

UNIVERSIDADE DE ARARAQUARA
MESTRADO PROFISSIONAL EM ENGENHARIA DE PRODUÇÃO

Elias Shohei Kamimura

**MODELO DE *CREDIT SCORING* PONDERADO POR DECISÕES DE
INVESTIMENTO E FINANCIAMENTO: ESTUDO DE CASO EM UMA
LOJA DE MATERIAIS DE CONSTRUÇÃO**

Dissertação apresentada ao Programa de Mestrado Profissional em Engenharia de Produção da Universidade de Araraquara – UNIARA – como parte dos requisitos para obtenção do título de Mestre em Engenharia de Produção, Área de Concentração: Gestão Estratégica e Operacional da Produção.

Prof. Dr. Anderson Rogério Faia Pinto
Orientador

Araraquara, SP – Brasil
2021

FICHA CATALOGRÁFICA

K62m Kamimura, Shohei Elias

Modelo de Credit Scoring ponderado por decisões de investimento e financiamento: estudo de caso em uma loja de materiais de construção/
Shohei Elias Kimimura. – Araraquara: Universidade de Araraquara, 2021.
65f.

Dissertação (Mestrado) - Mestrado Profissional em Engenharia de
Produção – Universidade de Araraquara - UNIARA

Orientador: Prof. Dr. Anderson Rogério Faia Pinto

1. Naïv Bayes. 2. Credit Scoring. 3. Risco de crédito. 4. Micro e pequenas
Empresas. I. Título.

CDU 62-1

REFERÊNCIA BIBLIOGRÁFICA

KAMIMURA, E. S. **Modelo de *Credit Scoring* ponderado por decisões de investimento e financiamento: Estudo de caso em uma loja de materiais de construção**. 2021. Número de folhas 65 f. Dissertação de Mestrado em Engenharia de Produção – Universidade de Araraquara, Araraquara-SP.

ATESTADO DE AUTORIA E CESSÃO DE DIREITOS

NOME DO AUTOR: Elias Shohei Kamimura

TÍTULO DO TRABALHO: Modelo de *Credit Scoring* ponderado por decisões de investimento e financiamento: Estudo de caso em uma loja de materiais de construção.

TIPO DO TRABALHO/ANO: Dissertação / 2021

Conforme LEI Nº 9.610, DE 19 DE FEVEREIRO DE 1998, o autor declara ser integralmente responsável pelo conteúdo desta dissertação e concede a Universidade de Araraquara permissão para reproduzi-la, bem como emprestá-la ou ainda vender cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta dissertação pode ser reproduzida sem a sua autorização.



Elias Shohei Kamimura

Universidade de Araraquara – UNIARA

Rua Carlos Gomes, 1217, Centro. CEP: 14801–340, Araraquara-SP

Email (do autor): eliasshohei@gmail.com



UNIVERSIDADE DE ARARAQUARA - UNIARA
MESTRADO PROFISSIONAL EM ENGENHARIA DE PRODUÇÃO

FOLHA DE APROVAÇÃO

Dissertação apresentada ao Programa de Mestrado Profissional em Engenharia de Produção da Universidade de Araraquara – UNIARA – para obtenção do título de Mestre em Engenharia de Produção.

Área de Concentração: Gestão Estratégica e Operacional da Produção.


NOME DO AUTOR: ELIAS SHOHEI KAMIMURA

TÍTULO DO TRABALHO:

"MODELO DE CREDIT SCORING PONDERADO POR DECISÕES DE INVESTIMENTO E FINANCIAMENTO: ESTUDO DE CASO EM UMA LOJA DE MATERIAIS DE CONSTRUÇÃO."

Assinatura do(a) Examinador(a)

Conceito


Prof(a). Dr(a). Anderson Rogério Faia Pinto (orientador(a))
Universidade de Araraquara - UNIARA

(X)Aprovado () Reprovado


Prof(a). Dr(a). Bruna Cristine Scarduelli Pacheco
Universidade de Araraquara - UNIARA

(X)Aprovado () Reprovado


Prof(a). Dr(a). Marcelo Seido Nagano
Universidade de São Paulo - USP

(X)Aprovado () Reprovado

Versão definitiva revisada pelo(a) orientador(a) em: 03 / 11 / 2021


Prof(a). Dr(a). Anderson Rogério Faia Pinto (orientador(a))

AGRADECIMENTOS

A Deus primeiramente pela oportunidade.

Ao empresário e empreendedor João Carlos Barbosa pelo apoio.

Ao Prof. Dr. Anderson Rogério Faia Pinto pela orientação e dedicação.

Existe o risco que você não pode jamais correr,
e existe o risco que você não pode deixar de
correr.

Peter Drucker

RESUMO

A realidade econômica é caracterizada por uma acirrada concorrência entre Micro e Pequenas Empresas (MPEs) que representam 99% de todas as empresas do Brasil. É neste cenário que as vendas por meio da concessão de crédito são uma importante estratégia para a geração de valor. A venda a prazo é comum no mercado, facilita a realização de negócios e é fundamental para as transações financeiras. Em muitos casos, as políticas de crédito também estão voltadas à obtenção de recursos para uma decisão de investimento cujo objetivo é a sobrevivência ou cobrir desenhos de caixa. A literatura traz diferentes modelos para análise do risco de crédito que incluem desde as abordagens clássicas (métodos estatísticos) até as mais sofisticadas (inteligência computacional) e que são comumente conhecidos como Modelos de *Credit Scoring* (MCSs). Todavia, os MCSs propostos ainda estão longe do alcance da maioria das MPEs do Brasil. Isso ocorre seja pelo desconhecimento ou incapacidade técnica dos microempresários ou ainda pelos altos custos e complexidade dos referidos MCSs. É frente a estas lacunas que esta pesquisa tem como objetivo propor um MCS para uma MPE do ramo de comércio de materiais de construção do interior do estado de São Paulo (SP). A pesquisa é aplicada e de natureza quantitativa, e caracterizada como um estudo de caso em uma MPE. Para o cálculo da probabilidade de inadimplência e adimplência, o MCS é fundamentado no algoritmo de *Naive Bayes*. Todos os resultados são ponderados nas soluções das ferramentas da Moderna Teoria de Finanças com utilização do Custo de Conceder e do Custo de Não Conceder. Este MCS é adaptado à realidade em questão e foi desenvolvido (entradas/saídas de dados e as análises dos resultados) por meio do Microsoft Office Excel 2016. As análises dos resultados demonstraram que o MCS é capaz de ponderar o risco nas decisões de investimento e financiamento na concessão de vendas a crédito da referida MPE. No caso dos clientes com crédito aprovado, o MCS apresentou acertos de 70% e para os clientes com crédito negado os acertos foram de 66%. Experimentos utilizando as aprovações das vendas no crediário realizadas em 2021 demonstraram um resultado financeiro positivo de R\$ 32,20 mil com a aplicação do MCS. Quando avaliado o impacto da aplicação do MCS no DRE da empresa, o lucro líquido foi maior em 124,2%. As principais contribuições desta pesquisa estão em prover um MCS de baixo custo, de fácil manuseio e implementação no Microsoft Office Excel 2016. Esta pesquisa é um passo inicial na aplicação de MCSs aos processos de concessão de crédito realizados por MPE. Interessante sugestão para futuras pesquisas é o desenvolvimento de MCSs com foco nas peculiaridades das MPEs.

Keywords: *Naive Bayes*; *Credit Scoring*; Risco de Crédito; Micro e Pequenas Empresas

ABSTRACT

Economic reality features a flaring competition among Micro and Small Enterprises (MSEs), which represent 99% of all businesses in Brazil. In this scenario, the sales made by means of credit granting are an important strategy towards value generation. Sales in installment plans are very common in the market as they facilitate business and play a crucial role in financial transactions. In many cases, credit policies are also focused on obtaining resources for an investment decision whose goal is either surviving or bridging cash flow gaps. Literature tells of different models for credit risk analysis, which range from classical approaches (statistic methods) to the most sophisticated (computational intelligence), widely known as Credit Scoring Models (CSMs). However, the proposed CSMs are still far from reach for most MSEs in Brazil. This occurs due to either microentrepreneurs' unawareness or technical incapacity or the high costs and complexity of the aforementioned CSMs. Facing these gaps, this research aims at proposing a CSM for an MSE in the segment of construction materials commerce in the countryside of São Paulo (SP) state. The study applies and is of quantitative nature, being classified as a case study of an MSE. In the CSM the foundation of the compliance and noncompliance probability calculations is laid on the Naive Bayes classifier. All results are pondered on solutions provided by the Modern Finance Theory tools, making use of the Cost of Granting and the Cost of Not Granting. This CSM is adapted to the reality at issue and was developed (data inputs/outputs and analysis of results) via Microsoft Office Excel 2016. The analyses of results have shown the CSM is able to weight the risk in decisions about investment and installment plans for sales based on credit granting for the MSE spoken of above. In the case when clients had their credit approved, the CSM had 70% of hits, and for clients whose credit was denied, hits were 66%. Trials on approvals of credit sales closed in 2021 demonstrated a positive financial result of R\$32,200 when the CSM was applied. Evaluating the impact of applying the CSM on the enterprise's Statement of Operations, the net profit was higher in 124.16%. The main contributions of this study consist of providing a low cost CSM that is easy to use and implement in Microsoft Office Excel 2016. This research is a first step in the application of CSMs to the credit granting processes that SMEs conduct. An interesting suggestion for future studies is to develop CSMs rather focused on SMEs' specifications.

Keywords: *Naive Bayes; Credit Scoring; Credit Risk; Micro and Small Enterprises*

Lista de Figuras

Figura 1 – Métodos Usados pelos Modelos de <i>Credit Scoring</i>	29
Figura 2 – Classificação dos Modelos de <i>Credit Scoring</i>	30
Figura 3 – Bases de Dados Usadas pelos Modelos de <i>Credit Scoring</i>	31
Figura 4 – Abordagem das Pesquisas Seleccionadas	34
Figura 5 – Mapa de Relações das Palavras-Chave	34
Figura 6 – Ocorrência de Palavras-Chave por Ano	35
Figura 7 – Distribuição Temporal das Publicações	35
Figura 8 – Categorização das Publicações de <i>Credit Scoring</i>	36
Figura 9 – Periódicos com Abordagens de <i>Credit Scoring</i>	37
Figura 10 – Composição do Capital da Empresa	41
Figura 11 – Evolução das Vendas das no Crediário	44
Figura 12 – Volume de Vendas no Crediário por Ano	45
Figura 13 – Inadimplência das Vendas no Crediário	46
Figura 14 – Fluxo do Modelo de Risco de Crédito	46
Figura 15 – Fluxo do Modelo de Análise do Custo de Capital	47
Figura 16 – Fluxo de Elaboração do Custo de Conceder Adaptado	47
Figura 17 – Fluxo de Elaboração do Custo de Negar Adaptado	47
Figura 18 – Interface do Modelo de <i>Credit Scoring</i> Proposto	56
Figura 19 – Erros e Acertos de Pedidos com Crédito Aprovado	57
Figura 20 – Erros e Acertos de Pedidos com Crédito Reprovado	57

Lista de Quadros

Quadro 1 – Relação de Artigos e Técnicas Aplicadas.....	28
Quadro 2 – Esquemas Metodológicos Aplicados às Modelagens.....	29
Quadro 3 – Abordagens para Construção de Sistemas Híbridos.....	30
Quadro 4 – Classificação das Pesquisas de Revisão de Literatura.....	36
Quadro 5 – Periódicos com Publicações de <i>Credit Scoring</i>	38

Lista de Tabelas

Tabela 1 – Indicadores dos Periódicos com Publicações de <i>Credit Scoring</i>	39
Tabela 2 – Demonstrativos do Resultado do Exercício (DRE).....	42
Tabela 3 – Histórico da Participação do Crediário nas Vendas	44
Tabela 4 – Perfil dos Clientes que Compram no Crediário.....	45
Tabela 5 – Valores em Atraso das Vendas no Crediário.....	46
Tabela 6 – Distribuição das Vendas por Faixa de Valor	48
Tabela 7 – Classificação de Clientes da Empresa	48
Tabela 8 – Matriz de Cálculo da Probabilidade (P).....	49
Tabela 9 – <i>Escores</i> dos Clientes da Empresa	50
Tabela 10 – Cálculo do Beta Realavancado e <i>Capital Asset Pricing Model</i>	51
Tabela 11 – Cálculo do <i>Weighted Average Cost of Capital</i>	51
Tabela 12 – Distribuição de Pendências Cadastrais por Pedido.....	54
Tabela 13 – Distribuição de Pedidos por Faixa de Valor	55
Tabela 14 – Distribuição dos Prazos de Pagamentos dos Pedidos	55
Tabela 15 – Totalidade de Clientes por Faixa e Quantidade de Pedidos	55
Tabela 16 – Resultados do Modelo de <i>Credit Scoring</i> Proposto.....	58
Tabela 17 – Demonstrativo do Resultado Ajustado ao Modelo de <i>Credit Scoring</i>	58

