fonte: Clearsale (2018)

No tocante à tentativa de fraude por categoria, tem-se, como alvo primário, a compra de celulares em comércios eletrônicos. Em segundo lugar, jogos eletrônicos e, por fim, em terceiro, o comércio eletrônico de bebidas, conforme figura a seguir:

Figura 41 - Tentativa de fraude por categoria de *e-commerce* (%) no Brasil em 2017



Fonte: Clearsale (2018)

O subcanal CNP *E-Commerce*, em uma divisão por corredor de transação (doméstico e *crossborder*), teve as seguintes porcentagens em 2017, tanto levando em conta a quantidade de transações fraudulentas como também o volume financeiro fraudulento que foi transacionado:

Tabela 1 - Porcentagem da quantidade e do valor financeiro de fraude em *e-commerce* por corredor transacional no Brasil em 2017

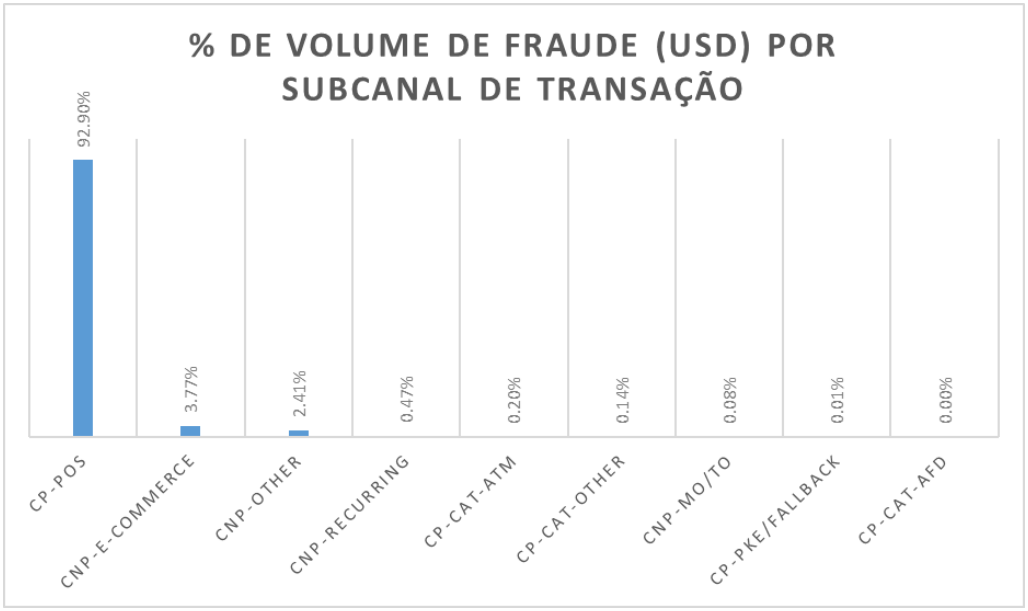
Porcentagem de quantidade de transações fraudulentas e valor financeiro fraudulento por mês/2017	Crossborder	Doméstico
% Quantidade de transações fraudulentas	14%	86%
% Valor Financeiro de transações fraudulentas	27%	73%

Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Pode-se perceber então que, em 2017, para cartões emitidos no Brasil, 86% da quantidade de transações fraudulentas em CNP *E-Commerce* foi doméstica e 14% foi *crossborder*, e que 73% do volume financeiro das transações fraudulentas foi *crossborder* e 27% foi doméstico.

Ademais, no ano de 2017, o subcanal transacional CNP *E-Commerce* ficou em segundo lugar em termos de volume de fraude em transações com cartão, com 3.77% do volume de fraude (em dólares) total nesse mesmo ano, apenas abaixo da fraude no subcanal transacional CP *PoS*, o qual representa 92.90% da fraude no Brasil nesse ano.

Figura 42 - Porcentagem de volume financeiro de fraude (USD) por subcanal transacional no Brasil em 2017



Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Contudo, é um equívoco analisar apenas as porcentagens do volume de fraude de um país em um determinado ano, para cada classificação. Aliado ao volume de fraude, deve-se ter em análise, também, o volume financeiro total das transações por cartões, no país, no dado período.

Assim, analisando-se tal volume, tem-se que o subcanal CP *PoS* representa 83.28% do volume total transacionado por cartões, no Brasil, em 2017. Sendo assim, é por isso que ele apresenta uma porcentagem maior de fraude.

Tabela 2 - Porcentagem do volume financeiro (USD) transacionado por cartões no Brasil em 2017 por subcanal transacional

Subcanal	% do Volume Total transacionado por cartões no Brasil em 2017
CP-PoS	83.28%
CNP-E-Commerce	11.18%
CNP-Other	2.58%
CNP-Recurring	1.27%
CP-CAT-ATM	0.74%
CP-CAT-Other	0.60%
CNP-MO/TO	0.29%
CP-PKE/Fallback	0.05%
CP-CAT-AFD	0.00%
Total	**100.00%**

Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Finalmente, ao se comparar os *fraud basis points (fraud BPs)* dos dois primeiros subcanais, os quais são indicadores representados por 1 parte em 10000 da relação entre volume de fraude e volume total transacionado por cartão, em 2017 no Brasil, tem-se o seguinte:

Tabela 3 - BPs de fraude dos subcanais CNP-*E-Commerce* e CP-*PoS* no Brasil em 2017

Subcanal	Fraud Basis Points
CNP-E-Commerce	61.42
CP-PoS	1.83

Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Assim, a relação financeira fraude/total é muito maior em CNP E-Commerce quando comparado a CP-PoS.

3.1.2. Transações CNP frente às CP

A maioria das perdas e não-ganhos, atualmente, no mercado de meios de pagamento, estão associados às transações de cartão não presente, em consonância com o conteúdo exposto no capítulo 1.3 da relevância do trabalho.

Considerando o período de 2017, no Brasil, a relação de transações CNP frente às transações CP, em termos de volume, foi a seguinte:

Tabela 4 - Fração (%) do volume financeiro transacionado em relação ao transacionado no Brasil em 2017 por canal transacional

Canal	Volume Transacionado no canal em relação ao Volume Transacionado Total em 2017
CNP	15.33%
CP	84.67%

Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

É possível, no entanto, que transações sejam declinadas (não aprovadas) por motivos diversos, como por suspeita de fraude ou fundos insuficientes. A taxa de aprovação de cada um dos canais é mostrada abaixo:

Tabela 5 - Taxa de aprovação (%) no Brasil em 2017 por canal transacional

Canal	Taxa de Aprovação (%) em 2017
CNP	61.30%
CP	95.39%

Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Assim, CNP representa 15,33% do volume total transacionado por cartões emitidos no Brasil em 2017, sendo que 61,30% dessas transações foram aprovadas. Assim, o volume total aprovado de transações CNP representa a seguinte porcentagem do volume total transacionado no Brasil em 2017:

$$CNP\colon\ 61{,}30\% \times 15{,}33\% \cong \mathbf{9{,}40}\% \left[\frac{\boldsymbol{vol.aprovado\ CNP}}{\boldsymbol{vol.transacionado\ TOTAL}}\right] \quad (1)$$

Seguindo a mesma lógica para CP, tem-se:

$$CP\colon\ 95{,}39\% \times 84{,}67\% \cong \mathbf{80{,}77}\% \left[\frac{\boldsymbol{vol.aprovado\ CP}}{\boldsymbol{vol.transacionado\ TOTAL}}\right] \quad (2)$$

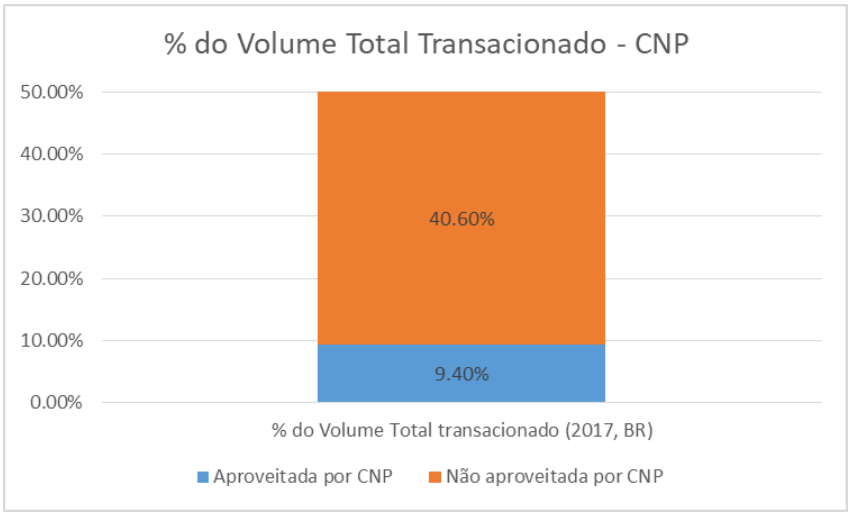
Ou seja, as transações CNP já são minoria. Aliado ao fato de terem uma taxa de aprovação ainda menor que as transações CP, representam apenas 9,40% do volume total transacionado no Brasil (tomando o ano de 2017 como período de referência).

Em outras palavras, considerando um **<u>cenário ideal</u> como sendo composto de um volume de 50% de transações CNP**, as quais apresentam **taxa de aprovação de 100%**, deixou-se de ganhar:

$$\%\ não\ aproveitada\ CNP\colon (100\% \times 50\%) - 9{,}40\% = \mathbf{40{,}60}\% \quad (3)$$

Assim, tem-se o seguinte gráfico:

Figura 43 - Procentagem do volume financeiro transacionado no Brasil em 2017 aproveitada e não aproveitada pelo canal CNP



Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Ademais, comparando os canais CP e CNP por corredor de transação (doméstico e *crossborder*), tem-se a segmentação abaixo:

Tabela 6 - Representatividade da quantidade de transações (%) por canal e por corredor transacional no Brasil em 2017

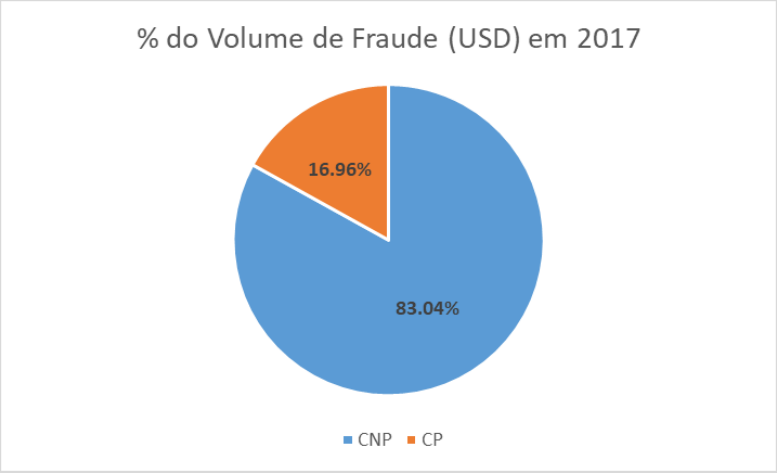
Canal / Corredor	Doméstico	Crossborder	Doméstico+Crossborder
CNP	87.57%	12.43%	15.33%
CP	98.43%	1.57%	84.67%
CP+CNP	96.76%	3.24%	100.00%

Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Ou seja, a maioria das transações está no corredor doméstico. No entanto, a proporção das transações CNP que está no corredor *crossborder* é maior do que a proporção das transações CP que está presente neste corredor.

Em termos de volume financeiro de fraude, a seguinte distribuição foi vigente no ano de 2017 no Brasil:

Figura 44 - Porcentagem do volume financeiro de fraude no Brasil em 2017 por canal transacional



Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Isto é, de 100% do volume financeiro de fraude, 83% se concentra no canal CNP e 17% no CP.

Contudo, analisando agora a **porcentagem que representa e relação entre o volume de fraude de um canal e do volume aprovado desse canal**, para cada um dos canais (CP e CNP), tem-se a seguinte tabela:

Tabela 7 - Volume financeiro de fraude em relação ao aprovado no Brasil em 2017 por canal transacional

Canal	Volume de Fraude do canal (USD) em relação ao Volume Aprovado do canal (USD) (2017, BR)
CNP	0.62%
CP	0.02%

Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Desse modo, em termos do volume financeiro total das transações em 2017, no Brasil, a porcentagem que representa o volume das transações CNP **aprovadas** e que **não foram fraude** é calculada a seguir:

$$\begin{aligned} &\%\, Transações\ CNP\ Aprovadas\ Não\ Fraudulentas \\ &= 61{,}30\%\ (taxa\ de\ aprovação\ das\ transações\ CNP) \\ &\times 15{,}33\%\ (transações\ CNP) \\ &\times (1 - 0{,}62\%) \cong \mathbf{9{,}34}\% \end{aligned} \qquad (4)$$

Resumidamente:

Tabela 8 - Fração de transações não fraudulentas aprovadas em relação ao volume financeiro total aprovado no Brasil em 2017 por canal transacional

Canal	A = % do Volume Total	B = Taxa de Aprovação (%)	C = (A x B) = % Aprovada do Volume Total	D = Volume Fraude / Volume Aprovado do canal (%)	E = (C x D) = % Fraude Aprovada do Volume Total	F = (C – E) = % Não Fraude Aprovada do Volume Total
CNP	15.33%	61.30%	9.40%	0.62%	0.06%	**9.34%**
CP	84.67%	95.39%	80.77%	0.02%	0.02%	**80.75%**

Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Já relacionando o **volume fraudulento de um canal** pelo **volume aprovado desse canal**, para os dois canais, e exibindo tal relação em *Basis Points* (uma parte em dez mil, em vez de porcentagem), observa-se o seguinte:

Tabela 9 - BPs de fraude no Brasil em 2017 por canal transacional

Canal	Fraud Basis Points (BPs) (2017, BR)
CNP	61.93
CP	2.29

Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Percebe-se, então, que o canal CNP apresenta uma relação fraude/aprovada maior que a do canal CP.

Ainda no cenário de fraude, comparando os canais por corredor de transação (doméstico e *crossborder*), tem-se a seguinte segmentação:

Tabela 10 - Representatividade do valor financeiro de fraude (%) por canal e por corredor transacional no Brasil em 2017

Canal / Corredor	Doméstico	Crossborder	Doméstico+Crossborder
CNP	72.95%	27.05%	83.04%
CP	77.89%	22.11%	16.96%
CP+CNP	73.79%	26.21%	100.00%

Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Ou seja, a maioria da fraude está concentrada no corredor doméstico. No entanto, a proporção de fraude CNP que está no corredor *crossborder* é maior do que a proporção de fraude CP que está presente neste corredor.

Finalmente, é possível perceber que grande parte do valor não-ganho (devido às taxas de aprovação não tão altas) e das perdas com fraude estão concentradas no canal de CNP. A maior porcentagem dessa fraude está no subcanal CNP *E-Commerce*, conforme o capítulo 3.1.1 do cenário de e-commerce no Brasil. Assim, faz-se necessária uma estratégia de detecção de fraude capaz de classificar, com maior assertividade, as transações fraudulentas e não fraudulentas em um curto intervalo de tempo após a ocorrência de um ataque de fraude.