Lista de Siglas e Abreviaturas

AUC – *Area Under the ROC Curve*
CBR – *Case-Based Reasoning*
CNDL – *Confederação Nacional de Dirigentes Lojistas*
DA – *Discriminant Analysis*
DT – *Decision Tree*
EMP – *Expected Maximum Profit*
GA – *Genetic Algorithm*
IHT – *Instance Hardness Threshold*
KDD – *Knowledge-Discovery in Databases*
k-NN – *k-Nearest Neighbors*
LLE – *Locally Linear Embedding*
LR – *Logistic Regression*
LRM – *Linear Regression*
MCS – *Modelo de Crédito Scoring*
MPE – *Micro e Pequena Empresa*
MR – *Multivariate Regression*
NB – *Naive Bayes*
NMF – *Non-negative Matrix Factorization*
NN – *Neural Network*
P2P – *Person to Person*
PCA – *Principal Component Analysis*
RF – *Random Forest*
ROC – *Receiver Operating Characteristic Curve*
SEBRAE – *Serviço Brasileiro de Apoio às MPes*
SMOTE – *Synthetic Minority Oversampling Technique*
SPC – *Serviço de Proteção ao Crédito*
SVC – *Support Vector Clustering*
SVM – *Support Vector Machine*
TIR – *Taxa Interna de Retorno*
WACC – *Weighted Average Cost of Capital*

SUMÁRIO

1 INTRODUÇÃO	12
1.1 Problema de pesquisa	14
1.2 Justificativa da Pesquisa	14
1.3 Objetivos da Pesquisa.....	15
1.3.1 Objetivo Geral	15
1.3.2 Objetivos Específicos	15
1.4 Delimitação da Pesquisa.....	16
1.5 Estrutura da Dissertação	16
2 FUNDAMENTAÇÃO TEÓRICA.....	17
2.1 Política de Crédito como Decisão de Investimento.....	17
2.2 Processo de Concessão de Crédito	18
2.3 Gestão do Risco de Crédito	19
2.4 Algoritmo <i>Naive Bayes</i>	20
2.5 <i>Credit Scoring</i>	20
3 REVISÃO E ANÁLISE DA LITERATURA	22
3.1 Revisão da Literatura.....	22
3.2 Análise da Literatura	27
4 METODOLOGIA DE PESQUISA	33
4.1 Classificação.....	33
4.2 Método de Pesquisa	33
4.4 Procedimentos Operacionais	40
5 MODELO DE <i>Credit Scoring</i> PROPOSTO	41
5.1 A Empresa	41
5.2 Cadastro de Clientes	42
5.3 Procedimento das Vendas.....	43
5.4 Modelo de <i>Credit Scoring</i>	46
5.4.1 Modelagem do Risco de Crédito	47
5.4.2 Modelagem do Custo de Capital	51
5.4.3 Formulação do Modelo de <i>Credit Scoring</i>	52
6 EXPERIMENTAÇÕES E ANÁLISE DOS RESULTADOS.....	54
6.1 Experimentações.....	54
6.2 Análise dos Resultados.....	57
7 CONSIDERAÇÕES FINAIS.....	60
REFERÊNCIAS	61

1 INTRODUÇÃO

Nas três últimas décadas, a participação das Micro e Pequenas Empresas (MPEs) ganhou cada vez mais relevância na economia do Brasil. Em 1985, estas empresas representavam 21% de participação no Produto Interno Bruto (PIB), chegando a 30% em 2017 (SEBRAE, 2020). No Brasil, há mais de 19 milhões de empresas, sendo 17 milhões de MPEs (SEBRAE, 2020). Estas MPEs apresentam um quadro no qual os seus respectivos patrimônios normalmente se confundem com o dos proprietários (SILVA, 2016). Portanto, as funções básicas de finanças (decisões de investimento, financiamento e incremento do valor para os acionistas) cabem ao empresário, ao empreendedor ou aos sócios proprietários (BREALEY; MYERS, 1996; SECURATO, 2012; SILVA, 2016). Logo, a estrutura de capital e as decisões relativas às fontes de financiamentos estão diretamente relacionadas com a política de crédito da empresa e de seus tomadores de decisão (SECURATO, 2012; SILVA, 2016). É possível afirmar que o crédito é fundamental para as transações financeiras, já que, proporciona liquidez para as atividades econômicas e para os financiamentos de operações e investimentos (DOUMPOS et al., 2018).

É neste sentido que as operações de crédito atuam como facilitadoras do negócio à medida que possibilita tanto a aquisição de um bem pelos clientes quanto incrementos de vendas para os comerciantes. A classificação dos clientes para a concessão de crédito é, portanto, uma função empresarial tão relevante quanto a realização das vendas (DE AMORIM, 2019). Em algumas atividades comerciais, a receita obtida com o financiamento das mercadorias chega a ser maior que as suas margens com a comercialização mercadorias (SECURATO, 2012; SILVA, 2016). Pesquisa publicada pela CNDL/SPC (2019) revelou que 30% dos brasileiros entrevistados recorreram ao crediário nos últimos 12 meses (26% todos os meses). Entre os principais motivos para isso estão a falta de condições para o pagamento à vista (35%) e a pouca burocracia exigida pelos comerciantes (25%), o que representou um incremento de 12% em relação a 2018 (CNDL/SPC, 2019). Em termos gerais, a concessão de crédito é uma decisão sob condições de incerteza e o risco ao credor é gerado pela probabilidade de os mutuários não cumprirem com todas as suas obrigações. Portanto, se o credor puder estimar a probabilidade da ocorrência de uma eventual perda, então, as tomadas de decisões serão mais confiáveis (ANDRIOSOPOULOS et al., 2019; SICSU, 2012).

Assim sendo, a gestão eficaz do risco é um fator decisivo para as instituições de crédito e tema de interesse para empresas não financeiras e consumidores (ANDRIOSOPOULOS et al., 2019; DOUMPOS et al., 2018; SARIANNIDIS et al., 2020). Logo, os *scores* de crédito buscam quantificar o risco e, para isso, se baseiam em características sociodemográficas para categorizar os clientes com características que permitam classificá-los como um “bom pagador”

ou “mau pagador” (LOUZADA; ARA; FERNANDES, 2016; SICSÚ, 2012). Estes *escores* podem ser simbolizados por letras, números ou rótulos específicos, e representam uma nota ou ainda uma ou mais qualidades de determinados clientes. Então, o primeiro passo nas análises de decisão de crédito é estimar qual a probabilidade de que um requerente sustente um bom risco. Feito isso, os clientes cujos escores respondem com uma alta probabilidade de recebimento são aceitos enquanto aqueles com baixa probabilidade são rejeitados (consulte Finlay (2009), Louzada, Ara e Fernandes (2016) e Sicsú (2012). Há, portanto, que se caracterizar as medidas do risco e as políticas de crédito como dois elementos interdependentes e fundamentais das Finanças Corporativas. Além disso, uma importante ferramenta muito utilizada para a medida do risco de crédito é o *Credit Scoring*. Em termos técnicos, o *Credit Scoring* tem como finalidade quantificar o risco na tomada de decisão de concessão ou não de crédito (SECURATO, 2010). A política de crédito já consiste em uma decisão de investimento, ou seja, a grande questão é como oferecer crédito sem comprometer a situação financeira da empresa (RENDE, 2013; SECURATO, 2012).

Entretanto, é importante ressaltar que os credores não se importam somente com a probabilidade de inadimplência dos clientes. Há de se considerar que o lucro proporcionado pelo empréstimo também depende das perdas com as inadimplências e das taxas de juros dos empréstimos (ANTONAKIS; SFAKIANAKIS, 2009; SERRANO-CINCA; GUTIÉRREZ-NIETO, 2016). À vista disto, empréstimos mais arriscados apresentam uma maior probabilidade de inadimplência, contudo, oferecem taxas de juros mais altas para compensar o credor pelo risco. Exemplos são os microcréditos focados em pessoas excluídas financeiramente, ou seja, podem ser arriscados, mas, dadas as suas altas taxas de juros, podem ser muito rentáveis (SERRANO-CINCA; GUTIÉRREZ-NIETO, 2016). No contexto das MPEs, as decisões que envolvem a alocação de recursos na concessão de crédito podem levar à realização de classificações equivocadas dos clientes. Essas classificações normalmente resultarão em perdas que geram a necessidade de capital de terceiros, comprometimento com fornecedores e aumento do risco de falência (DE AMORIM, 2019; POTRICH et al., 2012). Nessa lógica, há uma notória demanda por modelos de análise de crédito que considerem tanto os aspectos do risco quanto as decisões de investimento e financiamento mais rentáveis às MPEs. É fundamentado na realidade prática das MPEs e na busca por uma maior rentabilidade na realização de vendas à crédito que esta pesquisa propõe um Modelo de *Credit Scoring* (MCS). Por conseguinte, este MCS será aplicado ao processo de análise de vendas à crédito de uma MPE do ramo de comércio varejista de materiais de construção localizada no interior do estado de São Paulo (SP).

1.1 Problema de Pesquisa

Esta pesquisa aborda um problema real de análise de risco e *Credit Scoring* para as vendas à crédito de uma MPE. A empresa em estudo atua no ramo do comércio varejista de materiais de construção e está localizada em uma estância turística no interior do estado de São Paulo (SP). A gestão da empresa é centralizada no sócio administrador que toma todas as decisões, em especial, as que cuidam de investimentos e financiamentos. Este sócio é também quem realiza a aprovação ou não das vendas à crédito, porém, com base em análise julgamental e empírica, baseada em sua experiência e *feeling*. Fato é que nenhum dos sócios possui conhecimento técnico em torno do tema de análise de risco e *Credit Scoring*. Logo, não são aplicadas quaisquer técnicas para análise do risco de crédito ou MCSs. Entretanto, as vendas a crédito, se comparadas ao faturamento total da empresa, representaram 15,39% (até o mês de setembro de 2021). Este percentual corresponde a R\$ 852,75 mil, entretanto, 44% deste valor se refere à clientes que estão em situação de inadimplência, ou seja, R\$ 327,12 mil. É evidente que estas questões podem comprometer a sobrevivência da empresa, quer seja ao priorizar um negócio com maior risco ou pela não realização de uma venda rentável. Junta-se a isso os aspectos econômicos e financeiros que demandam decisões de investimento em um contexto de intensa concorrência e altas taxas de mortalidade como é a realidade do Brasil. É com base nesse contexto que surge a seguinte questão de pesquisa: Como desenvolver um MCS para uma MPE considerando o risco de inadimplência e as decisões de investimento e financiamento?

1.2 Justificativa da Pesquisa

A maioria das grandes empresas conta com completas bases de dados de clientes e complexos modelos para a análise de risco e precificação. Em geral, possuem uma estrutura de capital diversificada com fontes de financiamento (grandes bancos e financeiras) a custos baixos e consideráveis investimentos em tecnologia de ponta. No contexto das MPEs, a geração de valor está, muitas vezes, voltada para a obtenção de recursos visando a sobrevivência ou com o objetivo de cobrir desencaixes de caixa. Dessa maneira, muitas MPEs apresentam limitações financeiras e, conseqüentemente, dificuldades para a realização de investimentos (SILVA, 2016; DE AMORIM, 2019). Além disso, as MPEs estão inseridas em ambientes complexos e competitivos, em especial, no Brasil. Logo, a concessão de crédito é uma importante estratégia para a manutenção das carteiras de clientes e para os acréscimos nos níveis de faturamento, visto que, facilita a concretização das vendas (RENDE, 2013; DE AMORIM, 2019). Porém, uma série de especificidades na estrutura e nos processos de gerenciamento são observadas nas MPEs. Exemplos são as quantidades insuficientes de colaboradores e a

administração familiar que pode prejudicar à especialização de funções relevantes (DE AMORIM, 2019). Isto é, muitas MPEs apresentam dificuldades no gerenciamento de atividades, incluindo as financeiras e, obviamente, a gestão do risco de crédito (DE AMORIM, 2019; SANTOS; CARMONA; LEISMANN, 2010).