Ou seja, uma abordagem assertiva de detecção à fraude ajuda na redução do número de falsos positivos (transações taxadas como fraude sendo que não são fraude), portanto, corrobora com a diminuição de transações declinadas por suspeita de fraude – aumentando, assim, a taxa de aprovação do canal e reduzindo os não-ganhos.

Por outro lado, também ajuda no aumento do número de verdadeiros positivos e verdadeiros negativos e na redução do número de falsos negativos (transações taxadas como não-fraude sendo que são fraude), portanto, colabora com a diminuição das perdas com fraudes que não foram detectadas ou foram detectadas erroneamente.

Os conceitos de verdadeiros e falsos positivos e negativos, assim como outros indicadores de performance de detecção de fraude, são explicados com maiores detalhes na revisão bibliográfica.

A proposta deste trabalho foi comparar duas abordagens de *machine learning* e determinar qual é a melhor em termos de assertividade. A abordagem vencedora foi utilizada como um modelo de detecção à fraude, trabalhando com variáveis de transações para definir quais transações são fraudulentas e quais não são. O ideal é ter o menor número de falsos positivos e falsos negativos de forma a diminuir as perdas com fraude e os não-ganhos por transações erroneamente declinadas por suspeita de fraude.

Os conceitos sobre metodologias de *machine learning* são, a priori, apresentadas no capítulo seguinte (revisão bibliográfica), bem como os indicadores de performance de fraude e de performance das próprias técnicas de *machine learning*. Posteriormente, a comparação de duas metodologias de *machine learning* (regressão logística e *random forest*) foi realizada no capítulo de metodologia. Foram apresentados os passos de cada uma das técnicas e o resultado final da comparação.

4. REVISÃO BIBLIOGRÁFICA

Neste capítulo, são apresentadas as bases teóricas do trabalho. Assim, este capítulo foca no significado de aprendizado de máquina e dos seus conceitos, assim como caracterização das metodologias dessa área, conceitos e indicadores que compõem os métodos utilizados para detecção de fraudes.

4.1. Machine learning: métodos estatísticos e aprendizagem

Em consonância com o exposto por Resende (2003), o aprendizado de máquina é uma subárea da Inteligência Artificial que tem como objetivo estudar e produzir técnicas ou sistemas computacionais capazes de adquirir conhecimento de forma automática. Um sistema de aprendizado é um programa de computador que toma decisões baseado em experiências acumuladas através da solução bem-sucedida de problemas anteriores. Os diversos sistemas de aprendizado de máquina possuem características particulares e comuns que possibilitam sua classificação quanto à linguagem de descrição, modo, paradigma e forma de aprendizado utilizado.

De acordo com Dietterich (1997), o aprendizado de Máquina é uma ferramenta poderosa, mas não existe um único algoritmo que apresente um bom desempenho para todos os problemas. Sendo assim, faz-se fundamental a compreensão do poder e da limitação dos diferentes algoritmos, usando alguma metodologia de avaliação que permita realizar a comparação de tais algoritmos.

Atualmente, existem vários grupos de métodos de *machine learning* disponíveis para serem utilizados. A escolha do melhor grupo de métodos é dependente de uma gama de variáveis como, por exemplo, do tipo de problema a ser endereçado, o desempenho do método, a disponibilidade de dados e sua natureza etc.

Conforme Bolton e Hand (2002), os métodos estatísticos frequentemente utilizados para detecção de fraude podem ser classificados em não supervisionados ou supervisionados. Tais agrupamentos de métodos são apresentados a seguir.

4.1.1. Conceitos básicos de *machine learning*

A seguir, são expostos alguns conceitos básicos relativos a *machine learning*, de acordo com as palavras de Oshiro (2013):

a) Classificador:

Dado um conjunto de exemplos de treinamento, um indutor (ou algoritmo de aprendizado) gera como saída um classificador (ou hipótese, ou descrição de conceito) de forma que, dado um novo exemplo, ele possa predizer precisamente sua classe. No aprendizado supervisionado, todo exemplo (x_i, y_i) possui um atributo especial y_i, o rótulo ou classe, que descreve o fenômeno de interesse, isto é, a meta que se deseja aprender e poder fazer previsões a respeito. Um exemplo não rotulado x_i consiste do exemplo, exceto o rótulo, ou seja, um vetor de valores dos atributos.

b) Atributos e conjuntos de exemplos:

Um conjunto de dados ou de exemplos (ou, em inglês, *dataset*) é composto por exemplos contendo valores de atributos bem como a classe associada:

Figura 45 - Esquematização das variáveis de um conjunto de dados

	X_1	X_2	$\cdots$	X_a	Y
z_1	x_{11}	x_{12}	$\cdots$	x_{1a}	y_1
z_2	x_{21}	x_{22}	$\cdots$	x_{2a}	y_2
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
z_n	x_{n1}	x_{n2}	$\cdots$	x_{na}	y_n

Fonte: Oshiro (2013)

Mostra-se, na tabela acima, o formato padrão de um conjunto de exemplos “T” com “n” exemplos e “a” atributos. Nessa tabela, a linha “i” refere-se ao i-ésimo exemplo (i = 1,2,...,n) e a entrada x_{ij} se refere ao valor do j-ésimo (j = 1,2,...,a) atributo X_j do exemplo “i”.

4.1.2. *Ondersampling*, *undersampling* e *overfitting*

De acordo com Kaur e Gosain (2018), *oversampling* (ou sobreamostragem) é o processo de aumentar o número de amostras da classe minoritária, aleatoriamente, através da replicação dos mesmos dados ou gerando sinteticamente usando alguma técnica para melhorar a taxa de desequilíbrio, de modo que os os algoritmos de classificação possam ser usados para classificar os dados. A vantagem dessa técnica é que não há perda de nenhuma informação importante do conjunto de dados e o conjunto de dados original é retido, embora novas informações sejam adicionadas a ele de forma a equilibrar os dados. A limitação dessa técnica é que leva mais tempo para ser executada em comparação com a abordagem de *undersampling*, pois estamos aumentando o número de instâncias. Também pode causar o problema de *overfitting* no caso de replicar as mesmas amostras.

Overfitting significa uma situação em que o modelo se “lembra” muito dos padrões do *dataset* de treinamento e, assim, não consegue fazer previsões certeiras com os dados que recebe futuramente no momento do teste.

Já *undersampling*, ao contrário da técnica de *oversampling*, algumas amostras da classe majoritária são removidas aleatoriamente ou usando alguma técnica para balancear as classes.

4.1.3. Métodos Não Supervisionados

Conforme exposto por Russel e Norvig (2002), no aprendizado não supervisionado, os algoritmos assumem que não se conhece a classe à qual os exemplos pertencem e procuram encontrar nos valores de atributos similaridades ou diferenças que possam, respectivamente, agrupar os exemplos pertencentes à mesma classe ou dispersar os exemplos de classes distintas.

Em outras palavras, nos algoritmos agrupados como não supervisionados (ou descritivos), não há um alvo ou variável de *output* para prever ou estimar. Tais algoritmos são usados para agrupar (*clusterizar*) uma população em diferentes grupos. Um exemplo de aplicação desses métodos seria a segmentação de clientes de uma marca ou determinado produto em diferentes grupos e, posteriormente, aplicar uma intervenção ou estratégia de venda ou *marketing* específica a cada um dos grupos.

Alguns exemplos de métodos não supervisionados são o *clustering* e a decomposição em valores singulares.

Ao se tratar de fraude, de acordo com Beraldi (2014, p. 29), os métodos não supervisionados buscam informações de contas, clientes, números de cartões, entre outras, que possuam um comportamento diferente do normal e que, muitas vezes, são designados como anomalias ou *outliers* e caracterizados como uam forma básica não padrão de observação.

4.1.4. Métodos Supervisionados

Em consonância com o exposto por Carvalho (2014), a abordagem de aprendizagem supervisionada, basicamente, consiste na utilização de uma gama de exemplos, denominados instâncias, já classificados, de modo a induzir um modelo que seja capaz de classificar novas instâncias de forma precisa, com base no aprendizado obtido mediante treinamento com os dados de treinamento (*dataset*). É comum que este modelo seja chamado também de classificador. Faz-se importante existir um conjunto de dados de treinamento de qualidade, para que o modelo criado possa ser capaz de predizer novas instâncias de forma eficiente.

De acordo com McCue (2018), a diferença entre os métodos de aprendizado supervisionados e não-supervisionados reside no fato de que os métodos não-supervisionados não precisam de uma pré-categorização para os registros, ou seja, não é necessário um atributo alvo. Tais métodos geralmente usam alguma medida de similaridade entre os atributos. As tarefas de agrupamento e associação são consideradas como não-supervisionadas. Já no aprendizado supervisionado, os métodos são providos com um conjunto de dados que possuem uma variável alvo pré-definida e os registros são categorizados em relação a ela. As tarefas mais comuns de aprendizado supervisionado são a classificação (que também pode ser não-supervisionado) e a regressão.

Em termos gerais, neste tipo de aprendizado, o programa é treinado ao utilizar um conjunto de dados pré-definidos. Durante tal treinamento, aprende-se o que é necessário de modo a tomar futuras decisões ao receber novos dados. Após efetuado o treinamento, o algoritmo deve ser capaz de prever resultados de novos exemplos e, por conseguinte, melhorar sua eficácia cada vez mais.

Beraldi (2014, p. 30) deixa claro que, nos métodos supervisionados, as amostras dos casos fraudulentos e legítimos são usadas para construir modelos que permitem atribuir novas observações a uma das duas classes. Certamente, isso exige discriminar um conjunto de variáveis que possam classificar, de maneira

correta, entre essas classes, os dados originais utilizados para construir os modelos. Além disso, só pode ser usado para detectar fraudes que tenham ocorrido anteirormente em um determinado espaço de tempo.

Geralmente, de forma a detectar fraudes, os métodos supervisionados se dividem em dois grupos: sistemas baseados em regras e modelos de pontuação. Assim, conforme Beraldi (2014, p. 30):

a) Sistemas baseados em regras:

Produzem classificadores utilizando regras com a seguinte premissa:

Se **(CONDIÇÃO = VERDADEIRA)** então **(AÇÃO)**

- Pontos positivos:
  - Configuração fácil e rápida (desenvolvimento e implantação) e controle;
  - Atualização dinâmica;
  - Baixo custo.
- Pontos negativos:
  - Requer atualizações frequentes;
  - Grande volume de regras;
  - Necessidade de empregados experientes para o desenvolvimento de regras;
  - Reflete um padrão limitado;
  - Difícil entendimento da relação entre as regras e potencial de duplicidade.

b) Modelos de pontuação (*scoring*):

Baseiam-se em técnicas estatísticas para retornar uma pontuação para determinada transação. Geralmente, quanto maior tal pontuação, maior a probabilidade (ou suspeita) de tal transação ser fraudulenta. Ademais, tendo o valor da pontuação, os casos com as maiores pontuações podem ser priorizados no processo de investigação, já que é muito caro realizar uma investigação detalhada de todos os casos. A pontuação geralmente é medida em uma escala de 0 a 100 ou 0 a 1000. A pontuação é refletida por características como hora da transação, ramo de atividade do estabelecimento, valor etc.

- Pontos positivos:
  - Utilizado por toda a indústria financeira;
  - Alta efetividade na decisão;
  - Ideal para grandes volumes de transações para as quais uma decisão necessita ser tomada imediatamente;
  - Abrange perfis de comportamentos individuais.

- Pontos negativos:
  - Não acompanha tendências de fraudes recentes;
  - Dependendo da modelagem, pode não refletir características da fraude local;
  - Alto custo;
  - Sem controle para mudanças repentinas;
  - Pode gerar resultados inexplicáveis.

Para atuação em fraude, é ideal que se utilizem métodos supervisionados de classificação. Conforme experimento feito por Niu, Wang e Yang (2019), os métodos supervisionados tem melhor desempenho do que os não-supervisionados em bases de dados com transações de cartão de crédito.

Desses métodos classificadores, resumidamente, os mais conhecidos são os seguintes:

a) Regressão logística: similar à regressão linear, contudo, utilizada quando a variável dependente não é um número, mas sim uma resposta binária (como “fraude” / “não fraude”). A função logística (ou *logit*) é representada por uma curva sigmoidal e, para qualquer variável de entrada X, a função logística garante que a saída seja um valor entre 0 e 1 – o qual dita a probabilidade de ocorrência de um determinado evento (classe). No caso deste trabalho, o evento é uma transação ser fraude. A técnica é chamada de regressão apenas por ter um funcionamento similar à regressão linear, no entanto, ela é um método de classificação;

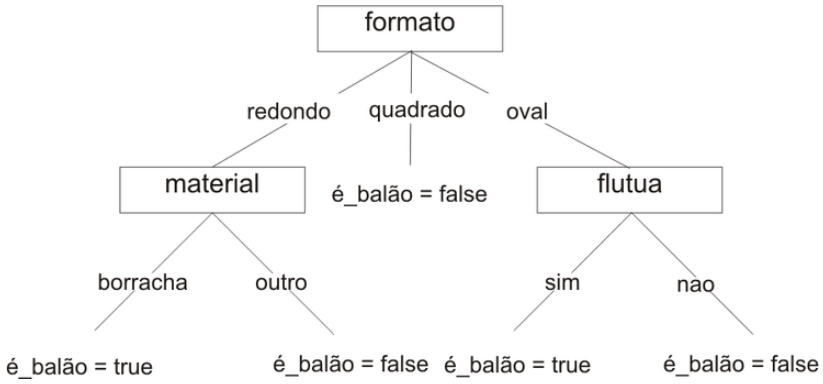
b) K-NN (K-*Nearest Neighbors*): é um algoritmo de classificação simples utilizado para identificar pontos de dados. O algoritmo separa tais pontos em classes

para predizer a classificação de um novo ponto que porventura alimente o sistema no futuro. É um método de aprendizado lento e realiza a classificação de novos casos baseado em mensuração de similaridade (como, por exemplo, utilizando funções de distância entre pontos);

c) Naive Bayes: trabalha com probabilidades, baseado-se no teorema bayesiano. Assume que a presença de uma *feature* (variável/caracterísitca) em uma classe é independente, ou seja, não está relacionada com nenhuma outra variável da mesma classe;

d) Árvore de decisão: trabalha com um modelo de classificação baseado em uma estrutura de árvore. Basicamente, ramifica o *dataset* em partes menores, até atingir nós de decisão, que permitem realizar a classificação desejada. Um exemplo de árvore decisão é exibido abaixo:

Figura 46 - Exemplo de árvore de decisão

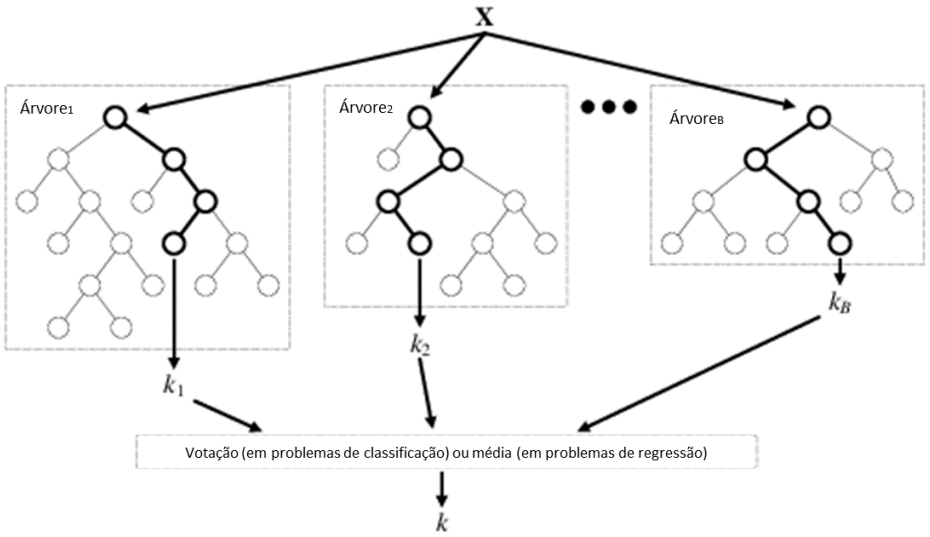


Fonte: Araujo (2004)

e) Modelos combinados (*ensembled methods*): é uma combinação de várias abordagens como as supracitadas. Um dos algoritmos mais conhecidos é o da **floresta aleatória** (*random forest*), sendo este uma combinação (*ensemble*) de árvores de decisão. Em termos gerais, tal algoritmo cria várias árvores de decisão e as combina de forma a obter uma predição de maior acurácia e mais estável. Basicamente, o *random forest* adiciona aleatoriedade ao modelo, quando está criando as árvores. Ou seja, em vez de procurar pela melhor característica na etapa de fazer a partição de nós, o algoritmo faz essa procura em um subconjunto aleatório das características. Tal processo cria

uma diversidade que leva à geração de modelos melhores. Uma esquematização da metodologia *random forest* é ilustrada na figura abaixo, em que “x” é uma *feature* (atributo) de entrada e “k” é a classe mais popular escolhida para tal atributo, com base nas classes k_i, i de 1 a B, das B árvores.

Figura 47 - Representação do método *Random Forest*



Fonte: adaptado de Verikas et al. (2016)

4.2. Indicadores de performance em modelos de aprendizado de máquina voltados à detecção de fraude

De acordo com Stolfo et al. (1997), uma forma de avaliar a performance de um modelo de aprendizado de máquina em detectar fraudes é por meio de números como o verdadeiro positivo, verdadeiro negativo, falso positivo e falso negativo.

Em fraude, o verdadeiro positivo diz respeito ao número de transações fraudulentas (classe positiva) corretamente identificadas, enquanto o falso positivo concerne ao número de transações fraudulentas (classe positiva) que foram erroneamente classificadas como não fraudulentas (classe negativa) – por isso, recebe o nome de “falso”. O mesmo raciocínio vale para verdadeiro negativo e falso negativo, sendo negativa a classe das transações não fraudulentas.

Com tais números – verdadeiros e falsos positivos e negativos – é possível calcular indicadores como a acurácia. No entanto, não apenas a acurácia é utilizada para mensurar a performance dos métodos, como também outros indicadores que dependem de tais números.

A seguir, são apresentados indicadores que são usados com frequência para a medição da performance de detecção de fraude, conforme Montague (2004).

a) Taxa de detecção / Cobertura / Sensibilidade: diz respeito ao percentual de transações fraudulentas observadas e classificadas como fraudulentas em relação ao total de transações fraudulentas observadas;

b) Taxa de verdadeiro negativo / Especificidade: relacionado ao percentual de transações não fraudulentas observadas e classificadas como não fraudulentas em relação ao total observado de transações não fraudulentas;

c) Taxa de falso positivo / (1 – Especificidade): concerne ao percentual de transações não fraudulentas observadas e classificadas como fraudulentas em relação ao total de transações não fraudulentas observadas;

d) Relação não-fraude por fraude: significa o número de transações não fraudulentas observadas e classificadas como fraudulentas dividido pelo número de transações observadas e classificadas como fraudulentas;

e) Precisão de fraude: é o percentual de transações fraudulentas observadas e classificadas como fraude em relação ao total de transações classificadas como fraudulentas;

f) Precisão de não-fraude: é o percentual de transações não fraudulentas observadas e classificadas como não fraudulentas em relação ao total de transações classificadas como não fraudulentas;

g) Acurácia: diz respeito à quantidade de transações classificadas corretamente como fraudulentas e não fraudulentas, em relação ao observado, dividido pelo número total de transações.

Na matriz da figura a seguir, conhecida como matriz de confusão (*confusion matrix*), exibe-se, de forma resumida, os indicadores de performance supracitados e suas expressões para cálculos:

Figura 48 - Matriz de confusão

Fonte: Beraldi (2014)

Conforme Hossin e Sulaiman (2015), as vantagens da acurácia (ou taxa de erro = 1 – acurácia) são: métrica fácil de ser computada (baixa complexidade); métrica fácil de ser entendida e aplicável não apenas a problemas de classificação binácia como também a problemas de multi-classificação.

Contudo, de acordo com os mesmos autores, também há limitações em se utilizar tal indicador: produz valores menos distintos/discriminados e fornece pouca informação, precisando ser utilizado em conjunto com outros indicadores;

Ainda conforme tais autores, os indicadores mais utilizados para a medição de performance dos métodos de aprendizado de máquina são os seguintes:

Figura 49 - Indicadores de performance para métodos de *machine learning*

Métrica	Cálculo	Descrição
Acurácia	(VP+VN) / (VP+FP+VN+FN)	Taxa de previsões corretas sobre o total de previsões feitas
Taxa de Erro	(FP+FN) / (VP+FP+VN+FN)	Taxa de previsões incorretas sobre o total de previsões feitas
Sensitivity = *Recall* (sensitividade)	VP / (VP+FN)	Fração de positivos corretamente classificados (em relação ao número total de positivos)
Specificity (especificidade)	VN / (VN+FP)	Fração de negativos corretamente classificados (em relação ao número total de negativos)
Precisão	VP / (VP+FP)	Fração de positivos corretamente classificados (em relação ao total de previsões positivas feitas)
F-*Measure* (Medida "F")	2 * Precisão * *Recall* / (Precisão + *Recall*)	Média harmônica entre o *recall* e a precisão

Fonte: adaptado de Hossin e Sulaiman (2015)

4.3. Área sob a curva

A área sob a curva é uma das métricas comumente utilizadas para ranquear métodos de aprendizado de máquina voltados a classificações. Existem duas curvas principais para se utilizar o classificador AUC (*Area Under Curve*):

a) Curva ROC (*Receiver Operating Characteristic*), ou curva sensitividade x (1-especificidade);

b) Curva *precision* x *recall* (Curva PR).

De acordo com Da Silva (2006), uma curva ROC é uma demonstração bidimensional da performance de um classificador. Para comparar classificadores é preciso reduzir a curva ROC a um valor escalar. Um método comum para realizar esta redução é calcular a área abaixo da curva ROC. Como a AUC (*Area Under Curve*) é uma porção da área do quadrado unitário (espaço ROC), seus valores vão de 0.0 a 1.0.

O eixo das ordenadas leva os valores da métrica de Sensitividade enquanto o eixo das absicssas contém os valores de uma métrica chamada de Taxa de Falsos Negativos, a qual é equivalente a 1-Especificidade.

Davis e Goadrich (2006) afirmam que as curvas ROC são comumente usadas para apresentar resultados para problemas de decisão binária em aprendizado de máquina, entretanto, ao lidar com conjuntos de dados altamente distorcidos ou desbalanceados (como é o caso de um conjunto de dados de fraude em transações de crédito), as curvas *Precision-Recall* (curvas PR) fornecem um quadro mais informativo do desempenho de um algoritmo.

Em bases desbalanceadas, com um número grande de classes negativas, há um certo desinteresse na habilidado do modelo em prever/classificar tal classe corretamente. Assim, faz-se importante a utilização de métricas como a precisão e o *recall*, já que estas não apresentam o número de verdadeiros negativos em seus cálculos. Tais métricas estão focadas na predição correta da classe minoritária, isto é, a classe positiva, das transações fraudulentas. Nos gráficos de curva PR, a precisão se encontra no eixo das ordenadas e o *recall* no eixo das abscissas.

Finalmente, em consonância com as palavras de Zhu e Wang (2010), os valores de AUC podem ser classificados conforme os itens a seguir:

a) Excelente: 0.9 < AUC <= 1;

b) Bom: 0.8 < AUC <= 0.9;

c) Baixo: 0.7 < AUC <= 0.8;

d) Ruim: AUC <= 0.7.

Assim, é importante buscar um método que produza resultados excelentes em termos de AUC, ou seja, este deve estar situado entre 0.9 e 1. Contudo, ser um método excelente em termos de AUC é uma condição mínima, mas não suficiente. É imprescindível que a área sob a curva *Precision-Recall* resulte em um número próximo de 1 para o método ser considerado de alto desempenho.

5. METODOLOGIA

A metodologia de um trabalho científico pode estar norteada por duas vertentes, chamadas de métodos qualitativos e métodos quantitativos. Ambos os métodos qualitativos e quantitativos precisam ser delineados de modo a alcançar os objetivos propostos, fornecendo resultados com a possibilidade de confirmar ou negar as hipóteses lançadas.

Os métodos qualitativos descrevem uma relação entre o objetivo e os resultados que não podem ser interpretadas através de números, nomeando-se como uma pesquisa descritiva. Todas as interpretações dos fenômenos são analisadas indutivamente (FERNANDES e GOMES, 2003).

Por outro lado, estão os métodos quantitativos que acreditam que tudo deve ser quantificado para promover resultados confiáveis. Trabalham com dados numéricos e técnicas estatísticas tanto para classificar como para analisar os resultados, desta forma são mais empregados em pesquisas nas áreas biomédicas e exatas, nomeando-se como uma pesquisa tanto descritiva como analítica (FERNANDES e GOMES, 2003).

Neste trabalho, devido a sua natureza de dados numéricos (transações de cartão de crédito), utilizaram-se métodos quantitativos.

5.1. Identificação do problema e estratificação

Com base nos dados fornecidos na etapa de relevância do problema, item 1.3, é possível identificar o problema em questão e, sobretudo, estratificá-lo em menores problemas (ou subproblemas) de modo a se ter um ponto focal de atuação. Com o problema estratificado, é mais simples a obtenção de sua(s) causa(s) e, por fim, a aplicação de uma ou mais soluções para o mesmo.

Para tanto, foram utilizados os dados de proporção de volume transacionado, de fraude por canal de transação e suas taxas de declínio, em uma simples análise realizada no item 6.2 de aplicação da metodologia.

5.2. Causa do problema

Em seguida, identificou-se a causa principal que leva à ocorrência dos estratos do problema. A causa é a mesma para os dois, e foi brevemente explorada no subcapítulo 6.2 deste trabalho por meio de uma tabela dos “5 Porquês”.

Resumidamente, esta causa está relacionada à baixa asservidade na detecção de transações fraudulentas e de transações legítimas, prejudicando a decisão de quando negar uma transação e quando aceitar (pois, o que está ocorrendo, é que transações fraudulentas não estão sendo identificadas e transações legítimas estão sendo bloqueadas por errônea suspeita de fraude).

Em seguida, no subcapítulo 5.3, explica-se como ocorre o enfoque do problema. É imprescindível saber em qual categoria de transações *online* estão ocorrendo as perdas com fraude e com transações legítimas erroneamente recusadas, já que existem diversas categorias de transações desse tipo. Assim, faz-se importante realizar um *drill-down* e encontrar um foco, de modo a não dispender recursos em outras categorias de transações em que os estratos do problema não ocorrem.

5.3. Enfoque do problema

Tendo o problema deste trabalho – relacionado às altas perdas financeiras em transações de cartão de crédito em ambiente *online* – utilizou-se uma das ferramentas da qualidade, o **gráfico de Pareto**, para identificar qual ou quais partes desse problema devem ser endereçadas.

Conforme Carvalho e Paladini (2012), o gráfico de Pareto foi criado por Vilfredo Pareto, nascido em Paris, o economista considerado político e sociólogo ao desenvolver o estudo sobre a distribuição de renda de seu país, percebeu que a distribuição não se dava de maneira igualitária onde a riqueza nacional estava concentrada numa pequena parcela da população, sendo assim, este modelo de Pareto foi adaptado por J. M. Juran transformando-se numa das ferramentas mais conhecidas da qualidade.

Assim, de acordo com Johnston, Chambers e Slack (2002), o gráfico de Pareto é uma técnica relativamente direta que classifica os itens de informação nos tipos de problemas, por ordem de importância, podendo ser usado para destacar áreas em que as investigações poderão ser úteis.

O gráfico de Pareto dispõe a informação de forma a permitir a concentração dos esforços para melhoria nas áreas one os maiores ganhos podem ser obtidos (WERKEMA, 2006).

Ainda, de acordo com o mesmo autor – Werkema (2006) – o gráfico de Pareto é um gráfico de barras no qual as barras são ordenadas a partir da mais alta até a

mais baixa e é traçada uma curva que mostra as porcentagens acumuladas de cada barra.

5.4. Soluções para atuação nas causas do problema

Com os estratos do problema (alta perda financeira em transações de cartão de crédito em ambiente *online*) – resumidos, basicamente, às transações fraudulentas erroneamente aceitas e às transações não fraudulentas erroneamente recusadas – e com a causa desses estratos conhecida, a proposta deste trabalho de formatura é utilizar um modelo de detecção de fraude, baseado em técnicas estatísticas associadas a *machine learning*, para garantir alta assertividade na detecção do que é de fato fraude (e pode ser bloqueada no momento da transação) e do que não é. Compararam-se duas técnicas neste documento: regressão logística e *random forest*.

No âmbito do aprendizado de máquina, conforme as palavras de Perlich (2017) – professora adjunta na Universidade de Nova Iorque – para uma matéria da Forbes, há um consenso nomeado de “teorema do Não Há Almoço Grátis" (*No Free Lunch Theorem*). Basicamente, tal ideia dita que nenhum algoritmo é o melhor para todos os problemas ou funciona melhor para aprendizado supervisionado. Por exemplo, as redes neurais ou regressão logística não são sempre melhores que as árvores de decisão ou vice-versa. Há muitos fatores a serem considerados, como o tamanho e a estrutura do *dataset*. Como resultado, vários algoritmos diferentes devem ser testados no problema. Justamente por isso existe um “conjunto de testes” (*testset*) de dados para avaliar o desempenho (baseado em algumas métricas pré-selecionadas como tempo de execução, acurácia etc) e selecionar o vencedor, conforme feito mais à frente neste trabalho.