Entretanto, no que tange a literatura, a quase totalidade das pesquisas abordam somente a melhoria da gestão de crédito para as instituições financeiras (DE AMORIM, 2019; SANTOS; CARMONA; LEISMANN, 2010). É raro encontrar pesquisas que tem como objetivo propor MCSs ou estudos que abordam a gestão do risco de crédito em MPEs (SANTOS; CARMONA; LEISMANN, 2010). Em sua maioria, as pesquisas tratam extensas bases de dados com muitas variáveis, e cujo foco é a busca por uma maior precisão na previsão do risco de inadimplência (consulte Lessmann et al. (2015), Louzada, Ara e Fernandes (2016) e Andriosopoulos et al. (2019)). Há, portanto, duas questões práticas, pontos chaves desta pesquisa, ainda não abordadas pela literatura das MPEs no Brasil: i) a demanda por modelos mais aptos à realidade das MPEs e; ii) os modelos existentes são caros e/ou complexos e não contribuem para o processo de tomada de decisão nas MPEs. É frente a estas lacunas que esta pesquisa irá propor um MCS mais apto e acessível às MPEs do Brasil. Uma visão mais ampla em relação às vendas a crédito, entendida como uma função de finanças e não só com vistas a probabilidade de inadimplência, será contemplada pelo referido MCS. Este MCS será aplicado aos processos de análise do risco e de viabilidade econômica das vendas a crédito da MPE em estudo localizada no interior de São Paulo (SP). A possibilidade de fornecer uma ferramenta de gestão melhor factível à realidade desta MPE justifica a formulação e aplicação de um MCS.

1.3 Objetivos da Pesquisa

1.3.1 Objetivo geral

Propor um Modelo de *Credit Scoring* ponderado por decisões de investimento e financiamento na concessão de vendas à crédito de uma empresa comercial do ramo de materiais de construção do interior de São Paulo.

1.3.2 Objetivos específicos

- Aperfeiçoar o processo de classificação de clientes quanto ao risco;
- Prover uma ferramenta de apoio à tomada de decisão para vendas à crédito;
- Aprimorar o processo de análise de decisões de investimento e financiamento;
- Ponderar o risco de inadimplência e as decisões de investimento e financiamento.

1.4 Delimitação da Pesquisa

Esta pesquisa tem como escopo propor um MCS específico à realidade de uma pequena empresa comercial do ramo de materiais de construção localizada no interior do estado de São Paulo (SP). É evidente que há diferentes técnicas computacionais, estatísticas e modelagens aplicadas ao risco de crédito e na construção de MCSs. Isso inclui o desenvolvimento de sofisticados algoritmos que tem como o objetivo obter agilidade e precisão na mensuração dos riscos de inadimplência como, por exemplo, o *Neural Network* (NN) e o *Genetic Algorithm* (GA). Entretanto, esta pesquisa não tem a intenção de discutir, aprimorar ou realizar análises comparativas relacionadas às teorias, aplicações ou a acurácia de ferramentas propostas pela literatura dos MCSs. A escolha pelo algoritmo *Naive Bayes* (NB) deu-se pelo fato de o NB ser um método de classificação muito simples, eficaz e de fácil implementação por meio do Microsoft Office Excel 2016. A mecânica do NB pode ser facilmente compreendida e executada rapidamente, e ainda pode superar os métodos de classificação mais sofisticados (ANTONAKIS; SFAKIANAKIS, 2009; HILL; LEWICKI, 2007; JIE; BO, 2011). Por conseguinte, a premissa básica desta pesquisa está em mitigar o risco das vendas à crédito ponderando-as pela rentabilidade do negócio mediante a aplicação de ferramentas da moderna Teoria de Finanças. Para isso, busca-se o desenvolvimento de um MCS de baixo custo, de fácil utilização e que apresente soluções viáveis para as MPEs. Além disso, que considere a estrutura, o acesso à informação e os custos da empresa em estudo. Em termos gerais, esta pesquisa visa fornecer uma ferramenta de gestão que apoie o empresário a tomar decisões mais assertivas e rentáveis nas vendas à crédito, logo, propõem um MCS.

1.5 Estrutura da Dissertação

Esta dissertação é dividida em sete seções resumidas da seguinte forma; a seção 2 expõe os aspectos teóricos que fundamentam a pesquisa; a seção 3 trata da revisão e análise da literatura dos MCSs; a seção 4 apresenta a metodologia de pesquisa; a seção 5 detalha a representação e modelagem do MCS; a seção 6 exhibe os experimentos e análise dos resultados do MCS; a última seção expõe as considerações finais relativas à pesquisa e as sugestões para futuros estudos em torno dos MCSs.

2 FUNDAMENTAÇÃO TEÓRICA

Esta seção expõe o embasamento teórico relativo à política de crédito como decisão de investimento no âmbito do *Credit Scoring*. Além disso, traz uma breve explanação sobre a concessão de crédito, a gestão do risco de crédito, o algoritmo *Naive Bayes* e o *Credit Scoring*.

2.1 Política de Crédito com Decisão de Investimento

A política de crédito é uma das principais decisões de investimento das empresas (SECURATO, 2012; SILVA, 2016). Essencialmente, as operações de crédito consistem na entrega de um bem ou valor presente com a expectativa de recebimento de um montante em uma data futura (SILVA, 2016). É nestas operações que incidem uma remuneração denominada de juros e que são cobradas por um determinado período como garantia de minimizar a inadimplência (SECURATO, 2012; SILVA, 2012). Portanto, as decisões que abordam a origem e a destinação de recursos devem ser norteadas pelas características operacionais e setoriais, além do ramo de negócio e do mercado onde está inserida a empresa. É ante a este contexto que os investimentos deverão buscar a satisfação dos objetivos da empresa, dado que, o retorno sobre o capital investido consiste no indicador mais utilizado na avaliação de uma empresa (SECURTAO, 2012).

Visando maximizar o retorno das aplicações, considerado determinado nível de risco, surgiram os trabalhos de *Modigliani & Miller*, *Markowitz*, *Solomon*, *Sharp e Lintner* que deram origem a Moderna Teoria de Finanças (BERNASTEIN, 2008; CASTRO JUNIOR; FAMÁ, 2002). Logo, as decisões econômico-financeiras objetivam potencializar a riqueza dos proprietários, sendo assim, a política de crédito adotada por uma empresa é uma relevante decisão de investimento (SECURATO, 2012; SILVA, 2016). Atualmente, o modelo de precificação de ativos é denominado de *Capital Asset Pricing Model* (CAPM) e permite estimar a taxa mínima de atratividade para o capital próprio, ou melhor, a taxa exigida pelos investidores (ABREU FILHO, 2007). Dessa forma, o CAPM é obtido conforme a expressão 1.

$$CAPM = R_f + \beta \cdot (R_m - R_f) \quad (1)$$

em que R_f é o retorno sobre investimento livre de risco, R_m corresponde ao retorno médio de mercado e o coeficiente β mede a sensibilidade do ativo frente aos movimentos de mercado. No cerne desta questão, está também o expressivo papel de ferramentas como o *Weighted Average Cost of Capital* (WACC), usualmente traduzido como *Weighted Average Cost of Capital* (WACC) e obtido pela expressão 2 (ABREU FILHO, 2007).

$$WACC = K_s \cdot C_s + (1 - IR) \cdot K_d \cdot C_d \quad (2)$$

onde K_s é o custo do capital próprio, *Credit Scoring* é a participação do capital próprio no capital total (expresso em percentual), K_d é o custo de capital de terceiros e C_d é a participação do capital de terceiros no capital total (expresso em percentual). Por conseguinte, a Taxa Interna de Retorno (TIR) consiste na taxa de retorno do investimento, e pode ser definida como a taxa hipotética que anula o Valor Presente Líquido (VPL) expresso pela equação 3 (SAMANEZ, 2009).

$$VPL = -I + \sum_{t=1}^n \frac{FC_t}{(1+i)^t} = 0 \quad (3)$$

onde FC_t representa o fluxo de caixa no t -ésimo período, I é o investimento inicial, K simboliza o custo do capital. Já o Custo de Conceder Crédito (CDC) é dado pela expressão 4 (SILVA, 2016).

$$CDC = P(I).VO + CI + CB \quad (4)$$

sendo $P(I)$ a probabilidade de inadimplência ou risco de crédito, VO o valor original do empréstimo ou do financiamento (principal), CI o custo do investimento (resultado do $VO \times$ taxa títulos públicos) e CB o custo de cobrança e outras despesas. Em concomitância, o Custo de Negar (CDN) é mensurado pela expressão 5 (SILVA, 2016).

$$CDN = P(R).VO.TJ \quad (5)$$

onde $P(R)$ é a probabilidade de recebimento, isto é, $P(R) = 1 - P(I)$, VO é o valor original do empréstimo ou do financiamento e TJ é a taxa de juros. Em suma, todo o arcabouço teórico das medidas do risco e as políticas de crédito estão estreitamente relacionados com os elementos interdependentes e fundamentais das Finanças Corporativas.

2.2 Processo de Concessão de Crédito

No processo de concessão de crédito, os Modelos de *Credit Scoring* (MCSs) mais utilizados são aqueles baseados em julgamentos realizados por especialistas, principalmente na falta de dados históricos, ou ainda para tipos especiais de avaliações (ANDRIOSPOULOS et al., 2019; MARQUÉS; GARCÍA; SÁNCHEZ, 2013). Prática comum é considerar o modelo qualitativo dos clássicos 3 C's, 4 C's ou 5 C's: caráter, capacidade, capital, garantia e condições (MARQUÉS; GARCÍA; SÁNCHEZ, 2013). Entretanto, estes MCSs de análises julgamentais geram custos com formação de profissionais, decisões incorretas e inconsistências consideradas por diferentes analistas frente a mesma proposta (MARQUÉS; GARCÍA; SÁNCHEZ, 2013). Já os MCSs quantitativos são utilizados na maioria dos casos, visto que, baseiam-se na análise de

dados de inadimplência de empréstimos usando estimativas analíticas (ANDRIOSOPOULOS et al., 2019). Nessa direção, verifica-se um foco na utilização dos modelos quantitativos com suporte a decisão, tanto no que diz respeito à concessão como ao que cuida da gestão das carteiras (BRITO; ASSAF NETO, 2008). Em geral, a concessão de crédito pelas MPEs é uma estratégia para a conquista de clientes e estímulo das vendas, assim como uma das principais fontes de retorno e crescimento das vendas (DE AMORIM, 2019; POTRICH et al., 2012; ROSS; WESTERFIELD; JORDAN, 1995). Logo, os parâmetros de concessão correspondem às condições financeiras e ao histórico de crédito que o solicitante deve apresentar para tornar-se apto ao crediário (POTRICH et al., 2012).

2.3 Gestão do Risco de Crédito

Nas empresas não financeiras, o risco de crédito é igualmente importante, visto que, em suas operações diárias tais organizações também fornecem crédito à clientes (DOUMPOS et al., 2018). Logo, as empresas que colaboram com parceiros de baixa credibilidade podem enfrentar severas dificuldades financeiras. Por conseguinte, conforme Brito e Assaf Neto (2008), o risco de crédito pode ser compreendido a partir do: i) *risco de default* – diz respeito à probabilidade de ocorrer um evento de *default* em um certo período de tempo; ii) *risco de exposição* – decorre da incerteza em relação ao valor do crédito concedido quando ocorre o *default* e iii) *risco de recuperação* – se refere à falta de certeza no que trata do valor que pode ser recuperado na ocorrência do *default*. Estes consistem na possibilidade de o credor incorrer em perdas devido ao não cumprimento das obrigações assumidas pelo tomador de crédito (BRITO; ASSAF NETO, 2008).

Isto posto, sob o olhar de uma instituição financeira, conforme Silva (2016), os riscos podem ser classificados em quatro grupos: i) *risco do cliente* - relacionado as características do tomador de crédito; ii) *risco da operação* - tem como elementos o produto, o montante, o prazo, a forma de pagamento, as garantias e o preço; iii) *risco de concentração* - deriva da composição da carteira crédito e; iv) *risco da administração de crédito* – cuida dos recursos humanos e materiais da instituição e o papel destes nos níveis de inadimplência da carteira de crédito. Em termos gerais, os modelos de classificação de risco mensuram o risco de um tomador ou operação e expressam o resultado classificação ou pontuação que reflete a expectativa de risco de *default* (BRITO; ASSAF NETO, 2008). Ou seja, o objetivo de um MCS é classificar como “bons pagadores”, aqueles clientes onde é previsto que realização os pagamentos dentro do prazo, e como “maus pagadores”, aqueles em que é prevista a inadimplência (LI; CHEN, 2020; SICSÚ, 2010).

2.4 Algoritmo *Naive Bayes*

Fundamentado no Teorema de Bayes (1764), um dos classificadores mais antigos é o algoritmo *Naive Bayes* (NB). É reconhecido como um método simples, eficiente e popular, e muito utilizado em técnicas de *Data Mining* (LIU; BO, 2011; WU et al., 2008). Trata-se de um classificador estatístico fundamentado em diversos conceitos da teoria da probabilidade, como a probabilidade condicional, a independência condicional e a regra da distribuição conjunta de probabilidades (WU et al., 2008). Logo, é amplamente aplicado em muitas áreas com alta eficiência computacional e precisão satisfatória à realidade (LIU; BO, 2011).

É, também, um método de aprendizagem constantemente aplicado em muitos problemas de classificação supervisionados como, por exemplo, os encontrados nos MCSs (ANTONAKIS; SFAKIANAKIS, 2009). Portanto, dada as propriedades categóricas e a probabilidade condicional contrária, o Teorema de Bayes calcula a probabilidade condicional de um evento ocorrer com base em outras informações que podem ser relevantes (ZILHÃO, 2007; HARTSHORNA, 2016). A equação 6 traz o Teorema de Bayes.

$$P(A/B) = \frac{P(A).P(B/A)}{P(B)} \quad (6)$$

onde A e B são eventos, $P(A)$ e $P(B)$ são as probabilidades de A e de B sem consideração de um pelo outro, sendo que $P(A/B)$ é a probabilidade de A dado que B é verdadeiro e $P(B/A)$ é a probabilidade de B dado que A é verdadeiro (HARTSHORNA, 2016). Segundo Antonakis e Sfakianakis (2009), muitos estudos sugerem que o algoritmo NB é muito simples para a construção de regras de classificação se comparado com os métodos que têm uma melhor capacidade preditiva. Além disso o algoritmo NB também não tem estado no foco da pesquisa no que trata do *Credit Scoring* (ANTONAKIS; SFAKIANAKIS, 2009).

Nessa lógica, alguns autores veem o *Credit Scoring* (*Credit Scoring*) como uma aplicação direta do Teorema de Bayes dado que, o objetivo de um MCS é, com base em amostras de casos classificados como “bons” e “maus”, estimar a probabilidade de o cliente ser um “bom pagador” (SILVA, 2016). A utilização desta técnica apresenta a vantagem de o escore ser padronizado em um número de escala de 0 a 1 e que, observadas as devidas ressalvas, pode ser entendido como a probabilidade de o cliente ser um “bom pagador” (SILVA, 2016).

2.5 *Credit Scoring*

Na prática, uma das mais importantes ferramentas para a avaliação do risco na concessão de crédito e na inadimplência é o *Credit Scoring*. Esta ferramenta trata de uma denominação

genérica, com origem no mercado, dada para as fórmulas de cálculo da pontuação de crédito e que tem como objetivo quantificar o risco (LOUZADA; ARA; FERNANDES, 2016; SICSÚ, 2010). Assim sendo, o *Credit Scoring* consiste em uma estimativa baseada em modelos que buscam mensurar a probabilidade de um mutuário mostrar algum comportamento considerado indesejável no futuro (LESSMANN et al., 2015). Portanto, o *Credit Scoring* tem como o objetivo utilizar modelos estatísticos para transformar dados relevantes em medidas numéricas que orientem as decisões de crédito (ABDOU; POINTON, 2011).

A maioria dos MCSs visa identificar as características que influenciam os comportamentos que levam à efetivação ou não dos pagamentos de modo que seja possível classificar o cliente como um “bom pagador” ou “mau pagador” (LOUZADA; ARA; FERNANDES, 2016). Silva (2016) destaca que alguns autores entendem o *Credit Scoring* como uma aplicação direta do Teorema de Bayes, visto que, com base em alguns atributos cadastrais e amostras de casos “bons” e “ruins”, estima-se a probabilidade de o cliente ser um “bom pagador” ou “mau pagador”. Um modelo simplificado do *Credit Scoring* é dado pela expressão 7.

$$Theta = C + X_1p_1 + X_2p_2 + X_3p_3 + X_4p_4 + X_5p_5 \quad (7)$$

onde *Theta* é o valor obtido pela soma de uma constante com os produtos resultantes da multiplicação de cada variável *X* com base em seu respectivo peso *p* (SILVA, 2016). Assim, os clientes que o MCS apresenta com alta probabilidade de ser um “bom pagador” são aceitos enquanto aqueles com baixa probabilidade são rejeitados (FINLAY, 2009). A metodologia de *Credit Scoring* pode ser usada para diferentes situações, dentre outras, são exemplos os cartões de crédito, os pedidos de empréstimos pessoais, os empréstimos para casa, os empréstimos para pequenas empresas etc. (VUKOVIC et al., 2012).

3 REVISÃO E ANÁLISE DA LITERATURA

Esta seção expõe a revisão e análise da literatura sobre os mais recentes métodos de otimização utilizados na formulação dos MCSs. A seção se inicia com a revisão de literatura e se encerra com as análises dos MCSs.

3.1 Revisão da literatura

A demanda por uma análise para a concessão de crédito surgiu no início do comércio com vendas para pagamentos futuro (LOUZADA; ARA; FERNANDES, 2016). A aplicação da pontuação estatística na distinção entre os candidatos “bons” e “maus” foi apresentada, possivelmente pela primeira vez por Durand (1941) em um estudo que aborda os elementos de risco no financiamento ao consumidor (ABDOU; POINTON, 2011). Após tempo considerável, Altman (1968) aborda as taxas financeiras, a *Discriminant Analysis* (DA) e a previsão de falência corporativa para trazer o primeiro modelo de pontuação operacional com base em cinco índices selecionados de oito variáveis das demonstrações financeiras corporativas. Assim sendo, Altman (1968) realiza uma combinação linear dos referidos índices financeiros que resultou na apresentação de uma função discriminante Z . Por conseguinte, Orgler (1970) desenvolve um modelo geral de pontuação de crédito para empréstimos comerciais que aborda a questão das variáveis dependentes e independentes com a utilização da *Multivariate Regression* (MR) com a variável dependente restrita a valores zero ou um. A pesquisa traz uma aplicação prática de um MCS para empréstimos comerciais com dados obtidos a partir dos arquivos de examinadores bancários em vários estados da costa leste dos Estados Unidos.

Posteriormente, Eisenbeins (1978) analisa os tipos de MCSs, em especial, às abordagens metodológicas que têm sido empregadas e os problemas estatísticos associados a esses modelos usando DA. Segundo o autor, a questão da minimização das taxas de inadimplência é apenas uma face do problema na concessão de crédito, já que, sob o foco do credor, o objetivo de um MCS deveria ser a maximização do lucro ou a minimização do custo. Já na década de oitenta, Capon (1982) propõem um olhar crítico sobre a base lógica dos sistemas de *Credit Scoring*, pois a questão estatística empregadas a estes sistemas pode causar graves problemas jurídicos aos credores. As decisões de crédito são baseadas em variáveis que apresentam uma relação explicativa em vez de estatística. Isso porque, estas variáveis provavelmente não estão relacionadas a fatores econômicos e históricos de *Credit Scoring*. É somente após dez anos que Leonard (1992) modela o processo de decisão de crédito com o emprego de DA e LR. Para isso, o autor utiliza pedidos de empréstimos, para pequenas empresas, de um grande banco do

Canadá. A pesquisa demonstra que limitar a análise a um pequeno subconjunto homogêneo de toda a população permitirá a exploração de muitas características preditivas que estão presentes apenas em pequenos grupos avaliados por MCSs.

Entretanto, foi a partir dos anos dois mil que surgiram novos tipos de abordagem com a publicação de Baesens et al. (2003). Os autores avaliaram o risco de crédito de três conjuntos de dados utilizando três técnicas de extração de regras baseadas em NN (*Neurorule*, *Trepan* e a *Nefclass*). Estas ferramentas permitem ao analista explicar a classificação de determinado candidato com o conhecimento aprendido pelas redes. Além disso, permitiu a construção de sistemas de apoio a decisão avançados e fáceis de usar para avaliação do risco de crédito. Entretanto, foi somente após cinco anos que Sinha e Zhao (2008) comparam, com e sem a incorporação de conhecimento de domínio, o desempenho de sete métodos de classificação de *Data Mining* (NB, LR, DT, *Decision Table*, NN, KNN e *Support Vector Machine* (SVM)). A pesquisa foi aplicada em empréstimos bancários indiretos com incorporação do conhecimento de classificação de crédito ao processo de aprendizagem. O objetivo principal foi examinar o papel do conhecimento do domínio de aprendizagem no classificador de crédito. Os resultados corroboraram com a coesão entre o conhecimento de domínio e a *Data Mining*. Evidenciou-se que a experiência de domínio capturada na forma de uma base de conhecimento parcial pode melhorar significativamente o desempenho de uma ampla variedade de classificadores em pequenos conjuntos de dados relativos ao *Credit Scoring*.

Na sequência, Antonakis e Sfakianakis (2009) examinam a eficácia do teorema de *Bayes* como método para construção de regras de classificação para triagem de candidatos a crédito. Neste estudo, os pesquisadores utilizaram dois conjuntos de dados reais para comparar o Teorema de Bayes com a *Linear Discriminant Analysis*, a LR, a KNN, a *Classification Trees* e a NN. Os resultados demonstraram que o Teorema de Bayes tem uma capacidade preditiva inferior que alguns métodos abrangidos pela pesquisa. Porém, tem seus méritos por ser um método simples, de fácil compreensão e pode ser executado mais rapidamente que as alternativas mais sofisticadas. Finlay (2009) utiliza o *Genetic Algorithm* (GA) para criar um conjunto de modelos de pontuação linear direcionados por medidas individuais de interesse empresarial. No estudo, não houve diferenças relevantes entre os modelos derivados do GA com aqueles construídos utilizando a LR e a *Linear Regression Model* (LRM). Isso porque, o GA foi treinado para otimizar um único objetivo do negócio enquanto os modelos estatísticos foram desenvolvidos para otimizar funções mais genéricas do *Credit Scoring*.

Šušteršič, Mramor e Zupan (2009) desenvolveram um MCS, para consumidores com dados limitados e/ou não disponíveis de crédito, empregando NN e para *Variable Selection*

utilizou-se o *Principal Component Analysis* (PCA) e o GA. Segundo os autores, outros dados disponíveis dentro das instituições financeiras como, por exemplo, dados contábeis e saldos de contas de clientes foram utilizados na pesquisa. Nas análises dos resultados, o método proposto obteve 79,3% de precisão quando comparados aos modelos desenvolvidos com a LR. Ince e Aktan (2009) pesquisaram o desempenho de MCSs com aplicação de abordagens tradicionais e de *Machine Learning* como a DA, a LR, a NN e a DT. A pesquisa foi aplicada em um conjunto de dados com nove variáveis preditoras de um conjunto de 1.260 registro de clientes que solicitaram cartão de crédito em um banco da Turquia. Os resultados demonstram que a NN exibiu erros tipo II mais baixos enquanto o modelo que apresentou os melhores resultados em comparação às outras técnicas aplicadas foi o que utilizou a DT.

No ano de 2010, com o exponencial crescimento das pesquisas em *Credit Scoring*, as medidas financeiras contínuas como a inadimplência, a receita e a contribuição para os lucros são modeladas por Finlay (2010). O estudo concluiu que tais medidas baseadas em GA superam as funções de pontuação mais gerais como a soma do erro quadrático. Liu e Bo (2011) utilizam a ferramenta Otimização Global e o algoritmo *Simulated Annealing* juntamente com GA para realizar a seleção ideal de atributos, em bases de dados reais, do classificador NB. Os autores concluem que o NB baseado em GA pode ser mais eficaz e veloz se comparado com um classificador tradicional. Vukovic et al. (2012) traz um sistema com 4 modelos do *Case-Based Reasoning* (CBR) e para a escolha das funções de preferência e definição dos pesos dos atributos adota o GA. Este MCS é comparado ao KNN aplicado em três conjuntos de dados de crédito alemão e australiano do repositório da *University of California Irvine* (UCI), e dados do tutorial do software SPSS *Statistics Base para Windows* da IBM. É demonstrado que o MCS proposto superou os modelos tradicionais da abordagem CBR. Bravo, Maldonado e Weber (2013) apresentam uma metodologia de concessão e acompanhamento de crédito para microempreendedores aplicando a LR e o *Knowledge-Discovery in Databases* (KDD). Logo, foi apresentada uma metodologia de ponto de corte que determina a perda esperada no aceite de um candidato “ruim” e o custo de rejeitar um candidato “bom”.