Assim, esses dois métodos – regressão logística e *random* forest (floresta formada por árvores de decisão) – foram escolhidos para comparação por serem comumente usados em cenários de base desbalanceada, como ocorre com bases de transações de cartão de crédito, isto é, em que a maior parte das linhas do *dataset* é da classe não fraudulenta e menor parte é da classe fraudulenta. O método de regressão logística é ainda o mais comum dentre todos os métodos supervisionados de classificação no âmbito do aprendizado de máquina. O método de árvores de decisão também é um método comumente utilizado para problemas de classificação, e sua “forma” *ensembled* (ou seja, a técnica *random forest*, que reúne várias árvores

de decisão) apresenta melhores resultados (em termos de métricas) quando em comparação a uma árvore de decisão sozinha.

Ademais, algoritmos como a regressão ou árvores de decisão / *random forest* (as quais contém árvores de decisão) produzem resultados fáceis de interpretar, enquanto outros métodos são menos interpretáveis.

De acordo com a Cybersource (2016), uma análise por meio de regressão, como a regressão logística, é uma técnica estatística popular de longa data que mede a força das relações de causa-efeito em *datasets* estruturados. A análise de regressão tende a se tornar mais sofisticada quando aplicada à detecção de fraudes devido ao número de variáveis e ao tamanho dos conjuntos de dados. Ela é uma técnica que agrega valor ao avaliar o poder preditivo de variáveis individuais ou combinações de variáveis. Nessas técnicas, as transações autênticas são comparadas com as transações fraudulentas para criar um algoritmo. Este modelo (algoritmo) irá prever se uma nova transação é fraudulenta ou não.

Já as árvores de decisão, ainda de acordo com a Cybersource (2016), constituem uma família de algoritmos de aprendizado de máquina usada para automatizar a criação de regras para tarefas de classificação. Os algoritmos da Árvore de Decisão podem ser usados para problemas de modelagem preditiva de classificação ou regressão. No caso deste trabalho, tais algoritmos estão sendo usados para classificação, assim como a técnica de regressão logística (que, apesar do nome, não é de regressão – e possui tal nome apenas por estar embasado no conceito estatístico de regressão linear). As árvores de decisão são essencialmente um conjunto de regras que são treinadas usando exemplos de fraude que os clientes estão enfrentando. A criação de uma árvore ignora recursos irrelevantes e não requer uma ampla normalização dos dados. Uma árvore pode ser inspecionada e, assim, pode-se entender por que uma decisão foi tomada seguindo a lista de regras acionadas.

A técnica Random Forest usa uma combinação de várias árvores de decisão para melhorar o desempenho da classificação. Isso permite suavizar o erro que pode existir em uma única árvore. Portanto, essa técnica aumenta o desempenho geral e a precisão do modelo, ao mesmo tempo em que mantém a capacidade de interpretar os resultados. Os tempos de execução aleatórios da *random forest* são bastante rápidos e são capazes de lidar com dados desequilibrados e ausentes (*missing data*).

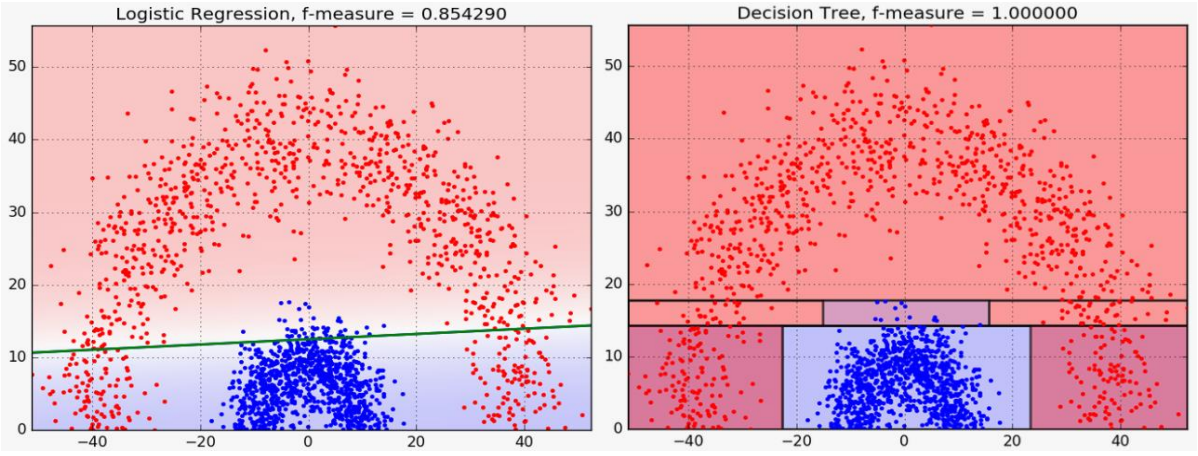
A seguir, são mostrados algumas notáveis vantagens e uma desvantagem da utilização de *random forests*, conforme a Altexsoft (2017):

a) **Vantagens**: além de sua simplicidade e velocidade, as *random forests* podem ser usadas com diferentes tipos de dados, incluindo números de cartão de crédito, datas, endereços IP, códigos postais etc. Eles são considerados preditores precisos que podem funcionar mesmo com conjuntos de dados que possuem registros ausentes;

b) **Desvantagem**: às vezes, os usuários da técnica enfrentam alguns problemas relativos a *overfitting* – o que significa que o modelo lembra muito dos padrões do *dataset* de treinamento e, assim, não consegue fazer previsões certeiras com os dados que recebe futuramente no momento do teste.

Ademais, de acordo com exemplos estudados por Lee (2016), a regressão logística e as árvores de decisão método *random* forest diferem na maneira como geram os limites de decisão, isto é, as linhas que são desenhadas para separar as diferentes classes. Para ilustrar essa diferença, analisou-se os resultados dos dois tipos de modelo no seguinte problema de duas classes:

Figura 50 - Primeiro exemplo de comparação dos limites de decisão dos métodos de Regressão Logística e *Random Forest*

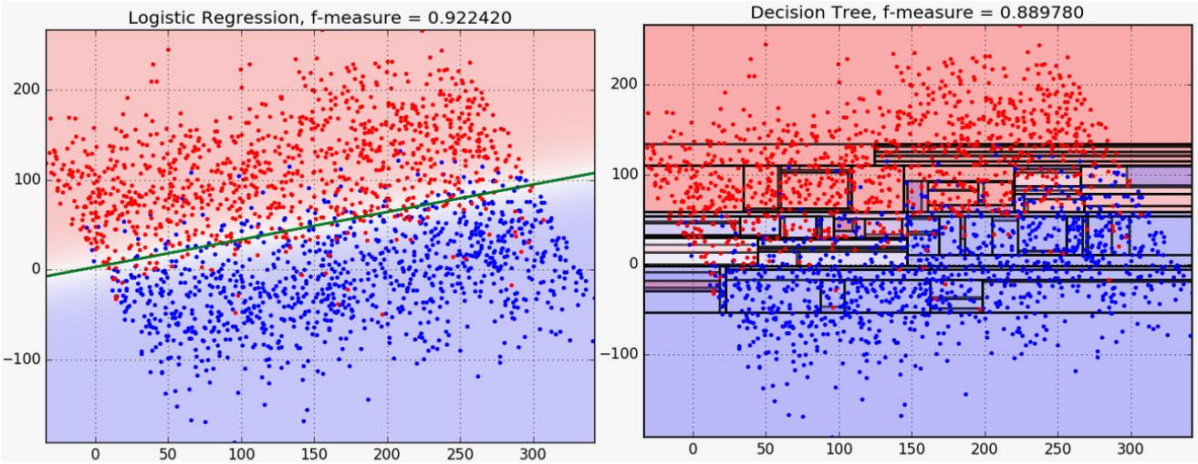


Fonte: Lee (2016)

Lee afirma que as árvores de decisão dividem o espaço em regiões cada vez menores, enquanto a regressão logística ajusta uma única linha para dividir o espaço

exatamente em duas. É claro que, para dados de maior dimensão, essas linhas se generalizariam e se tornam planos, hiperplanos etc. Um único limite linear pode, às vezes, ser limitante para a regressão logística. Neste exemplo, onde as duas classes são separadas por um limite decididamente não linear, vemos que as árvores podem capturar melhor a divisão, levando a um desempenho de classificação superior. No entanto, quando as classes não são bem separadas, as árvores são suscetíveis a superdimensionar os dados de treinamento, de modo que o limite linear simples da Regressão Logística se generalize melhor.

Figura 51 - Segundo exemplo de comparação dos limites de decisão dos métodos de Regressão Logística e *Random Forest*



Fonte: Lee (2016)

Por fim, a cor de fundo desses gráficos representa a confiança de predição. Cada nó de uma árvore de decisão atribui um valor de confiança constante a toda a região que ele abrange, levando a uma aparência de retalhos de valores de confiança por todo o espaço do gráfico. Por outro lado, a confiança de predição para a regressão logística pode ser calculada em forma fechada para quaisquer coordenadas de entrada arbitrárias, resultando em um resultado mais refinado e, por conseguinte, oferecendo maior segurança ao usuário no que concerne os valores de confiança do modelo.

5.5. Programação e bibliotecas

De modo a carregar a base de dados, analisar suas variáveis e, por fim, treinar e testar os métodos de classificação supervisionada de regressão logística e *random*

forest, foi utilizada a linguagem Python em um ambiente criado pela aplicação Jupyter Notebook, pertencente ao Anaconda Navigator.

Conforme Corrêa (2019), nos últimos anos, Python consolidou-se como uma das tecnologias mais difundidas na área ciência de dados, apesar de não ter sido originalmente projetada para este fim. Essa conquista não se deveu apenas ao fato de a linguagem facilitar a criação de programas de “rosto bonito”. Na realidade, o sucesso do Python para ciência de dados está relacionado a outras de suas características, conforme apresenta-se a seguir:

a) Como Python é uma linguagem interpretada, os iniciantes podem aprender alguns comandos e começar a executar comandos muitas vezes complexos (ex.: aplicar funções matemáticas e estatísticas sobre conjuntos de dados) quase que imediatamente, sem esbarrar em problemas relacionados à compilação de código. Ademais, o interpretador Python pode ser utilizado de forma interativa, onde cada comando digitado é imediatamente traduzido e executado. Isto oferece aos programadores uma forma ágil e simples para examinar em tempo real os resultados intermediários obtidos em qualquer passo de um processo de análise de dados;

b) Python é uma linguagem livre (open source). No *website* da *Python Software Foundation*, é possível baixar gratuitamente o arquivo que instala o interpretador Python e a sua biblioteca padrão (a famosa *standard library*). Juntos, estes componentes formam o “coração” do ambiente Python, oferecendo um rico conjunto de estruturas de dados (como listas e dicionários) e centenas de módulos voltados para a execução dos mais diversos tipos de tarefas, desde o uso de funções matemáticas e estatísticas até o processamento de arquivos texto em diferentes formatos (CSV, JSON etc.);

c) A linguagem Python pode ser facilmente estendida através da incorporação de outros pacotes. Atualmente, existem milhares de pacotes disponíveis no repositório central do Python (*Python Package Index – PyPI*). Muitos deles são voltados para ciência de dados, tais como o *NumPy* (manipulação de vetores e matrizes), *SciPy* (integração e cálculo

numérico), *pandas* (manipulação de DataFrames), *matplotlib* (geração de gráficos) e *sklearn* (algoritmos de mineração de dados e aprendizado de máquina).

Para tanto, dentro do ambiente de programação, foram importadas bibliotecas pelo comando "*import*", necessárias para plotagem de gráficos, cálculos envolvendo tabelas, treinamento e teste de cada um dos métodos a serem estudados etc. Dentre essas bibliotecas, estão a *sklearn*, *matplotlib*, *NumPy* e *pandas*.

5.6. Levantamento de dados (*dataset*)

A base de dados a ser utilizada para treinamento e testes dos modelos de regressão logística e *random forest* foi obtida por meio da plataforma Kaggle, sendo esta um repositório de projetos de data science e *datasets* para utilização em modelos de aprendizado de máquina.

Tal base teve seus dados coletados e analizados durante uma pesquisa colaborativa realizada pela *Worldline* (processadora de pagamentos por cartão de crédito) e a ULB (*Université Libre de Bruxelles*) com foco em detecção de fraude e *big data mining*.

O *dataset* e suas variáveis são analisados durante a etapa de aplicação da metodologia, especificamente no item 6.4.

5.7. Tratamento do *dataset* e aplicação das soluções

É comum, no âmbito de *machine learning*, dividir o conjunto de dados em um subconjunto de treinamento e em um de teste. A razão para tal ação é evidente: se um indivíduo tentasse avaliar seu modelo com base em dados utilizados justamente para treinar esse modelo, certamente estaria fazendo algo irrealista. O ponto principal de um modelo de aprendizado de máquina é poder trabalhar com dados imprevisíveis, ou seja, um modelo não pode utilizar os mesmos dados que usou no seu treinamento para o seu teste, senão estaria simplesmente coincidindo dados e não utilizando o treinamento para prever uma situação nova com dados nunca antes vistos.

Assim, faz-se necessária a divisão da base em dados de treinamento e em dados de testes. Geralmente, utiliza-se a divisão 80:20, em que 80% do *dataset* é utilizado para a etapa de treinamento do modelo de classificação e 20% para a etapa de testes. Em seguida, faz-se o *oversampling*, que é o balanceamento dos dados de

teste de modo a não deixá-lo desequilibrado em termos de ter mais transações não fraudulentas (classe 0) em relação às fraudulentas (classe 1). Tal processo é exibido no item 6.5 deste trabalho.

Em seguida, com a base já dividida e com a sub base de testes já tratada, executa-se os métodos de aprendizado de máquina (regressão logística e *random forest*) mediante utilização da biblioteca *sklearn* do *Phyton*. Tais passos serã oexecutados nos itens 6.6 e 6.7 do trabalho, sendo os resultados apresentados no capítulo 7.

6. APLICAÇÃO DA METODOLOGIA

Este capítulo apresenta o desenvolvimento da aplicação da metodologia apresentada no capítulo anterior. Identificou-se o problema a ser resolvido, e o mesmo foi estratificado. Posteriormente, utilizou-se o gráfico de Pareto para auxílio no enfoque do problema e, por fim, detectou-se a causa do problema elencou-se duas possíveis soluções a esta, no âmbito de *machine learning*. Tais soluções foram aplicadas em uma base de transações acadêmica, de modo a compreender qual apresenta a melhor performance em termos das métricas apresentadas no capítulo anterior de revisão bibliográfica.

6.1. Identificação do problema e estratificação

Retomando o que foi visto no item 1.3 de relevância do problema, identificou-se uma grande disparidade das transações de cartão presente frente às transações de cartão não presente, em termos de volume (em USD). Praticamente, em 2017 e 2018, 80% das transações se mantiveram no canal de Cartão Presente e 20% no canal de Cartão Não Presente:

Mesmo representando apenas 20% do volume total transacionado em 2017 e 2018, as transações no canal de Cartão Não Presente apresentam as maiores taxas de declínio (relação de transações declinadas pelo total de transações) de transação, com média de 36% do início de 2017 até o final de 2018. Por outro lado, observa-se uma média de 4% para o canal de Cartão Presente nesse mesmo período e 11% para o canal consolidado (CP+CNP).

Além disso, além de ser a classe transacional com a maior taxa de declínio, o CNP também apresenta as maiores perdas com fraude. Aproximadamente 83% do volume de fraude está concentrada no canal de Cartão Não Presente e 17% no canal de Cartão Presente.

Assim, é possível notar, apenas com esta simples estratificação, que o problema de altas perdas financeiras no canal CNP se concentra tanto na alta taxa de declínio das transações de CNP (transações erroneamente declinadas) como também nos altos índices de fraude (transações fraudulentas não detectadas – e aprovadas – que deveriam ser declinadas).

6.2. Causa do problema

Ao se analisar o **problema** no item anterior – **altas perdas financeiras em transações de cartão de crédito no ambiente CNP (*online)*** – foi possível estratificá-lo em duas partes:

1. Perdas financeiras com transações fraudulentas erroneamente aceitas;
2. Perdas financeiras com transações legítimas erroneamente recusadas.

Assim, a **causa** do problema pode ser identificada, de maneira simples, por meio da técnica dos “5 Porquês”:

Tabela 11 - Técnica dos "5 Porquês”

Técnica dos "5 Porquês"	
1. Por que há altas perdas financeiras em transações de cartão de crédito CNP (*online)*?	Porque há muitas transações fraudulentas e falsos-positivos na categoria de transações CNP (*online)*
2. Por que há muitas transações fraudulentas e falsos-positivos na categoria de transações CNP (*online)*?	Porque a atual estratégia de detecção a fraudes em transações CNP (*online)* não é adequada
3. Por que a atual estratégia de detecção a fraudes em transações CNP (*online)* não é adequada?	Porque os métodos de prevenção e detecção a fraudes em transações CNP (*online),* presentes na atual estratégia, não são os adequados
4. Por que os métodos de prevenção e detecção a fraudes em transações CNP (*online),* presentes na atual estratégia, não são os adequados?	**(Foco na detecção de fraudes)** Porque são métodos que apresentam baixa acurácia na identificação de padrões em transações no cenário CNP (*online)*
5. Por que os métodos atuais apresentam baixa acurácia na identificação de padrões em transações no cenário CNP (*online)*?	Porque são métodos baseados em regras, os quais recusam apenas transações cujas variáveis correspondem a valores (ou estão dentro de intervalos) pré definidos pelo Time Operacional. Tais métodos acabam por deixar passar uma quantidade expressiva de transações fraudulentas (falsos negativos) e bloquear transações legítimas (falsos positivos)

Fonte: elaboração do autor

É imprescindível que o método a ser aplicado para detecção de fraudes, na categoria de transações CNP (*online*) apresente uma assertividade considerável – isto é, uma acurácia de, no mínimo, 90%. As técnicas de aprendizado de máquina, conforme exibido no subcapítulo 5.4, são bastante adequados para detecção de padrões em dados de natureza desbalanceada, como os dados de transações de cartão de crédito *online*. No entanto, necessita-se dar enfoque ao problema, pois existem várias categorias de transações *online,* como as realizadas em *e-commerce*, por exemplo.

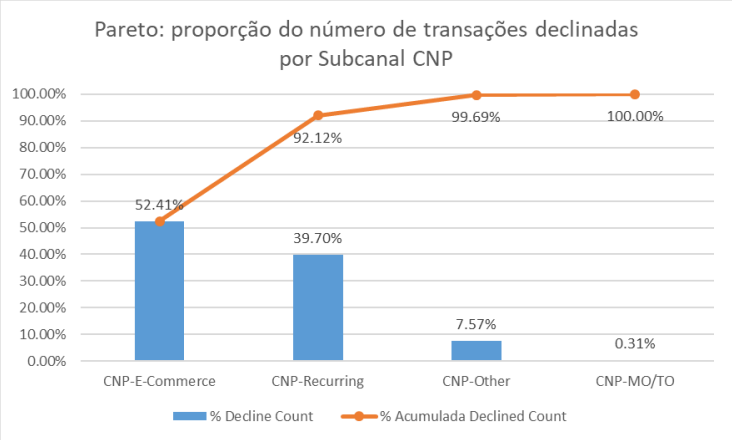
Assim, no item 6.2 a seguir, foi aplicada a metodologia adotada para dar enfoque ao problema, dentro do canal CNP, tanto para o primeiro estrato do problema (altas taxas de declínio no canal CNP, de transações *online*) como também para o segundo estrato (altos índices de fraude para as transações do canal CNP, de transações *online*).

6.3. Enfoque do problema

Já sabendo a causa do problema, é importante saber em que categoria de transações o problema está concentrado. Já se sabe que o problema ocorre em transações *online*, mas existem diversas categorias de transações *online*. Assim, faz-se necessário focalizar o problema.

Aplicando-se o gráfico de Pareto, explorado teoricamente no item 5.2, para a obtenção da proporção do número de transações declinadas por subcanal de transação do canal CNP (*online*), no período de 2018, no Brasil, tem-se o seguinte:

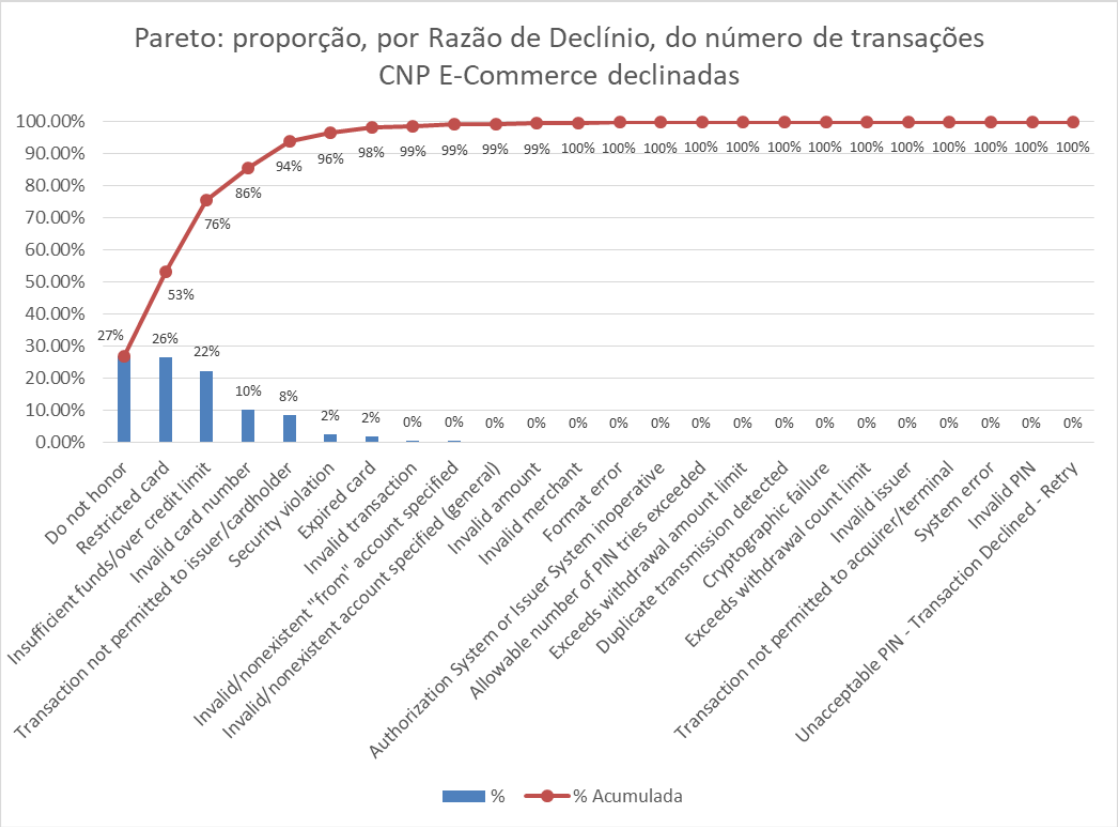
Figura 52 - Gráfico de Pareto com a proporção do número de transações declinadas por subcanal CNP (%) no Brasil em 2018



Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Ou seja, em termos de quantidade de transações para o canal CNP, o subcanal **CNP *E-Commerce*** é o mais expressivo, isto é, apresenta o maior número de transações CNP declinadas. Contudo, para se entender a razão de tais declínios, deve-se analisar as razões de declínio. Para tanto, foi criado um gráfico de Pareto com as proporções de transações CNP E-Commerce declinadas, dispostas por Razão de Declínio (Brasil, 2018):

Figura 53 - Gráfico de Pareto com a proporção do número de transações CNP E-Commerce declinadas, por razão de declínio, no Brasil em 2018



Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Com base no gráfico acima, é possível notar que as principais razões de declínio são:

- "*Do Not Honor*" (transação não honrada pelo portador – **motivo comumente associado a suspeita de fraude**);

- "*Restricted Card*" (cartão restrito, isto é, limitado para realizar certas transações) e, por fim;

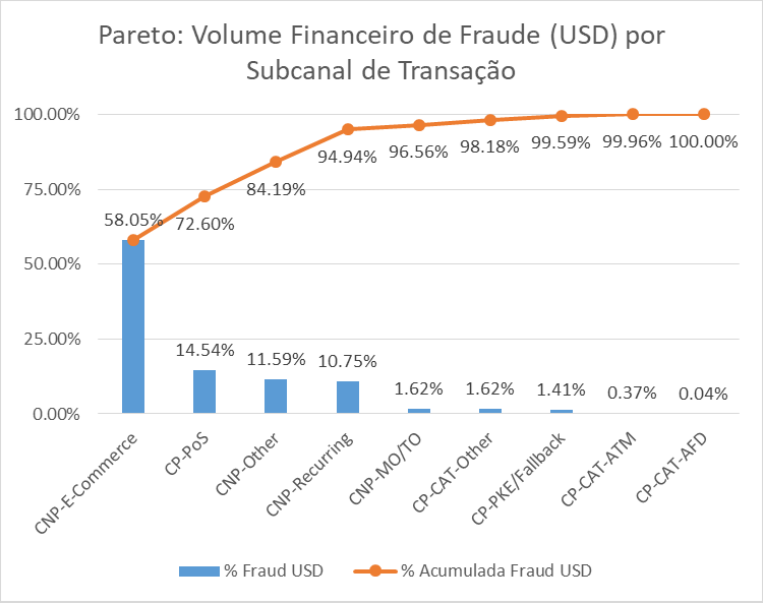
- "*Insufficient funds / over credit limit*" (fundos insuficientes / limite do cartão atingido).

Portanto, como o primeiro motivo de declínio está relacionado a fraude, a maioria das transações CNP E-Commerce está sendo declinada por suspeita de fraude – e o objetivo deste trabalho é, justamente, melhorar a assertividade do

modelo de detecção de fraude de forma no que tange a classificação das transações como fraudulentas ou não, de forma a evitar falsos positivos (nesse caso, evitar transações declinadas por suspeita de fraude sendo que não são fraude) e falsos negativos.

Ainda, por meio do gráfico de Pareto, é possível analisar o volume de fraude por subcanal transacional (Brasil, 2018):

Figura 54 - Gráfico de Pareto com a porcentagem do volume financeiro de fraude por subcanal transacional no Brasil em 2018



Fonte: elaboração do autor com dados autorizados pela empresa onde estagiou

Nota-se que, em termos de volume financeiro, assim como o número de transações declinadas, a fraude se concentra mais no subcanal **CNP *E-Commerce***, com uma proporção ainda maior que qualquer outro subcanal, inclusive o CP-*PoS*. Tal volume expressivo de fraude poderia ser tratado se o modelo de detecção de fraudes fosse mais assertivo. É fato que, se o modelo tivesse uma assertividade maior, o número de falsos negativos seria menor, ou seja, o número de transações fraudulentas erroneamente classificadas pelo modelo como “não fraude” seria menor, diminuindo a proporção das transações fraudulentas. Além disso, o número de falsos positivos também seria menor, diminuindo o número de transações não fraudulentas classificadas erroneamente como “fraudulentas” e, portanto, declinadas.

A proposta do presente trabalho é, então, atuar em um modelo que enderece tais necessidades – **melhorar a assertividade da detecção de transações de cartão de crédito fraudulentas e legítimas, no ambiente *online*, na categoria *e-commerce***.

6.4. Programação e bibliotecas

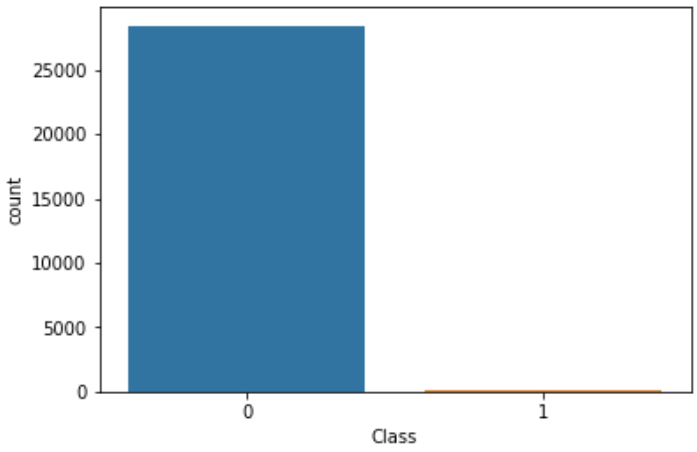
Dentro do ambiente de programação, foram importadas as seguintes bibliotecas pelo comando “import”, necessárias para plotagem de gráficos, cálculos envolvendo tabelas, treinamento e teste de cada um dos métodos a serem estudados etc:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from imblearn.over_sampling import SMOTE
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
```

6.5. Análise do *dataset*

O *dataset* contém transações efetuadas por portadores de cartões de crédito europeus, em dois dias de setembro de 2013, captadas pelos sistemas de processamento da Worldline. Do total de 284807 transações, 492 são fraudulentas. Assim, pode-se dizer que o conjunto de dados é altamente desbalanceado, já que a classe positiva (fraude) representa 0.172% de todas as transações. Conforme exposto por Rocca (2019), um *dataset* já é desbalanceado quando uma classe representa 10% dele e a outra 90%. A figura abaixo, gerada via Python, ilustra esse desequilíbrio do *dataset* de fraude:

Figura 55 - Desequilíbrio entre classes no *dataset*



Fonte: elaboração do autor

Ademais, a base contém apenas variáveis de entrada numéricas, que são o resultado de uma transformação chamada de PCA (*Principal Component Analysis*), um procedimento matemático baseado em transformação ortogonal para converter um conjunto de variáveis possivelmente correlacionadas em um conjunto de valores linearmente não correlacionados. Infelizmente, devido a problemas de confidencialidade, a Worldline não pôde fornecer mais informações sobre os dados e muito menos o valor original das variáveis.

As variáveis de atributo (*features*) V1, V2, ... V28 são os principais componentes obtidos pela transformação PCA. As únicas variáveis de atributo que não sofreram a transformação PCA são “Tempo” (*time*) e “Valor” (*value*). O atributo “Tempo” diz respeito aos segundos decorridos entre cada transação e a primeira transação do conjunto de dados. Já o atributo “Valor” concerne ao montante financeiro da transação. Já a “Classe” (*class*) é a variável de resposta e recebe o valor 1 no caso da transação ser fraude e 0 caso contrário.

Antes do *dataset* ser utilizado para treinamento e testes, deve-se verificar se alguma variável é nula em alguma(s) linha(s). Felizmente, nenhuma variável assume valor nulo, conforme mostra a figura a seguir, em que “data” é:

Figura 56 - Variáveis não nulas no *dataset*

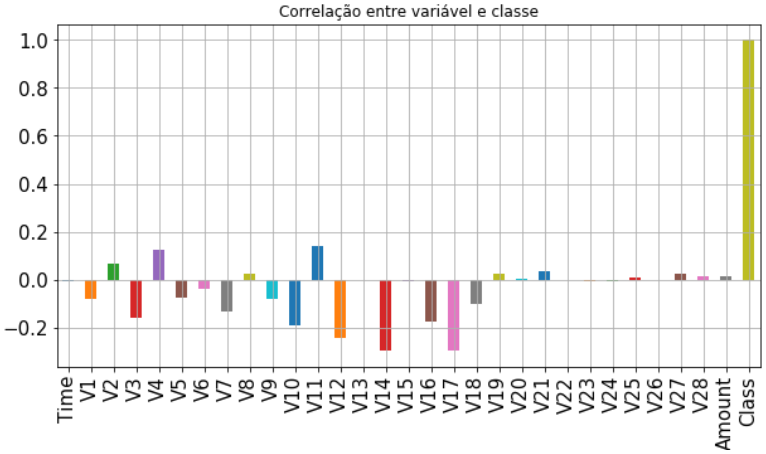
```
In [34]: data.isnull().any()