Kruppa et al. (2013) aprimoraram a estimativa de probabilidade dos riscos envolvidos no crédito ao consumidor utilizando *Machine Learning*. Foram aplicados, a um conjunto de dados de uma empresa que produz eletrodomésticos, métodos como *Random Forest* (RF) e o *k-NN* juntamente com a LR. A pesquisa concluiu que as abordagens de *Machine Learning* são consistentes quando aplicadas a um grande conjunto de dados superando a LR. Řezáč (2014) utilizou o *ESIS2 Algorithm* para avaliar o poder discriminatório dos MCSs. É proposto um novo

ESIS2 ajustado para o objetivo de estimar o valor da informação no *Credit Scoring*. A conclusões são de que o ESIS2 é a escolha correta quando se trata de estimar o valor da informação na avaliação do poder discriminatório dos MCSs. Verbraken et al. (2014) encontram uma compensação entre as perdas esperadas e a perda por inadimplência por meio da adaptação da medida *Expected Maximum Profit* (EMP). É demonstrado que a medida EMP leva a MCSs mais lucrativos e que supera as abordagens alternativas de *Credit Scoring*. Kozeny (2015) preenche parcialmente uma lacuna na utilização dos GA no *Credit Scoring* pois estes assumiam papel de suporte para outras técnicas como NN. Este estudo concentra-se na função *fitness*, já que, a codificação dos dados deve representar o problema em questão e são cruciais para o desempenho do GA. Lessmann et al., (2015) compararam 41 classificadores em base de dados do mundo real mediante uma atualização da pesquisa de Baensens et al. (2003). Foi examinado até que ponto as avaliações dos MCSs tradicionais diferem dos novos indicadores alternativos com precisão preditiva, incluindo a comparação com os Esquemas Únicos, Híbridos e Conjuntos. Para os autores, a intensificação na pesquisa sobre *Credit Scoring* levou a necessidade de atualizar o estudo de *benchmark* incluindo-se classificadores de última geração e técnicas avançadas como Esquemas Conjuntos. As simulações revelaram que os Esquemas Conjuntos ocuparam o *top 10* do ranking, sendo que para os Esquemas Únicos as primeiras posições ficaram com a NN.

A partir da segunda metade da década de 2000, pesquisas como a de Serrano-Cinca e Gutiérrez-Nieto (2016) não se limitaram somente em prever a probabilidade de inadimplência na formulação dos MCSs. Para tanto, os autores propuseram um sistema de apoio a decisão de pontuação de lucro para empréstimo *Person to Person* (P2P) baseado em *Multivariate Regression* (MR) e na Taxa Interna de Retorno (TIR). A pesquisa revelou que clientes com alta probabilidade de inadimplência podem levar a resultados lucrativos, conforme a rentabilidade esperada de investir em empréstimos P2P. A justificativa dos autores se baseia no fato de que os fatores que explicam a lucratividade são diferentes daqueles que buscam prever a inadimplência como os propostos pelos MCSs. Maldonado et al. (2017) desenvolveram uma estrutura baseada no lucro para a seleção de modelos de análise de custo-benefício e o cálculo das perdas financeiras para inadimplentes, baseada em *Support Vector Machine* (SVM). A pesquisa incorporou informações relacionadas ao negócio, considerando os benefícios e perdas ocasionadas pela classificação correta e incorreta dos modelos. Este modelo, por meio de uma análise detalhada de custo-benefício que incluiu o cálculo das perdas financeiras para inadimplentes, foi aplicado em dois conjuntos de dados de clientes microempresários de um

banco do Chile. Krichene (2017) empregou o algoritmo NB na previsão de inadimplência em empréstimos de curto prazo de um banco comercial da Tunísia. A pesquisa utilizou uma base de dados com 924 arquivos de créditos contendo variáveis de entrada classificadas em duas categorias relacionadas e não relacionadas ao fluxo de caixa para os anos de 2003 até 2006. A análise dos resultados com a utilização do NB demonstrou que uma boa taxa de classificação é da ordem de 63,85% com uma *Area Under the ROC Curve* (AUC) na ordem de 69%. A pesquisa também realizou a comparação com um modelo que emprega a NN e que apresentou uma AUC na ordem de 83%.

Por conseguinte, Bastani, Asgari e Namavari (2019) propuseram uma abordagem de pontuação em dois estágios por meio de uma pesquisa focada no processo de alocação de fundos no mercado de empréstimos P2P. Este estudo integra o *score* de crédito e de lucro com base no *Deep Learning Algorithm* desenvolvido por cientistas do Google. No primeiro estágio, a pontuação de crédito é realizada por meio da análise de probabilidade de inadimplência utilizando a LR. No segundo estágio, a previsão de rentabilidade é focada nos empréstimos identificados como não inadimplentes com utilização da TIR. Sariannidis et al. (2019) analisam os dados de pagamentos de uma carteira de cartões de crédito contendo cerca de 30.000 clientes com 23 atributos. A pesquisa compara a precisão da previsão de 7 métodos de classificação como a LR, o NB, a DT, a RF, o KNN, a SVC e a SVC Linear com precisão dos modelos que varia de 70% a 82,6%. A pesquisa pode ser aplicada por instituições financeiras ou empresas de cartão de crédito para categorizar potenciais clientes com condições de solvência. Kozodoi et al. (2019) se basearam em invólucro para tratar a lucratividade e a interpretabilidade utilizando número de atributos como duas funções de adequação para a seleção de características e a EMP. A estrutura de invólucro multi-objetiva foi desenvolvida utilizando o algoritmo NSGA-II com duas funções de aptidão, além do número de recursos e o EMP. A eficácia do modelo foi testada executando experimentos empíricos em 10 conjuntos reais de dados de *Credit Scoring*. Çiğşar e Ünal (2019) lidaram com a identificação e prevenção do risco de inadimplência utilizando algoritmos de classificação *Data Mining*. Assim, o trabalho emprega 6 algoritmos de classificação, sendo o NB, a RF, o *J48 Algorithm*, o *Bayesian Network* e o *Multi Layer Perceptron* com utilização do software de *Data Mining WEKA 3.9*. Para o estudo foram utilizados dados de 12 variáveis obtidos do instituto de estatística da Turquia. A pesquisa conclui que as instituições de crédito devem considerar as características e circunstâncias de seus clientes nos MCSs.

Mais recentemente, os pesquisadores passaram a reforçar a utilização e a combinação de mais de uma técnica nos MCSs. Trivedi (2020) traz um modelo de previsão e *Credit Scoring* aplicado a uma base de dados da Alemanha disponível ao público utilizando os classificadores NB, RF, DT e SVM. Trivedi (2020) realiza uma análise comparativa dos classificadores de seleção de características e compara três técnicas de seleção de recursos, Qui-Quadrado, Ganho de Informação e Taxa de Ganho. As análises demonstraram que o Qui-Quadrado foi considerado adequado com a maioria preditores informativos e para todos os modelos de *Machine Learning*. No que diz respeito aos modelos de *Machine Learning*, a RF foi, de forma geral, considerada a melhor entre outros classificadores cuja precisão foi de 93%. Nalić e Martinovic (2020) propuseram um MCS personalizado e de alto desempenho baseado no histórico de crédito denominado *Reliable Client* e empregaram o *Generalized Linear Algorithm* e o SVM em dados reais de um banco da Bósnia e Herzegovina. A pesquisa revelou que o nível de acerto nas previsões dos empréstimos inadimplentes no processo de treinamento foi de 91,07%. Assim, com a utilização do MCS o banco teria acertado em 511 decisões a mais, levando a uma queda nas perdas de 31,64%. Li e Chen (2020) trazem uma importante contribuição experimental sobre a escolha dos modelos de previsão de risco de crédito com avaliação comparativa de quatro *Ensemble Algorithms* (RF, a *AdaBoost*, o *XGBoost* e a *LightGBM*). Para tanto, aplicaram quatro algoritmos tradicionais (NN, a LR, a DT e o SVM) a um conjunto de dados de crédito do *Lending Club* nos Estados Unidos. As experimentações revelaram que o desempenho da aprendizagem em conjunto é melhor do as individuais, exceto para a *AdaBoost*. Além disso, a RF tem o melhor desempenho em termos de quatro métricas do *XGBoost* e da *LightGBM*. As conclusões são de que o melhor classificador que, para a maioria das métricas de avaliação, supera todo os outros quatro da linha de base é a LR.

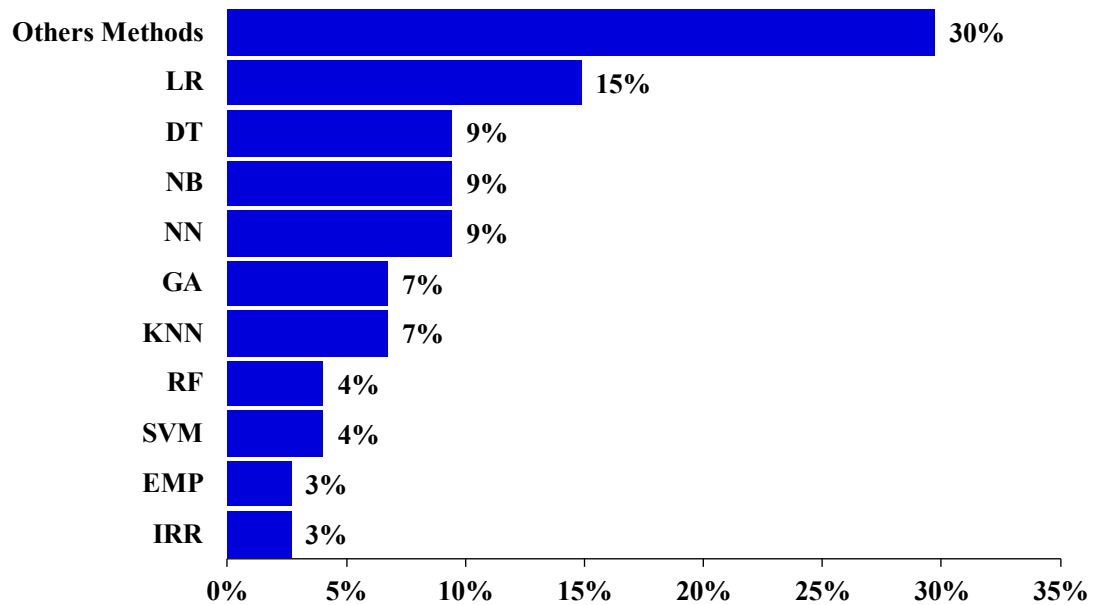
3.2 Análise da Literatura

A literatura demonstra que muitas técnicas são utilizadas no cálculo do nível de risco de crédito e na formulação do MCSs. Em especial, estão abordagens como a LR, a NN, o GA, a DT e o KNN, incluindo os modelos de Pesquisa Operacional como a *Linear Programming* e a *Quadratic Programming*. Além disso, também são empregadas outras técnicas como o NB e a *Factor Analysis*. As pesquisas trazidas pela revisão de literatura, bem como os métodos de solução propostos, são apresentadas pelo quadro 1. Na sequência, os referidos métodos com duas ou mais aplicações são ilustrados pelo gráfico da figura 1. Logo, os métodos de solução com apenas uma aplicação são agrupados no item *Others Methods*.