Out[34]: Time      False
         V1        False
         V2        False
         V3        False
         V4        False
         V5        False
         V6        False
         V7        False
         V8        False
         V9        False
         V10       False
         V11       False
         V12       False
         V13       False
         V14       False
         V15       False
         V16       False
         V17       False
         V18       False
         V19       False
         V20       False
         V21       False
         V22       False
         V23       False
         V24       False
         V25       False
         V26       False
         V27       False
         V28       False
         Amount    False
         Class     False
         dtype: bool
```

Fonte: elaboração do autor

O ticket médio das transações é de US$ 88.35, enquanto a transação de maior valor da base de dados é de US$ 25691.16. As características estatísticas de cada variável como média, desvio padrão, quartis, entre outros, se encontram no Apêndice A. Já os histogramas gerados para cada variável são exibidos no Apêndice B.

Na figura a seguir, ilustra-se a correlação entre cada um dos atributos e a classe da base de dados. Percebe-se maior correlação positiva para as variáveis V4 e V11:

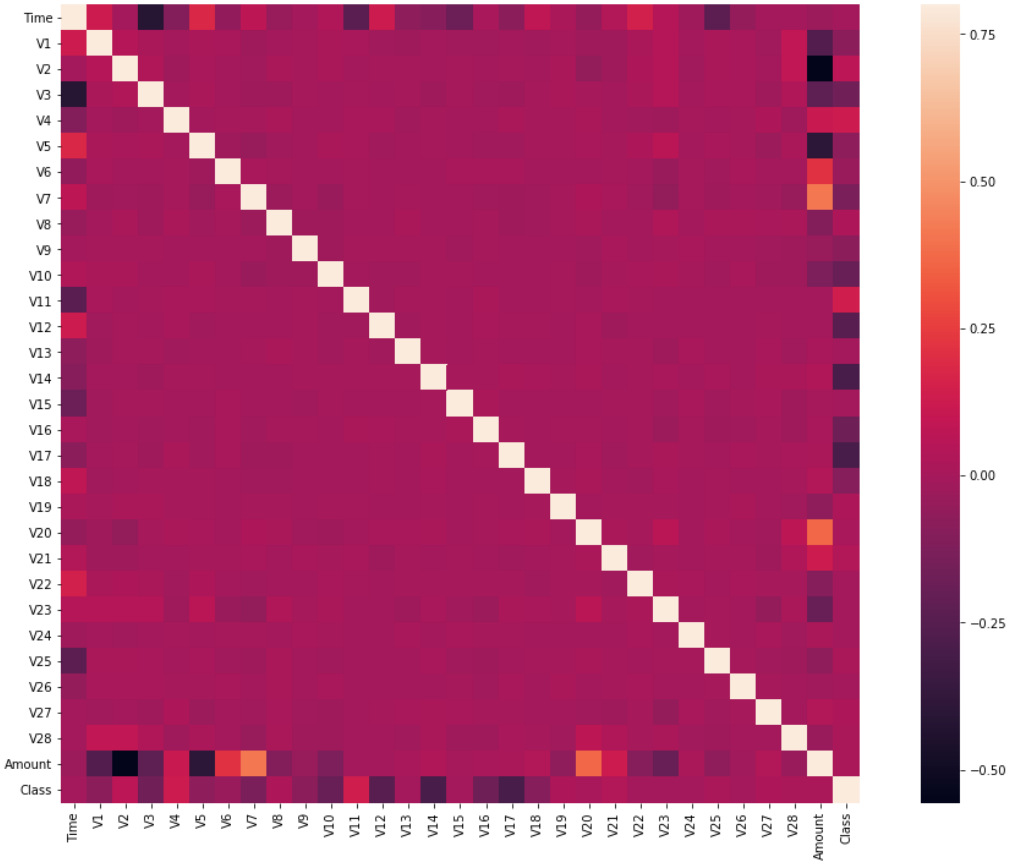
Figura 57 - Correlação entre variável e classe no *dataset*



Fonte: elaboração do autor

Já o mapa de calor a seguir ilustra a correlação entre cada uma das variáveis. As correlações Amount-V7, Amount-V20 merecem destaque:

Figura 58 - Correlações entre variáveis e classes no dataset



Fonte: elaboração do autor

6.6. Divisão e tratamento dos dados

Para a divisão dos dados estudados, utilizou-se a proporção 80:20, em que 80% do *dataset* é utilizado para a etapa de treinamento do modelo de classificação e 20% para a etapa de teste. Conforme Fuentes (2018), é comum utilizar a divisão 80:20 em bases de dados não tão grandes. Reservar 20% de uma base de 5 milhões de transações somente para teste seria desnecessário, visto que haveria 1 milhão de transações apenas para teste. Não existe uma regra específica para definir uma divisão, mas é sempre bom separar uma grande quantidade de dados para a etapa de treinamento, pois é nela que o modelo aprende os padrões dos dados e consegue, assim, realizar previsões futuramente – sobretudo com os dados de teste.

Por fim, foi criado um novo conjunto de dados, contendo os dados do subconjunto de testes, contudo, aplicando-se o método de *oversampling* para balanceamento desses dados. O número de amostras, para cada classe, do conjunto de testes antes e depois do *oversampling* se encontra na figura abaixo:

Figura 59 - Número de variáveis de cada classe após *oversampling* da base de testes

```
Conjunto de Testes: linhas com Classe 0
284315
Conjunto de Testes: linhas com Classe 1
492
Conjunto de Testes: total de linhas
284807