Quadro 1 – Relação de Artigos e Técnicas Aplicadas

Referências	Citações	Métodos Aplicados
Sinha e Zhao (2008)	158	NB, LR, DT, <i>Decision Table</i> , NN, <i>k</i> -NN
Antonakis e Sfakianakis (2009)	41	NB, <i>Linear Discriminant Analysis</i> , LR, <i>k</i> -NN, <i>Classification Trees</i> e NN
Finlay (2009)	35	GA
Ince e Aktan (2009)	89	DT, LR, NN e DT e LR
Šušteršič, Mramor e Zupan (2009)	144	NN e <i>Variable Selection</i> , PCA e GA
Finlay (2010)	64	GA
Liu e Bo (2011)	13	<i>Simulated Annealing Algorithm</i> , GA e NB
Vukovic et al. (2012)	84	CBR, GA e <i>k</i> -NN
Bravo, Maldonado e Weber (2013)	58	LR, KDD
Kruppa et al. (2013)	164	<i>Random Forest</i> , <i>k</i> -NN e LR
Řezáč (2014)	2	<i>ESIS2 Algorithm</i> e <i>Monte Carlo Simulation</i>
Verbraken et al. (2014)	159	EMP, NN e LR
Kozeny (2015)	80	GA
Serrano-Cinca e Gutiérrez Nieto (2016)	229	MR, DT e TIR
Maldonado et al. (2017)	83	SVM
Krichene (2017)	13	NB e NN
Bastani, Asgari e Namavari (2019)	49	SMOTE, IHT, LR e TIR
Çiğşar e Ünal (2019)	41	NB, <i>Bayesian Network</i> , <i>J48 Algorithm</i> , RF, <i>Multi Layer Perceptron</i> e LR
Sariannidis et al. (2019)	7	<i>k</i> -NN, LR, NB, DT, RF, SVC e SVC
Kozodoi et al. (2019)	52	EMP e <i>NSGA-II Algorithm</i>
Li e Chen (2020)	24	<i>Ensemble Algorithms</i> : RF, <i>AdaBoost</i> , <i>XGBoost</i> , <i>LightGBM</i> e NN, DT, LR, NB e SVM
Nalić e Martinovic (2020)	2	<i>Generalized Linear Algorithm</i> e SVM
Trivedi (2020)	15	NB, RF, DT, SVM

Fonte: Elaborado pelo autor. Referências às abreviações: NB – *Naive Bayes*; MR – *Multivariate Regression*; GA – *Genetic Algorithm*; NN – *Neural Network*; LR – *Logistic Regression*; LRM – *Linear Regression Model*; IHT - *Instance Hardness Threshold*; DT – *Decision tree*; RF – *Random Forest*; EMP – *Expected Maximum Profit*; KDD – *Knowledge - Discovery in Databases*; *k*-NN - *K-Nearest Neighbors*; LLE – *Locally Linear Embedding*; NMF - *Non-Negative Matrix Factorization*; PCA – *Principal Component Analysis*; SMOTE - *Synthetic Minority Oversampling Technique*; SVC – *Support Vector Clustering*; SVM- *Support Vector Machine*; TIR – *Taxa Interna de Retorno*.

Figura 1 – Métodos Usados pelos Modelos de *Credit Scoring*

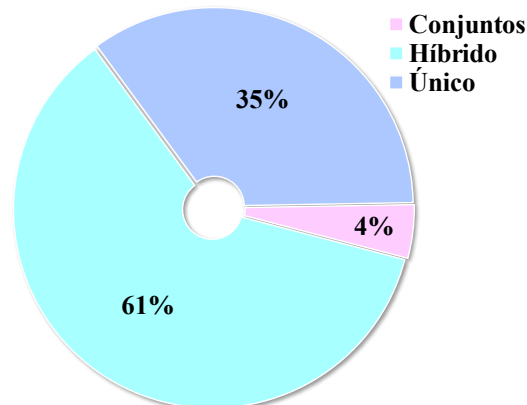
Fonte: Elaborado pelo autor

No quadro 1, é apresentada a quantidade total de citações de cada artigo que propõe os 33 métodos empregados na construção dos MCSs. Portanto, o gráfico da figura 1 expõe o percentual destes métodos de solução aplicados aos MCSs. Nesse sentido, fica claro que, independentemente de ser ou não combinado com outros métodos, a LR (11%) é o método mais utilizado, seguido da DT (9%), do NB (9%) e das NN (9%). Em relação à construção dos MCSs, três esquemas metodológicos (consulte Louzada, Ara e Fernandes (2016) e Andriosopoulos et al. (2019)) podem ser identificados, conforme demonstrado pelo quadro 2. Na sequência, a figura 2 exhibe os tipos de modelagem utilizados na pesquisa de acordo com a classificação de Louzada et al. (2016).

Quadro 2 – Esquemas Metodológicos Aplicados às Modelagens

Esquemas	Citações	Referências
Único	24	Li e Chen (2020)
Híbrido	1.127	Antonakis e Sfakianakis (2009), Bravo, Maldonado e Weber (2013), Çiğşar e Ünal (2019), Ince e Aktan, (2009), Kozodoi et al. (2019), Kruppa et al. (2013), Liu e Bo (2011), Nalić e Martinovic (2020), Řezáč (2014), Serrano-Cinca e Gutiérrez-Nieto (2016), Šušteršič, Mramor e Zupan (2009), Vukovic et al. (2012) e Verbraken et al. (2014).
Conjuntos	455	Finlay (2009), Finlay (2010), Kozeny (2015), Krichene (2017), Maldonado et al. (2017), Sariannidis et al. (2020), Sinha; Zhao (2008) e Trivedi (2020).

Fonte: Elaborado pelo autor.

Figura 2 – Classificação dos Modelos de *Credit Scoring*

Fonte: Elaborado pelo autor.

A classificação metodológica apresentada pelo quadro 2 e a figura 2 demonstram o número de citações e o valor percentual das abordagens que tratam os Esquemas Único, Esquemas Híbridos e Esquemas de Conjunto. Portanto, a construção de MCSs com a aplicação de um único método é denominada de Esquema Único (ANDRIOSOPOULOS et al., 2019). Já os Esquema Híbridos combinam diferentes técnicas visando uma melhora na capacidade de desempenho e de adaptabilidade dos MCSs (LOUZADA; ARA; FERNANDES, 2016; ANDRIOSOPOULOS et al., 2019). Nos Esquema Conjuntos muitas técnicas são exploradas na análise, entretanto, apenas uma é normalmente implantada na previsão final dos MCSs (CHEN et al., 2016). Lin et al. (2012) apresenta três abordagens para construir um Esquema Híbrido, conforme apresentado pelo quadro 3.

Quadro 3 – Abordagens para Construção de Sistemas Híbridos

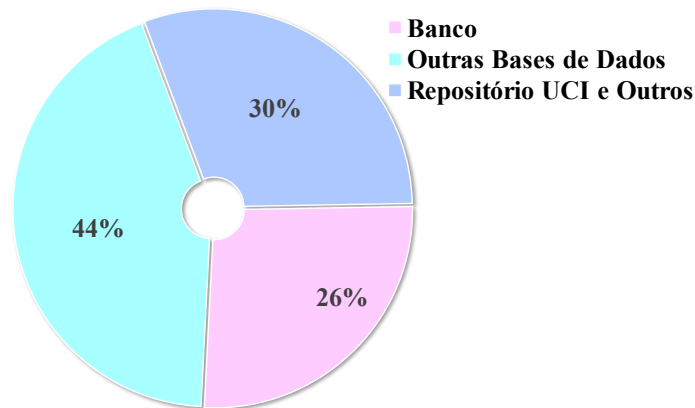
Abordagem	Técnicas Aplicadas
Modo de Combinação	As técnicas heurísticas são integradas a modelos de classificação para otimizar o desempenho de predição a partir de várias visões.
Modo Cascata	Classificadores diferentes em cascata, onde a saída do classificador de primeiro nível alimenta o classificador de segundo nível como entrada.
Modo de Clustering	Armazenamento em <i>cluster</i> é usado como uma etapa de pré processo de classificação para aumentar a precisão da previsão.

Fonte: Elaborado pelo autor.

À vista disso, Esquemas de Conjunto combinam classificadores de múltiplos modelos básicos que se utilizam de um único classificador ou vários algoritmos para obter melhores previsões. Logo, são frequentemente utilizados três modelos (ensacar, impulsionar e empilhar) e os resultados dependem da diversidade e da diminuição do viés de cada modelo (LOUZADA; ARA; FERNANDES, 2016; ANDRIOSOPOULOS et al., 2019). Em seguida, as análises dos

estudos demonstraram que os pesquisadores utilizaram grandes bases de dados com muitas variáveis na construção dos MCSs. Essas bases de dados podem ser sintetizadas em três categorias: i) Banco, ii) Outras Bases de Dados e; iii) Repositório UCI e Outros. A frequência das bases de dados utilizadas na pesquisa é ilustrada pela figura 3.

Figura 3 – Bases de Dados Usadas pelos Modelos de *Credit Scoring*



Fonte: Elaborado pelo autor

Em relação à origem dos dados, foram utilizadas 14 diferentes fontes, sendo que 6 pesquisas se basearam em bancos de várias nações, em especial, da Europa. Logo, outros 5 autores procuraram bancos de dados disponíveis no Repositório de Banco de Dados de *Machine Learning* da UCI. Além disso, 2 artigos manipularam dados da UCI e de plataformas como os bancos gregos PAKDD e Kaggle, e ainda de instituições financeiras da Benelux e do Reino Unido. Ademais, outros autores se utilizaram de outras fontes como, por exemplo, o *Experian UK*, o *Taiwan Economic Journal (TEJ)*, o *Germany and Australia Credit Bureau*, o Clube de Empréstimos nos Estados Unidos, Plataformas P2P na China etc.

É importante ressaltar que a pesquisa mais recente sobre *Credit Scoring* se baseia no lucro e na estimativa da rentabilidade dos empréstimos e não somente na probabilidade de inadimplência. Isto porque, as causas da rentabilidade são diferentes das causas da inadimplência, ou seja, clientes com alta probabilidade de inadimplência também podem ser rentáveis (SERRANO-CINCA; GUTIÉRREZ-NIETO, 2016; ONAY et al., 2018). Assim, os MCSs baseados na rentabilidade recorrem a abordagens como, por exemplo, o cálculo da TIR, o comportamento financeiro contínuo, a introdução de novas variáveis-alvo e a combinação de técnicas estatísticas e inteligência na construção de pontuações de lucro de empréstimo (consulte Finlay (2010), Bravo et al. (2013), Řezáč (2014), Verbraken et al. (2014), Serrano-Cinca e Gutiérrez-Nieto (2016), Onay et al. (2018) e Kozodoi et al. (2019)).

Estudos atuais também demonstram como a *big data* está levando os MCSs a mudanças disruptivas com a incorporação de um grande volume e variedade de dados. Além disso, está a

velocidade na coleta e armazenamento destes dados. Logo, há hoje uma consideração mais ampla de dados não apenas oriundos dos históricos de pagamentos e recebimentos dos requerentes ao crédito. Há também dados obtidos das redes sociais, das informações de aplicativos e das denominadas “pegadas digitais”. Assim, a *big data* permite a avaliação da qualidade do crédito em potenciais mutuários com histórico financeiro limitado (ONAY et al., 2018). Evidentemente, os bancos, as agências de crédito, os *fintechs* e outros provedores não bancários de serviços financeiros também se utilizarão de *big data* para níveis mais elevados de precisão em seus serviços. Fato é que essa realidade traz novos desafios regulatórios em termos de prevenção da discriminação e dos direitos do consumidor (ONAY et al., 2018).

4 METODOLOGIA DE PESQUISA

Esta seção apresenta todas as etapas que compõem a metodologia empregada para a elaboração desta pesquisa e do MCS. Além disso, é realizada uma breve análise bibliométrica das pesquisas e periódicos que abordam os MCSs.

4.1 Classificação

Esta é uma pesquisa aplicada de natureza quantitativa e caracterizada como um estudo de caso que explora análises documentais mediante uma abordagem bibliográfica (consulte Bertrand e Fransoo 2002; Haegeman et al. (2013), Linnenluecke, Marrone e Singh (2020) e Donthu et al. (2021). Isto posto, esta pesquisa se utiliza de técnicas e modelos matemáticos propostos pela literatura para, fundamentada por uma base de dados real, formular um MCS. Todos os dados utilizados são oriundos de cadastros e transações comerciais de uma MPE que atua no ramo do comércio de materiais de construção no interior de São Paulo (SP).