Conjunto de Testes Oversampled: linhas com Classe 0
227454
Conjunto de Testes Oversampled: linhas com Classe 1
227454
Conjunto de Testes Oversampled: total de linhas
454908
```

Fonte: elaboração do autor

6.7. Regressão Logística

De forma a aplicar as funções do módulo de regressão logística, a importação das bibliotecas do item 5.3.1 se faz fundamental.

A biblioteca de código aberto que contém uma gama de funções e metodologias de classificação em aprendizado de máquina é a *sklearn* (*science kit learn*) sua importação ocorre mediante a aplicação do seguinte comando:

from *sklearn.linear_model* ***import*** *LogisticRegression*

Desenvolvido pelos SCIKIT-LEARN DEVELOPERS (2019), o manual dessa biblioteca contém todos os detalhes acerca da importação das funções e aplicação dos métodos de aprendizado de máquina.

Assim, ao importar o módulo *LogisticRegression* da seção *linear_model* da biblioteca *sklearn*, permite-se a utilização de funções relacionadas ao método de regressão logística.

Em seguida, declara-se a variável ***clf_log*** como sendo a instância de classificação relacionada ao modelo de regressão logísitca. Define-se, como valor a esta instância, a função de regressão logística:

clf_log = ***LogisticRegression***(*random_state*=0)

Como *output* desta ação, tem-se a descrição da função inteira com suas variáveis de *input* (das quais foi definido apenas o *random_state*=0):

LogisticRegression (C=1.0, class_weight=None, dual=False, fit_intercept=True, intercept_scaling=1, max_iter=100, multi_class='warn', n_jobs=None, penalty='l2', random_state=0, solver='warn', tol=0.0001, verbose=0, warm_start=False)

Em seguida, agora que a variável ***clf_log*** assumiu o valor da instância ***LogisticRegression***, faz-se o treinamento do modelo com a base de treinamento *oversampled* que foi obtida no item 5.3.3, mediante o uso da função *fit* da instância ***LogisticRegression***, conforme comando abaixo:

clf_log.*fit*(***os_features***,***os_labels***)

Em que ***os_features*** são as variáveis da base de dados de teste *oversampled* e ***os_labels*** são as variáveis de classe dessa mesma base.

Após o treinamento, armazena-se o valor atual das variáveis-classe do *dataset* de testes na variável ***backup*** e executa-se a função *predict* nas *features* (variáveis não classificatórias) dessa mesma base, a qual foi gerada no item 5.3.3 (e representando 20% da base de dados original). Também é definida uma variável *logisticprediction* para armazenar os *outputs* da predição do modelo:

backup = ***labels_test***
logisticpredictions = ***clf_log***.*predict*(***features_test***)

Finalmente, após o modelo ser treinado e a predição ter sido realizada, já se tem os resultados produzidos pela regressão logística. Com tais resultados, é

possível gerar a matriz de confusão do modelo e obter números como falsos e verdadeiros positivos e negativos, além de calcular outras métricas com tais valores.

Em seguida, a técnica *random forest* é executada e seus resultados são comparados com os resultados da *regressão logística* no capítulo 7.

6.8. Random Forest

De forma a aplicar as funções do módulo de *random forest*, a importação das bibliotecas do item 5.3.1 também se faz fundamental assim como foi para a aplicação da regressão logística. A importação ocorre mediante a aplicação do seguinte comando:

from *sklearn.ensemble* ***import*** *RandomForestClassifier*

Assim, ao importar o módulo *RandomForestClassifier* da seção *ensemble* da biblioteca *sklearn*, permite-se a utilização de funções relacionadas ao método *Random* Forest de classificação.

Em seguida, declara-se a variável ***clf_rf*** como sendo a instância de classificação relacionada ao modelo de *Random Forest*.

Define-se, como valor a esta instância, a função *RandomForestClassifier*:

clf_rf= ***RandomForestClassifier***(*random_state*=0)

Como *output* desta ação, tem-se a descrição da função inteira com suas variáveis de *input* (das quais foi definido apenas o *random_state*=0):

RandomForestClassifier (bootstrap=True, class_weight=None, criterion='gini', max_depth=None, max_features='auto', max_leaf_nodes=None, min_impurity_decrease=0.0, min_impurity_split=None, min_samples_leaf=1, min_samples_split=2, min_weight_fraction_leaf=0.0, n_estimators=10, n_jobs=None, oob_score=False, random_state=0, verbose=0, warm_start=False)

Em seguida, agora que a variável ***clf_rf*** assumiu o valor da instância ***RandomForestClassifier***, faz-se o treinamento do modelo com a base de

treinamento *oversampled* que foi obtida no item 5.3.3, mediante o uso da função *fit* da instância ***RandomForestClassifier***, conforme comando abaixo:

clf_rf.*fit*(***os_features***,***os_labels***)

Em que ***os_features*** são as variáveis da base de dados de teste *oversampled* e ***os_labels*** são as variáveis de classe dessa mesma base.

Após o treinamento, armazena-se o valor atual das variáveis-classe do *dataset* de testes na variável ***backup*** e executa-se a função *predict* nas *features* (variáveis não classificatórias) dessa mesma base, a qual foi gerada no item 5.3.3 (e representando 20% da base de dados original). Também é definida uma variável *randomforestprediction* para armazenar os *outputs* da predição do modelo:

backup = ***labels_test***
randomforesrprediction = ***clf_rf***.*predict*(***features_test***)

Finalmente, após o modelo ser treinado e a predição ter sido realizada, já se tem os resultados produzidos pelo *Random Forest*. Com tais resultados, é possível gerar a matriz de confusão do modelo e obter números como falsos e verdadeiros positivos e negativos, além de calcular outras métricas com tais valores.

No capítulo 7, a seguir, são comparados os resultados gerados pela predição dos dois métodos de classificação.

7. SUMÁRIO DOS RESULTADOS E COMPARAÇÃO FINAL DAS TÉCNICAS

De forma a apresentar os resultados dos dois métodos de classificação testados, a priori, foi obtida a matriz de confusão de ambos para que, assim, os números da matriz fossem utilizados para o cálculo de métricas e das áreas abaixo das curvas (*AUC – Area Under Curve*) dos tipos ROC e PR.

As matrizes de confusão de ambos os métodos podem ser obtidas pelos comandos a seguir em Python:

confusion_matrix(***backup***,***randomforestpredictions***)
confusion_matrix(***backup***,***logisticpredictions***)

Lembrando que ***backup*** representa as variáveis-classe da base de testes antes da aplicação dos modelos classificatórios e ***logisticpredictions*** e ***randomforestpredictions*** são as variáveis que guardam as predições das classificações dos modelos de Regressão Logística e *Random Forest*, respectivamente. Assim, comparando tais variáveis de predição com as classes originais da variável ***backup***, o comando *confusion_matrix* consegue gerar a matriz de confusão, a qual contém os números de falsos e verdadeiros positivos e negativos. Em outras palavras, tais números dizem respeito ao que cada modelo conseguiu errar (“falso”) ou acertar (“verdadeiro”) na classificação das transações como fraude (“positivo”) e não-fraude (“negativo”).

Com os comandos executados, pode-se gerar as matrizes de confusão, para cada método, conforme abaixo:

Tabela 12 - Resultado da matriz de confusão da Regressão Logística

Método: Regressão Logística		**Observado**	
		Fraude	Não Fraude
Classificação do Modelo	Fraude	VP = 55858	FP = 1003
	Não Fraude	FN = 11	VN = 90

Fonte: elaboração do autor

Tabela 13 - Resultado da matriz de confusão da Random Forest

Método: *Random Forest*		**Observado**	
		Fraude	Não Fraude
Classificação do Modelo	Fraude	VP = 56846	FP = 15
	Não Fraude	FN = 17	VN = 84

Fonte: elaboração do autor

O método *Random Forest* gerou uma quantidade de falsos negativos (fraudes reais classificadas erroneamente como não-fraude) ligeiramente maior que o método de Regressão Logística, enquanto este último gerou um número de falsos positivos (não-fraudes reais classificadas erroneamente como fraude) muito maior que o gerado pela técnica *Random Forest*.

Assim, com os números das matrizes de decisão, é possível calcular indicadores para cada um dos métodos, conforme tabela abaixo:

Tabela 14 - Resultados das métricas de performance dos métodos Regressão Logística e Random Forest

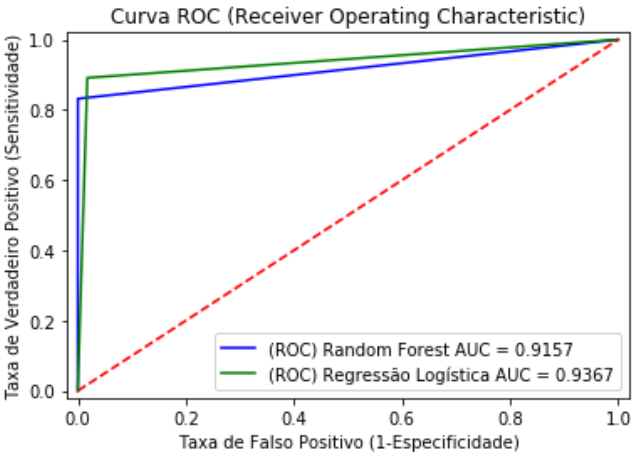
Métrica	**Métodos**		**Descrição**
	Regressão Logística	***Random Forest***	
Acurácia	98.2199%	99.9438%	Taxa de previsões corretas sobre o total de previsões feitas
Taxa de Erro	1.7801%	0.0562%	Taxa de previsões incorretas sobre o total de previsões feitas
Sensitivity = *Recall*	99.9803%	99.9701%	Fração de positivos corretamente classificados (em relação ao número total de positivos observados, isto é, a soma dos falsos negativos com verdadeiros positivos)
Specificity	8.2342%	84.8485%	Fração de negativos corretamente classificados (em relação ao número total de negativos observados, isto é, a soma dos falsos positivos com verdadeiros negativos)
Precisão	98.2360%	99.9736%	Fração de positivos corretamente classificados (em relação ao total de previsões positivas feitas, verdadeiras ou falsas)
F-*Measure*	99.1005%	99.9719%	Média harmônica entre o *recall* e a precisão

Fonte: elaboração do autor

Pode-se perceber que, em todas as métricas, com exceção da sensitividade, o método *Random Forest* foi o que performou melhor. Uma grande disparidade pode ser percebida na métrica de Especificidade, resultando numa melhor classificação de negativos (transações não fraudulentas) pelo *Random Forest*.

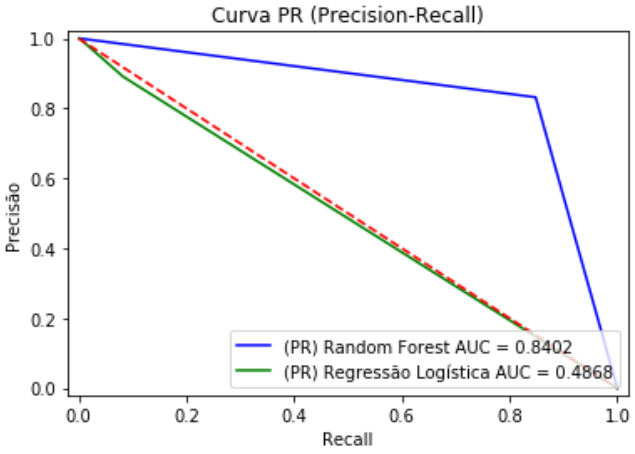
Em seguida, mediante a importação dos módulos *roc_curve*, *precision_recall_curve* e *auc* da biblioteca *sklearn.metrics*, foi possível a construção das curvas ROC e PR para os dois métodos testados. Para a geração das figuras das curvas, foi utilizada a biblioteca *matplotlib.pyplot*. As curvas e as respectivas áreas embaixo delas (AUC), de ambos os métodos, se encontram nas figuras abaixo:

Figura 60 - Resultados da curva ROC e AUC-ROC para os métodos Regressão Logística e *Random Forest*



Fonte: elaboração do autor

Figura 61 - Resultados da curva PR e AUC-PR para os métodos Regressão Logística e *Random Forest*



Fonte: elaboração do autor

A área sob a curva ROC pode ser utilizada para mensurar a capacidade de um modelo de classificação. Estando a métrica AUC entre 90% e 100%, um modelo de classificação é considerado como sendo de excelente capacidade de predição. Portanto, diz-se que ambos os modelos classificatórios, Regressão Logística e *Random Forest*, são excelentes para a natureza da base de dados em que foram testados. No entanto, o modelo de Regressão Logística obteve um resultado de AUC ligeiramente maior que o *Random Forest*. Isto significa que ele distingue melhor as classes negativas da positivas.

Ter um elevado AUC para a curva ROC é importante, pois garante que o modelo seja excelente em distinção de classes. No entanto, não é suficiente para garantir que o modelo possua alta capacidade de identificar classes positivas. Em uma base de dados desbalanceada, como as de fraude, o número de classes negativas (não-fraude) é predominante, então um modelo pode ter um AUC-ROC alto justamente por conseguir realizar uma boa predição das classes negativas. Sendo assim, é por isso que se faz necessário o cálculo do AUC para a curva PR, já que a Precisão e o *Recall*, componentes dos eixos do gráfico, não contêm verdadeiros negativos em seus cálculos. Assim, o resultado do AUC-PR diz respeito ao quão capaz um método é de realizar a predição correta de classes positivas (transações fraudulentas).

Por conseguinte, tem-se que o método *Random Forest* é 72.6% mais capaz que o método Regressão Logística de identificar classes positivas. Assim, para um cenário de transações de cartão de crédito, em que a base de dados é altamente desbalanceada (com mais classes positivas do que negativas), o *Random Forest* é o mais adequado.

8. CONCLUSÃO

De modo a encerrar o presente trabalho de formatura, procura-se tanto comparar as duas técnicas de *machine learning* quanto avaliar o aparato de análise proposto na metodologia e sua aplicação, sobretudo citando as características da base transacional e da divisão dos dados em *datasets* de treinamento e de testes.

8.1. Análise dos resultados

Este trabalho buscou realizar a comparação entre duas metodologias de aprendizado de máquina supervisionado de diferentes naturezas – a regressão logística, baseada em cálculos de probabilidade condicional, e a *Random Forest*, baseada em um *ensenble learning* de várias árvores de decisão – para aplicação em um modelo de detecção de transações de cartão de crédito. Como base de treinamento e testes para o modelo, foi utilizado um *dataset* utilizado em pesquisa colaborativa realizada pela Worldline (processadora de pagamentos por cartão de crédito) e a ULB (*Université Libre de Bruxelles*) com foco em detecção de fraude e *big data mining*. A base contém 284807 transações, sendo 492 fraudulentas, efetuadas por portadores de cartões de crédito europeus, em dois dias de setembro de 2013, captadas pelos sistemas de processamento da Worldline.

Após a divisão dos dados em 80% para treinamento dos métodos e 20% para testes, dada a natureza desbalanceada da base, isto é, que apresenta classes negativas (transações não fraudulentas) em sua grande maioria, foi necessário realizar um processo de *oversampling* nos dados de teste. Para o tratamento dos dados, utilizou-se a biblioteca *sklearn* em ambiente *Jupyter*, no qual se utiliza a linguagem de programação *Python*.

Com a aplicação das duas técnicas, percebe-se que a regressão logística apresenta as métricas de sensitividade e AUC-ROC superiores às da técnica *Random Forest*. Contudo, somente isso não garante uma capacidade elevada de predição de classes positivas.

Ambas as técnicas apresentam métricas aceitáveis – sobretudo em relação a acurácia – no entanto, somente a área sob a curva PR (AUC-PR) é o indicador que mede o quanto um método é capaz de predizer classes positivas corretamente. Tendo isso em consideração, percebe-se que o método *Random Forest* é o mais capaz de realizar a predição de classes positivas – a das transações fraudulentas, que são a minoria. Não somente isso, como também é capaz de predizer com maior corretude

as classes negativas – a das transações não-fraudulentas, que são a maioria – quando comparado ao método de regressão logística, uma vez que sua especificidade é mais de dez vezes maior que a deste último método.

Conclui-se, por conseguinte, que em bases de dados de transações de cartão de crédito realizadas em ambiente *online* – que apresentam natureza desbalanceada – ambos os métodos de aprendizado de máquina estudados neste trabalho são recomendados para aplicação em detecção de fraude. No entanto, em outros tipos de base, isto é, as que apresentam outras categorias de transações de cartão de crédito (por exemplo, as de Cartão Presente em máquinas de saque “ATM”), ou as que não apresentam transações de cartão de crédito, ou até em bases de natureza não balanceada, a assertividade de tais métodos não pode ser assegurada e, inclusive, seu desempenho pode ser consideravelmente menor que o exibido neste trabalho. Mas, caso a utilização seja restrita a bases transacionais de cartão de crédito em *e-commerce*, o desempenho de tais métodos é alto – com uma acurácia superior a 90% – sendo a técnica de *Random Forest* a mais adequada nesses casos.

8.2. Análise da metodologia

Em relação à metodologia adotada para o presente trabalho, primeiramente, retomou-se a definição do problema – altas perdas financeiras com transações de cartão de crédito no cenário de compras *online* – e, em seguida, tal problema foi estratificado em outros dois: perdas financeiras com fraude em transações *online* (i.e. falsos negativos, ou transações fraudulentas erroneamente aprovadas) e perdas financeiras com transações *online* recusadas (i.e. falsos positivos, ou transações legítimas erroneamente recusadas por suspeita de fraude).

Os dados obtidos na empresa de estágio do autor possibilitaram mostrar que a maior parte (aproximadamente 80%) do volume financeiro de transações de cartão de crédito está associada a transações presenciais (não *online*). Mesmo representando a maior parte do volume financeiro, as transações de Cartão Não Presente (*online*) apresentam uma taxa de 36% de recusas, contra 4% de taxa de recusa das transações de Cartão Presente (não *online*). Ou seja, em ambiente de transações *online*, recusa-se 36% das transações, considerada uma porcentagem alta. Ademais, ainda no ambiente de transações *online*, foi possível obter, também com dados fornecidos pela organização de estágio do autor, que este cenário concentra 83% do volume financeiro de fraude, contra 17% no canal de Cartão

Presente (transações não *online*). Assim, somente com essas porcentagens, é possível comprovar que, de fato, o problema de altas perdas financeiras em transações *online* é estratificado em outros dois, conforme descrito no parágrafo anterior: fraude (falsos negativos) e recusa de falsos positivos.

Em seguida, com a técnica dos “5 Porquês”, foi possível identificar a causa raíz para ambos os estratos do problema. No entanto, por questões de *compliance*, não foi possível obter, com a organização de estágio, as métricas de desempenho dos métodos de detecção de fraudes *online* utilizados pelos bancos emissores de cartão de crédito. Tais dados seriam de extrema valia para serem utilizados em comparações com as duas metodologias de aprendizado de máquina exploradas no presente trabalho.