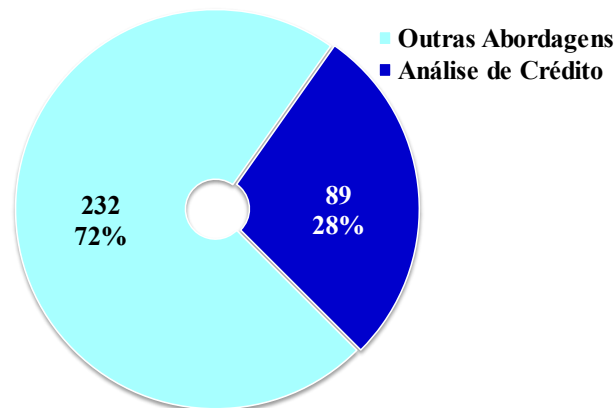
4.2 Método de Pesquisa

A pesquisa realizou-se com base em uma busca tradicional usando os bancos de dados *Scopus* e *Web of Science* (consulte Baker (2016), Winchester e Salji (2016)). As palavras-chave usadas nas buscas avançadas foram *Naive Bayes*, *Credit Scoring*, *Small Business*, *Small Sized Companies* e *Credit Risk Management*. Para que fossem encontrados resultados que contenham obrigatoriamente a combinação destas palavras-chave foi utilizado a expressão booleana “AND”. A seleção e classificação das publicações foram realizadas em três etapas por meio de uma avaliação criteriosa de relevância, acessibilidade, número de citações e enquadramento ao objetivo do estudo. Na primeira etapa, a investigação on-line (de 1968 a 2020) teve início em janeiro de 2021 e foi realizada até meados de março de 2021. Após as buscas e a exclusão dos artigos repetidos, foram selecionadas 647 obras de periódicos internacionais publicados somente no idioma inglês. Na segunda etapa, foram selecionados 321 artigos publicados entre os anos de 2008 até 2020. Estas datas foram definidas pelo crescimento exponencial das publicações segundo Louzada et al. (2016).

Na terceira etapa, foi realizada a seleção das publicações consideradas aderentes aos objetivos de pesquisa e formulação do MCS. Assim, foram selecionados 31 artigos publicados em revistas de alto impacto (9% dos 321). Todos os demais artigos foram descartados, já que, não se encaixam no protocolo estabelecido devido aos tipos de abordagem e/ou outras especificações inerentes aos MCSs. Entre os artigos selecionados, 8 são de revisão de literatura e apresentam abordagens com diferentes focos de pesquisa nos MCSs. As outras abordagens

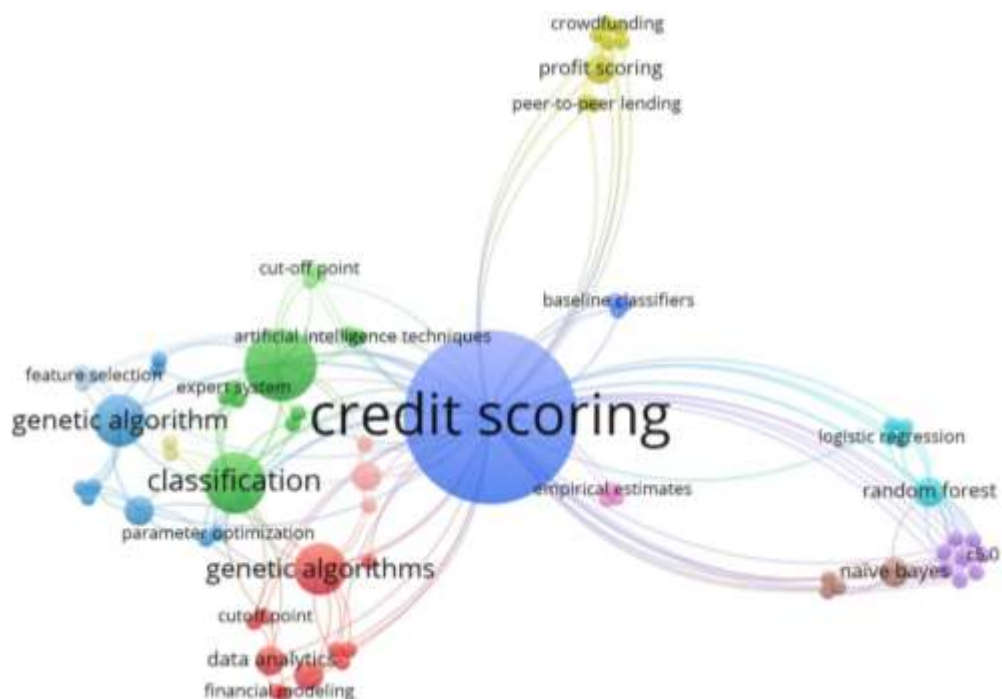
propõem MCSs, formulados mediante a aplicação de diferentes técnicas e métodos, para resolver uma diversidade de problemas de *Credit Scoring*. Logo, abarcam a classificação, a modelagem, a rentabilidade e a seleção de banco de dados para aplicação de MCSs. A quantidade de pesquisas que abordam a análise de crédito é ilustrada pelo gráfico da figura 4 e totalizam apenas 89 (28% das 321). Além disso, dentre estas 89 pesquisas, somente 43 (48%) abordam a análise e concessão de crédito no âmbito das MPEs. Na sequência, uma análise prévia identificou as palavras-chave mais utilizadas pelos artigos selecionados, como demonstrado pelo mapa de nuvem da figura 5 gerado pela versão VOSviewer 1.6.16 (<http://www.vosviewer.com/>).

Figura 4 – Abordagem das Pesquisas Seleccionadas



Fonte: Elaborado pelo autor.

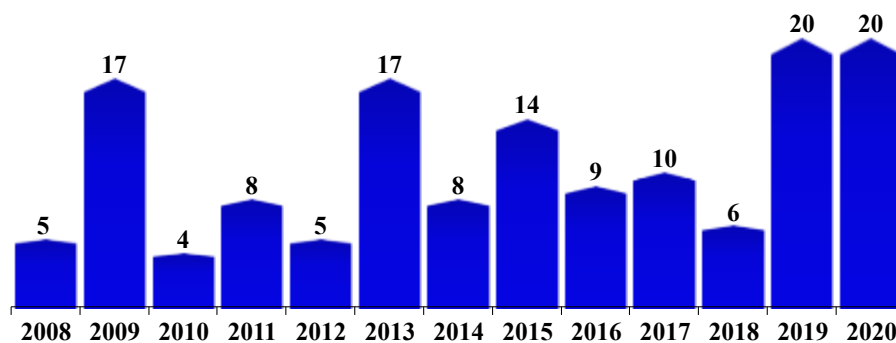
Figura 5 – Mapa de Relações das Palavras-Chave



Fonte: Elaborado pelo autor

A figura 5 demonstra que 90 palavras-chave foram listadas nas bases dos artigos selecionados (VOSviewer). Além disso, mostra que 72 *hits* consistem em palavras-chave relacionadas ao *Credit Scoring*. As cinco principais palavras-chave por quantidade de acessos entre 2008 até 2020 foram *Credit Scoring* (21), *Data Mining* (6), *Classification* (5), *Genetic Algorithm* (4) e *Genetic Algorithms* (4). A totalidade dos *hits* para todas as palavras-chave durante o período de pesquisa foi de 174 vezes, conforme ilustrado pelo gráfico da figura 6.

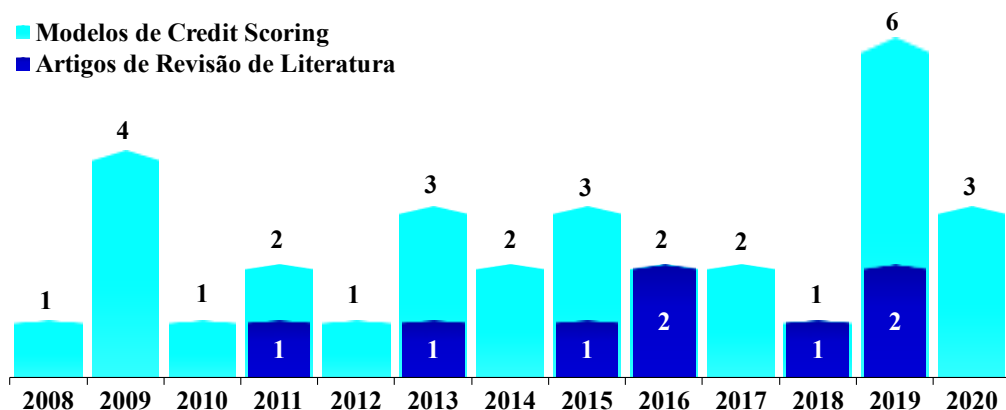
Figura 6 – Ocorrência de Palavras-Chave por Ano



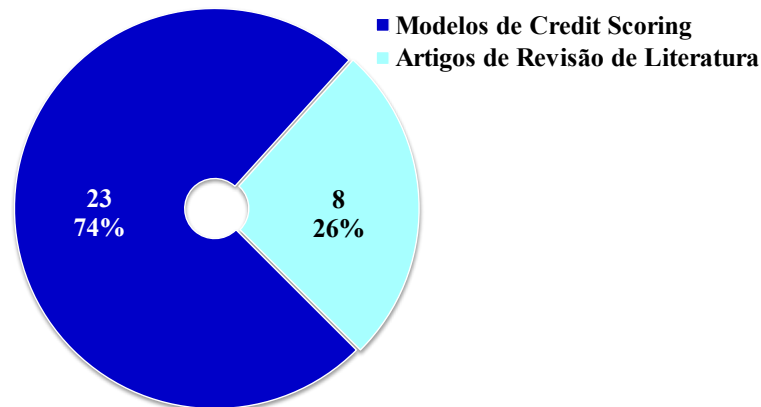
Fonte: elaborado pelo autor

As relações de intensidade e ocorrência das palavras-chave ilustradas pelo gráfico da figura 6 evidenciam que os artigos selecionados são pertinentes a uma revisão de literatura sobre MCSs. Por conseguinte, uma distribuição temporal dos artigos selecionados, que concatena tanto as quantidades das abordagens de revisão de literatura quanto daquelas que propõem MCSs, é apresentado pela figura 7. A figura 7 evidencia que durante a última década houve uma constância na quantidade de pesquisas propondo MCSs, com um aumento significativo em 2019. Na sequência, uma categorização entre estas diferentes abordagens é apresentada pelo gráfico da figura 8 e demonstra que dos 31 artigos selecionados tem-se que 8 (26%) são de revisão de literatura e 23 (74%) propõem MCSs.

Figura 7 – Distribuição Temporal das Publicações



Fonte: Elaborado pelo autor.

Figura 8 – Categorização das Publicações de *Credit Scoring*

Fonte: Elaborado pelo autor.

Em revisões de literatura, segundo Pae (2015) e Jahan et al. (2016), duas diferentes abordagens são observadas: i) Revisão Narrativa e; ii) Revisão Sistemática. A Revisão Narrativa traz à tona o estado da arte de um determinado assunto com um foco teórico ou contextual. A Revisão Sistemática utiliza métodos integrativos explícitos e sistemáticos que visam responder a uma ou mais perguntas específicas de forma a localizar, selecionar e avaliar tecnicamente a pesquisa. A classificação das pesquisas de revisão de literatura é apresentada pelo quadro 4.