Além do mais, o universo das transações CNP (*online*) de cartão de crédito é muito amplo, e contém outras subcategorias, como a CNP *E-Commerce*, CNP Recorrente, CNP MO/TO (*Mobile Order / Telephone Order*) e CNP-Outros. Assim, foi necessário obter mais dados, com a empresa de estágio do autor, com a finalidade de traçar um gráfico de Pareto e identificar qual é a subcategoria transacional com a maior participação nas transações declinadas (recusadas). Percebeu-se que a subcategoria mais expressiva em transações negadas é a CNP *E-Commerce*. Foi possível, também, utilizando o gráfico de Pareto, destacar quais as principais razões de recusa das transações declinadas pertencentes ao subcanal CNP *E-Commerce*. Dessas razões, as principais encontradas foram:

- “*Do Not Honor*” (transação não honrada pelo portador – **motivo comumente associado a suspeita de fraude**);

- “*Restricted Card*” (cartão restrito, isto é, limitado para realizar certas transações) e, por fim;

- “*Insufficient funds / over credit limit*" (fundos insuficientes / limite do cartão atingido).

A razão “*Do Not Honor*” se encontra em primeiro lugar das razões principais de recusa, sendo associada a suspeitas de fraude. Então, foi possível confirmar que

a maioria das transações recusadas no subcanal CNP *E-Commerce* foi devida a suspeitas de fraude.

Ainda com a utilização do gráfico de Pareto, identificaram-se, também, as subcategorias de transações CNP (*online*) com maior volume financeiro de fraude. Sem muitas surpresas, a subcategoria CNP *E-Commerce* se encontra em primeiro lugar.

Assim, com base nos dois parágrafos anteriores, fica evidente que as transações CNP *E-Commerce* são as que apresentam maior número de recusas por suspeita de fraude e maior volume de fraude (decorrente de suas transações erroneamente aceitas, i.e. falsos negativos), sendo esta categoria o foco do trabalho.

Ao se tratar das metodologias acima utilizadas, como a análise de proporção de transações, recusas e volume financeiro de fraude, assim como a técnica dos “5 Porquês” e os gráficos de Pareto, pode-se dizer que todas elas foram capazes de realizar o que foi objetivado, sem dificuldades. Tais métodologias são práticas, fáceis de serem usadas e fornecem muitas informações valiosas com pouco esforço em sua aplicação. Um problema enfrentado, entretanto, foi a definição dos períodos para extração dos dados. Em consulta a colegas de trabalho, à gerência da área de autorizações, da área de fraude e a materiais da organização de estágio, foi possível compreender que a performance, em termos de transações fraudulentas e de transações recusadas, se manteve similar nos últimos dois anos para a categoria CNP (*online*), sobretudo para a subcategoria CNP *E-Commerce*. Assim, utilizou-se os anos de 2017 e 2018.

Ademais, ao se tratar das técnicas de aprendizado de máquina utilizadas neste trabalho, ambas apresentam uma enorme praticidade de aplicação. A importação de bibliotecas públicas de plotagem, aprendizado de máquina e *data science*, criadas pela comunidade do Python, facilita a aplicação dos métodos de aprendizado de máquina em uma base de dados, ainda mais ao considerar o *dataset* obtido, o qual já estava sanitizado e pronto para ser utilizado (treinado e testado) por algoritmos de *machine learning*, com todas as suas variáveis já normalizadas.

Os métodos se comportaram conforme o esperado, e a aplicação de ambos foi extremamente similar – apenas se limitando à importação das bibliotecas e aplicação dos algoritmos como funções dessas bibliotecas.

Conforme já abordado no trabalho, no item 5.3, de acordo com a Cybersource (2016), uma análise por meio de regressão, como a regressão logística, é uma técnica

estatística popular de longa data que mede a força das relações de causa-efeito em *datasets* dados estruturados. A análise de regressão tende a se tornar mais sofisticada quando aplicada à detecção de fraudes devido ao número de variáveis e ao tamanho dos conjuntos de dados. Ela é uma técnica que agrega valor ao avaliar o poder preditivo de variáveis individuais ou combinações de variáveis.

Já as árvores de decisão, ainda de acordo com a Cybersource (2016), constituem uma família de algoritmos de aprendizado de máquina usada para automatizar a criação de regras para tarefas de classificação. Os algoritmos da Árvore de Decisão podem ser usados para problemas de modelagem preditiva de classificação ou regressão. As árvores de decisão são essencialmente um conjunto de regras que são treinadas usando exemplos de fraude que os clientes estão enfrentando. A criação de uma árvore ignora recursos irrelevantes e não requer uma ampla normalização dos dados. Uma árvore pode ser inspecionada e, assim, pode-se entender por que uma decisão foi tomada seguindo a lista de regras acionadas. A técnica Random Forest usa uma combinação de várias árvores de decisão para melhorar o desempenho da classificação. Isso permite suavizar o erro que pode existir em uma única árvore. Portanto, essa técnica aumenta o desempenho geral e a precisão do modelo, ao mesmo tempo em que mantém a capacidade de interpretar os resultados. De acordo com a Altexsoft (2017), além de sua simplicidade e velocidade, as *random forests* podem ser usadas com diferentes tipos de dados, incluindo números de cartão de crédito, datas, endereços IP, códigos postais etc.

Uma dificuldade presente no uso de ambas as técnicas é entender o que está ocorrendo no momento de sua execução. Algoritmos de computação simplesmente executam passos, sem informar o usuário o que cada um desses passos está, de fato, realizando. Para um usuário novo em aprendizado de máquina mas experiente em criação de regras de fraude, seria interessante compreender como o algoritmo funciona e o comportamento de cada um de seus passos com relação à base utilizada para estudos. Assim, uma opção seria a execução dos códigos em modo *verbose*, de forma a se obter mais informações acerca de cada passo do algoritmo, assim como de todos seus *outputs* de forma detalhada. O método *Random Forest* seria o mais interessante de ser visualizado detalhadamente, uma vez que o mesmo produz árvores de decisão e seria de grande valia, para um analista de dados e criador de regras, analisar os componentes de cada árvore e como o algoritmo está percorrendo cada nó.

Outro ponto a se destacar seria compreender melhor a razão pela qual o método de *Random Forest* conseguir identificar muito mais transações Falso-Positivo quando comparado ao método de regressão logística – cerca de 65.87 vezes maior. Os *outputs* das execuções dos algoritmos não fornecem detalhes acerca dessa divergência, então um estudo mais detalhado seria necessário.

Por fim, não menos importante, apesar dos métodos atuais serem amplamente utilizados para análise de bases com transações de cartão de crédito ou de natureza similar, conforme exibido no capítulo de Revisão Bibliográfica, seria pertinente realizar testes com outros algoritmos classificatórios. A junção de técnicas em um só método seria uma alternativa para incrementar a assertividade da análise transacional, uma vez que alguns métodos podem acabar suprindo carências de outros métodos (como os estudados nesse trabalho) ou ter melhor desempenho em outras situações.

8.3. Melhorias futuras

Um modelo de aprendizado de máquina deve sempre ser alimentado com novas transações de modo a aperfeiçoar seu treinamento. Assim, é importante carregá-lo com novas bases ou novas transações de uma mesma base à medida em que forem obtidas.

Ademais, é fundamental agregar novas variáveis (*features*) à base de dados do modelo. Principalmente ao se tratar de transações em *e-commerce*, existem muitas variáveis comportamentais do usuário que são possíveis de serem obtidas e, por conseguinte, usadas ao se extrair uma base de dados para aprendizado de máquina. Algumas delas são:

a) Tempo de permanência no *website*;
b) Tempo de permanência apenas na página de finalização da compra;
c) Velocidade de digitação (ou média de tempo por caractere);
d) Versões do navegador, java, sistema operacional, modelo de dispositivo etc;
e) Em caso de utilização de dispositivos móveis;
f) Ângulo de inclinação do dispositivo durante uso;
g) Dados de impressão digital do(s) polegar(es);
h) Dados de face;
i) Dados de voz;

j) Tempo médio em *stand-by* do aparelho;
k) Dados de redes sociais, caso autorizado pelo portador legítimo do cartão.

Com variáveis comportamentais, que dizem respeito a como o usuário se comporta e interage com o seu dispositivo ou com o *website* do comércio, o método de classificação pode ser treinado a ponto de identificar quaisquer mudanças de comportamento, as quais, possivelmente, podem indicar que outra pessoa está fazendo uma transação, e não o portador legítimo. Assim, qualquer comportamento fora do padrão em uma transação aumentaria a probabilidade de o modelo classificar tal transação como sendo de classe positiva (fraude).

REFERÊNCIAS BIBLIOGRÁFICAS

ABECS Mercado de Meios de Pagamento: Guia Prático. 20 de Setembro de 2018. Disponível em: <https://www.abecs.org.br/pdf/Cartilha-de-Meio-de-Pagamentos.pdf>. Acesso em: 13 de Maio de 2019.

ALEXSOFT Fraud Detection: How Machine Learning Systems Help Reveal Scams in Fintech, Healthcare, and eCommerce. 2017.

ARAUJO, R. M. Aprendizado de máquina em sistemas complexos multiagentes: estudo de caso em um ambiente sob racionalidade limitada. 2004.

BERALDI, F. Atualização dinâmica de modelo de regressão logística binária para detecção de fraudes em transações eletrônicas com cartão de débito. 2014.

BOLTON, R. J.; HAND, D. J. Statistical fraud detection: A review. *In Statistical Science*, 17:235-255. 2002.

CALDEIRA, E.; PEREIRA, A.; BRANDAO, G. Fraud Analysis and Prevention in e-Commerce Transactions. 2014.

CARVALHO, H. M. Aprendizado de Máquina voltado para Mineração de Dados: Árvores de Decisão. 2014.

CARVALHO, M. M.; PALADINI, E. P. Gestão da Qualidade. Teorias e Casos. 2ed, Rio de Janeiro, ABEPRO. 2012.

CHOI D.; LEE K. Machine learning based approach to financial fraud detection process in mobile payment system. CIST, Korea University. 2017.

CLEARSALE Mapa da Fraude 2018. 24 de Abril de 2018. Disponível em: <https://br.clear.sale/blog/post/mapa-da-fraude-2018-tenha-acesso-ao-estudo-da-fraude-mais-completo-mercado>. Acesso em: 30 de Abril de 2019.

CONCILIADORA A história do cartão de crédito. 31 de Março de 2015. Disponível em: <https://www.conciliadora.com.br/blog/a-historia-do-cartao-de-credito>. Acesso em: 22 de Abril de 2019.

CYBERSOURCE The Role of Machine Learning in Fraud Management. 2016.
DA SILVA, F. C. Análise ROC. 2006.

DAL POZZOLO, A. et al. Adaptive Machine learning for credit card fraud detection. ULB MLG PhD thesis. 2015.

DAL POZZOLO, A. et al. Calibrating Probability with Undersampling for Unbalanced Classification. In Symposium on Computational Intelligence and Data Mining (CIDM), IEEE. 2015.

DAL POZZOLO, A. et al. Credit card fraud detection: a realistic modeling and a novel learning strategy, IEEE transactions on neural networks and learning systems, 29, 8, 3784-3797, IEEE. 2018.

DAL POZZOLO, A. et al. Learned lessons in credit card fraud detection from a practitioner perspective, Expert systems with applications, 4915-4928, Pergamon. 2014.

DAVIS, J.; GOADRICH, M. The Relationship Between Precision-Recall and ROC Curves. 2006.

DIETTERICH, T. Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms. Neural Computation, 1895–1924. 1997.

CARCILLO, F. et al. Combining Unsupervised and Supervised Learning in Credit Card Fraud Detection Information Sciences. 2019.

CARCILLO, F. et al. Scarff: a scalable framework for streaming credit card fraud detection with Spark, Information fusion, 41, 182-194, Elsevier. 2018.

CARCILLO, F. et al. Streaming active learning strategies for real-life credit card fraud detection: assessment and visualization, International Journal of Data Science and Analytics, 5, 4, 285-300, Springer International Publishing. 2018.

FERNANDES, L.; GOMES, J. Relatório de pesquisa nas Ciências Sociais: Características e modalidades de investigação. ConTexto, Porto Alegre, 3, 4, 2003.

FUENTES, A. Hands-On Predictive Analytics with Python: Master the complete predictive analytics process, from problem definition to model deployment. Packt Publishing, 129-130, 2018.

HOSSIN, M.; SULAIMAN, M. N. A review on evaluation metrics for data classification evaluations. 2015.

JOHNSTON, R.; CHAMBERS, S.; SLACK, N. Administração da Produção. 2ed. São Paulo, Atlas, 2002.

KAUR, P.; Gosain, A. Comparing the Behavior of Oversampling and Undersampling Approach of Class Imbalance Learning by Combining Class Imbalance Problem with Noise. 2018.

LAU, Marcelo. Análise das fraudes aplicadas sobre o ambiente Internet Banking. Dissertação apresentada à Escola Politécnica de São Paulo, 2006.

LEBICHOT, B. et al. Deep-Learning Domain Adaptation Techniques for Credit Cards Fraud Detection, INNSBDDL 2019: Recent Advances in Big Data and Deep Learning, 78-88. 2019.

LEE, C. S. Logistic Regression versus Decision Trees. BigML, 28 de Setembro de 2016. Disponível em: <https://blog.bigml.com/2016/09/28/logistic-regression-versus-decision-trees>. Acesso em: 11 de Junho de 2019.

MASTERCARD Mastercard Processing Rules. 25 de Junho de 2019. Disponível em: <https://www.mastercard.us/content/dam/mccom/global/documents/transaction-processing-rules.pdf>. Acesso em: 6 de Maio de 2019.

MASTERCARD Mastercard Rules. 25 de Junho de 2019. Disponível em: <https://www.mastercard.us/content/dam/mccom/global/documents/mastercard-rules.pdf>. Acesso em: 6 de Maio de 2019.

MCCUE, C. Data Mining and Predictive Analysis - Intelligence Gathering and Crime Analysis. Elsevier. 2007.

MONTAGUE, D. A. Fraud prevention techniques for credit card fraud. 2004.
NIU, X.; WANG L.; YANG X. A comparison study of Credit Card fraud detection: Supervised versus Unsupervised. 2019.

OLIVEIRA, P. Detecção de fraudes em cartões: um classificador baseado em regras de associação e regressão logística. 2016.

OSHIRO, T. M. Uma abordagem para a construção de uma única árvore a partir de uma Random Forest para classificação de bases de expressão gênica. 2013.

PERLICH, C. What are the advantages of logistic regression over decision trees? Forbes, 19 de Junho de 2017. Disponível em: <https://www.forbes.com/sites/quora/2017/06/19/what-are-the-advantages-of-logistic-regression-over-decision-trees/#48b2d7c72c35>. Acesso em: 10 de Junho de 2019.

RESENDE, S. Sistemas Inteligentes: Fundamentos e Aplicações. Manolé, 2003.

ROCCA, B. Handling imbalanced datasets in machine learning. Towards Data Science, 27 de Janeiro de 2019. Disponível em: < https://towardsdatascience.com/handling-imbalanced-datasets-in-machine-learning-7a0e84220f28>. Acesso em: 24 de Junho de 2019.

RUSSEL, S. J.; NORVIG, P. Artificial Intelligence: A Modern Approach. Prentice-Hall, 2002.

SCIKIT-LEARN DEVELOPERS Scikit-Learn User Guide. 24 de Maio de 2019. Disponível em: <https://scikit-learn.org/stable/_downloads/scikit-learn-docs.pdf>. Acesso em: 3 de Julho de 2019.

STOLFO et al. Credit card fraud detection using meta-learning: Issuer and initial results. 1997.

TOTVS E-Commerce no Brasil: entenda o crescimento do segmento nos últimos anos. 14 de Fevereiro de 2019. Disponível em: <https://www.totvs.com/blog/e-commerce-no-brasil>. Acesso em: 29 de Abril de 2019.

VERIKAS, A. et al. Electromyographic patterns during golf swing: activation sequence profiling and prediction of shot effectiveness. 2016.

WERKEMA, M. C. C. As Ferramentas da Qualidade no Gerenciamento de Processos. Belo Horizonte: Editora de Desenvolvimento Gerencial, 1995.

ZHU N. Z. W.; WANG N. Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with pratical SAS implementations. 2010.

APÊNDICE A – ESTATÍSTICAS DAS VARIÁVEIS DO *DATASET*

	Time	V1	V2	V3	V4
count	28481.000000	28481.000000	28481.000000	28481.000000	28481.000000
mean	94705.035216	-0.001143	-0.018290	0.000795	0.000350
std	47584.727034	1.994661	1.709050	1.522313	1.420003
min	0.000000	-40.470142	-63.344698	-31.813586	-5.266509
25%	53924.000000	-0.908809	-0.610322	-0.892884	-0.847370
50%	84551.000000	0.031139	0.051775	0.178943	-0.017692
75%	139392.000000	1.320048	0.792685	1.035197	0.737312
max	172784.000000	2.411499	17.418649	4.069865	16.715537

	V5	V6	V7	V8	V9
count	28481.000000	28481.000000	28481.000000	28481.000000	28481.000000
mean	-0.015666	0.003634	-0.008523	-0.003040	0.014536
std	1.395552	1.334985	1.237249	1.204102	1.098006
min	-42.147898	-19.996349	-22.291962	-33.785407	-8.739670
25%	-0.703986	-0.765807	-0.562033	-0.208445	-0.632488
50%	-0.068037	-0.269071	0.028378	0.024696	-0.037100
75%	0.603574	0.398839	0.559428	0.326057	0.621093
max	28.762671	22.529298	36.677268	19.587773	8.141560

	...	V21	V22	V23	V24
count	...	28481.000000	28481.000000	28481.000000	28481.000000
mean	...	0.004740	0.006719	-0.000494	-0.002626
std	...	0.744743	0.728209	0.645945	0.603968
min	...	-16.640785	-10.933144	-30.269720	-2.752263
25%	...	-0.224842	-0.535877	-0.163047	-0.360582
50%	...	-0.029075	0.014337	-0.012678	0.038383
75%	...	0.189068	0.533936	0.148065	0.434851
max	...	22.588989	6.090514	15.626067	3.944520

	V25	V26	V27	V28	Amount
count	28481.000000	28481.000000	28481.000000	28481.000000	28481.000000
mean	-0.000917	0.004762	-0.001689	-0.004154	89.957884
std	0.520679	0.488171	0.418304	0.321646	270.894630
min	-7.025783	-2.534330	-8.260909	-9.617915	0.000000
25%	-0.319611	-0.328476	-0.071712	-0.053379	5.980000
50%	0.015231	-0.049750	0.000914	0.010753	22.350000
75%	0.351466	0.253580	0.090329	0.076267	78.930000
max	5.541598	3.118588	11.135740	15.373170	19656.530000

	Class
count	28481.000000
mean	0.001720
std	0.041443
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	1.000000

[8 rows x 31 columns]

APÊNDICE B – HISTOGRAMA DAS VARIÁVEIS DO *DATASET*