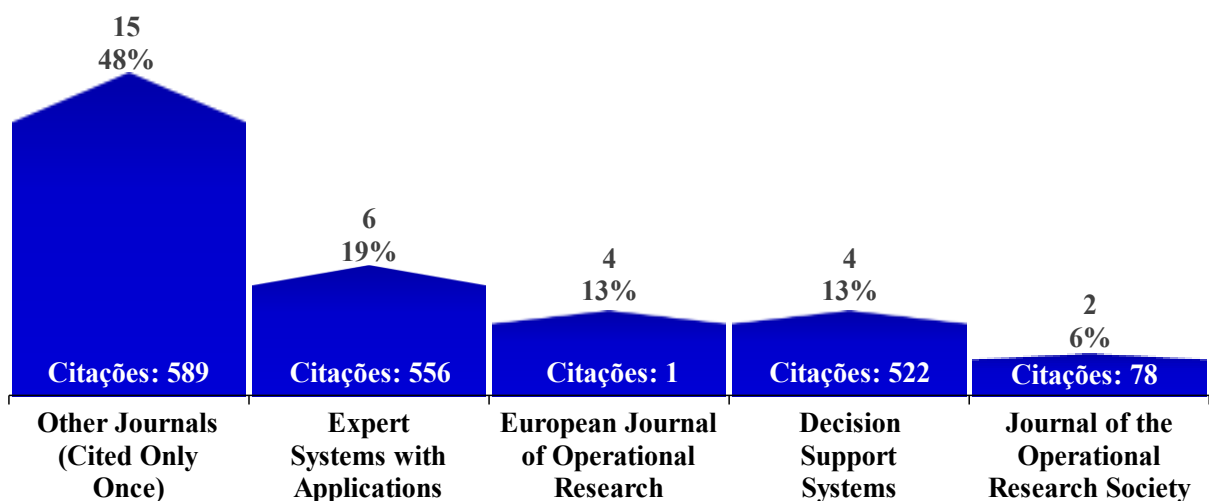
Quadro 4 – Classificação das Pesquisas de Revisão de Literatura

Referências	Citações	Metodologia de Pesquisa	Tipo	
			N	S
Abdou e Pointon (2011)	146	214 artigos/livros/teses de <i>Credit Scoring</i> para diversas áreas, especialmente para bancos e finanças.		✓
Marqués et al. (2013)	58	Artigos de conferência e de revistas publicados no período de 2000 até 2012.		✓
Lessmann et al. (2015)	750	41 classificadores em oito conjuntos de dados de <i>Credit Scoring</i> do mundo real.	✓	
Chen et al. (2016)	57	Não Especificado.	✓	
Louzada et al. (2016)	83	437 artigos (<i>Reaxys</i> , <i>Scopus</i> , <i>ScienceDirect</i> e <i>Engineering Information</i>).		✓
Onay et al. (2018)	27	299 artigos (<i>ProQuest</i> e <i>Emerald Research</i>).		✓
Andriosopoulos et al. (2019)	20	Não especificado.	✓	
Goh e Lee (2019)	29	Artigos com <i>Google Scholar</i> , <i>Science Direct</i> e <i>IEEE Xplore/Electronic Library Online</i> (1997 até 2018).		✓

Fonte: Elaborado pelo autor. Referências às abreviações: N – Revisão Narrativa; S – Revisão Sistemática.

A classificação das pesquisas de revisão de literatura demonstra que 3 (37,5%) artigos são de Revisão Narrativa e 5 (62,5%) de Revisão Sistemática. Além disso, o quadro 4 também apresenta a quantidade total de citações de cada artigo conforme as bases de dados *Scopus* e *Web of Science*. Note-se que o artigo de Lessmann et al. (2015) foi o mais citado, um total de 750 (64%). Também é importante destacar a opção pela exclusão de livros, resumos, editoriais e outros estudos que não foram classificados como artigos de Revisão Narrativa ou Revisão Sistemática. Portanto, a seleção dos artigos de revisão de literatura, consoante com Pae (2015) e Jahan et al. (2016), baseou-se no papel expressivo que estes desempenham na transferência de informações históricas e na atualização e aprimoramento dos MCSs. Em seguida, a figura 9 junto com o quadro 5 e a tabela 1 apresentam os indicadores bibliométricos, conforme as bases de dados *Scopus* e *Web of Science*, referentes aos periódicos que propuseram revisões de literatura e MCSs.

Figura 9 – Periódicos com Abordagens de *Credit Scoring*



Fonte: Elaborado pelo autor.

A tabela 1, resumida no gráfico 9, mostra que foram selecionados artigos de 19 diferentes periódicos, dos quais 15 foram classificados como *Others Journals* (Cited Only Once). A maioria destes periódicos, 16 (84%), estão concentrados na Europa e os únicos 3 (16%) periódicos de outros continentes são a *Scientific Programming* (Egito), a *Advances in Operations Research* (Estados Unidos) e o *International Journal of Software Engineering and Knowledge Engineering* (Cingapura).

Quadro 5 – Periódicos com Publicações de *Credit Scoring*

Periódico	Referências
Expert Systems with Applications	Finlay (2009), Šušteršič, Mramor e Zupan (2009), Vukovic et al. (2012), Kruppa et al. (2013), Kozeny (2015) e Bastani, Asgari e Namavari (2019)
Decision Support Systems	Sinha e Zhao (2008), Serrano-Cinca and Gutiérrez Nieto (2016), Maldonado et al. (2017) and Kozodoi et al. (2019)
European Journal of Operational Research	Finlay (2010), Bravo, Maldonado e Weber (2013), Verbraken et al. (2014) e Lessmann et al. (2015)
Journal of the Operational Research Society	Marqués, García e Sánchez (2013) e Andriosopoulos et al. (2019)
Mathematics	Li e Chen (2020)
Procedia Engineering	Liu e Bo (2011)
Technology in Society	Trivedi (2020)
Scientific Programming	Çiğşar e Ünal (2019)
Computational Economics	Řezáč (2014)
<i>Artificial Intelligence Review</i>	Antonakis e Sfakianakis (2009)
Journal of Applied Statistics	Chen, Ribeiro e Chen (2016)
Annals of Operations Research	Sariannidis et al. (2019)
Advances in Operations Research	Goh e Lee (2019)
Journal of Financial Regulation and Compliance	Onay e Ozturk (2018)
Journal of Business Economics and Management	Ince e Aktan (2009)
Journal of Economics Finance and Administrative Science	Krichene (2017)
Surveys In Operations Research and Management Science	Louzada et al. (2016)
Intelligent Systems in Accounting Finance and Management	Abdou e Pointon (2011)
International Journal of Software Engineering and Knowledge Engineering	Nalić e Martinovic (2020)

Fonte: Elaborado pelo autor.

Tabela 1 – Indicadores dos Periódicos com Publicações de *Credit Scoring*

Título do Periódico das Publicações Selecionadas	Total de Citações	Percentil Periódicos	Escore de Citação	Fator de Impacto	H-Index Periódicos	Quartil Periódicos
Expert Systems with Applications	293	98%	12.7	6.954	207.0	Q1
European Journal of Operational Research	557	97%	9.5	5.334	260.0	Q1
Decision Support Systems	284	98%	10.5	5.795	151.0	Q1
Journal of the Operational Research Society	20	87%	4.1	2.860	108.0	Q1
Mathematics	24	80%	2.2	2.258	32.0	Q1
Procedia Engineering	13	80%	4.0	1.880	74.0	Q1
Technology in Society	9	90%	4.2	4.192	51.0	Q1
Scientific Programming	27	41%	2.0	1.025	36.0	Q3
Computational Economics	2	72%	2.3	1.876	40	Q2
<i>Artificial Intelligence Review</i>	57	99%	10.4	8.139	78	Q1
Journal of Applied Statistics	41	62%	1.9	1.404	57	Q2
Annals of Operations Research	7	83%	5.2	4.854	105	Q1
Advances in Operations Research	19	55%	2.9	3.579	14	Q2
Journal of Financial Regulation and Compliance	16	40%	1.6	0.761	19	Q3
Journal of Business Economics and Management	89	78%	3.5	2.445	37	Q1
Journal of Economics Finance and Administrative Science	13	62%	1.5	1.270	13	Q2
Computers & Operations Research	-	93%	7.0	4.008	152	Q1
Intelligent Systems in Accounting Finance and Management	146	89%	4.1	5.500	11	Q1
International Journal of Software Engineering and Knowledge Engineering	-	48%	2.3	1.470	36	Q3

Fonte: Elaborado pelo autor.

4.3 Procedimentos Operacionais

Esta pesquisa lida com um problema real de *Credit Scoring* observado em uma MPE que atua no ramo do comércio de materiais de construção no interior do estado de São Paulo (SP). A tratativa do problema se fundamentou na revisão de literatura sobre a gestão do risco de crédito nas vendas a prazo e nas soluções propostas pelos MCSs. Logo, a pesquisa estruturou um embasamento teórico pertinente a respeito dos MCSs. Por conseguinte, foi realizado a coleta e análise de dados na empresa e que incluiu: i) Diagnóstico Financeiro; ii) Cadastros dos Clientes e; iii) Processo das Vendas à Crédito. A partir deste diagnóstico, foram definidos e detalhados a problemática e os objetivos da construção e aplicação do MCS. Todos os dados analisados foram então tratados considerando os aspectos da moderna Teoria das Finanças. Na busca pela construção de um MCS simples, eficaz e de fácil implementação no Microsoft Office Excel 2016 optou-se por utilizar o algoritmo *Naive Bayes* (NB). Assim sendo, a probabilidade de inadimplência e às questões relativas ao *Credit Scoring* ponderado por decisões de investimento e financiamento foram parametrizados e testados no MCS. Em suma, a pesquisa foi desenvolvida seguindo as seguintes etapas:

1. Identificação e escopo do problema de pesquisa;
2. Revisão de literatura sobre a aplicação dos MCSs;
3. Procedimentos metodológicos para aplicar o MCS;
4. Levantamento dos procedimentos de vendas à crédito;
5. Análise dos dados e diagnóstico financeiro da empresa;
6. Construção do MCS adaptado à realidade da empresa em estudo;
7. Experimentações, considerações finais e contribuições da empresa.

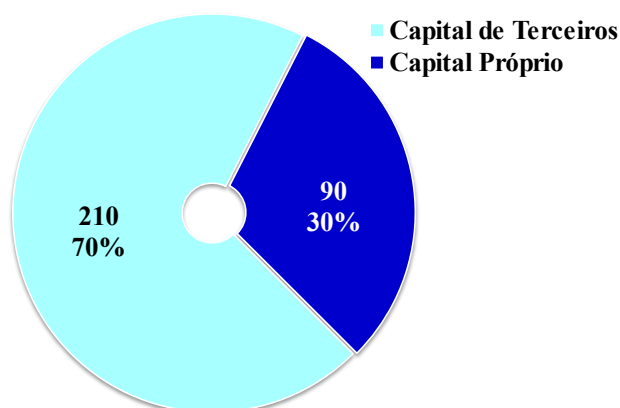
5 MODELO DE *CREDIT SCORING* PROPOSTO

Esta seção detalha a representação do problema, demonstra a modelagem matemática, realiza as experimentações e apresenta as análises dos resultados de forma que seja possível elucidar todos os pressupostos que norteiam a construção do MCS.

5.1 A Empresa

A empresa objeto de estudo atua no ramo da construção civil, especificamente no comércio de materiais de construção, e está localizada no interior de São Paulo (SP). Está inserida em um mercado de intensa concorrência cujas empresas do mesmo município que também atuam no comércio varejista de materiais de construção já somam 62. Trata-se de uma sociedade por cotas de responsabilidade limitada com dois sócios cotistas, sendo um o administrador do negócio, cada qual com um percentual de cotas de 50%. É optante pelo simples nacional e os serviços relativos às obrigações fiscais e de departamento pessoal são realizados por um escritório terceirizado. Todo o controle das vendas e dos estoques é feito pelo administrador que utiliza um software licenciado e para controle financeiro são usadas planilhas do Microsoft Office Excel 2016. A empresa foi fundada com o objetivo de atender o mercado local e regional na oferta de materiais de construção para o varejo. Iniciou suas atividades em 01/10/2019 com um capital inicial de R\$ 300 mil sendo que, R\$ 90 mil foram provenientes de recursos dos sócios e R\$ 210 mil tomados no mercado financeiro em empréstimos bancários. Este empréstimo foi tomado com taxa de juros nominal de 2,7% a.m., prazo de 36 meses e com carência de 60 dias para pagamento da primeira parcela. A composição de capital da empresa em percentuais é ilustrada pelo gráfico da figura 10.

Figura 10 – Composição do Capital da Empresa



Fonte: Elaborado pelo autor.

Em junho de 2020 foi tomado novo empréstimo junto à instituição financeira na monta de R\$ 100 mil com taxa de juros nominal de 2,7% a.m., prazo de 36 meses e com carência de 60 dias para pagamento da primeira parcela. O capital de terceiros é constituído por empréstimos bancários, modalidade de capital de giro, ambos com prazo de 36 meses e no ano de 2021 geraram o pagamento de mais de 126 mil reais por meio de parcelas mensais, incluídos os juros e amortização do saldo devedor. Para melhor compreensão destas questões, a tabela 2 traz o Demonstrativo de Resultado do Exercício (DRE) simplificado da loja, com valores expressos em milhares de reais, referente aos anos de 2019 a 2021. O modelo de DRE adotado é o proposto por SEBRAE (2017) adaptado para a inclusão da Provisão com Créditos de Liquidação Duvidosa (PCLD).

Tabela 2 – Demonstrativos do Resultado do Exercício (DRE)

Demonstrativo de Resultado	2019	2020	2021
Receita Bruta	R\$ 271,86	R\$ 3.234,83	R\$ 5.541,90
(-) Deduções e Abatimentos	-R\$ 22,50	-R\$ 90,00	-R\$ 90,00
(=) Receita Líquida	R\$ 249,36	R\$ 3.144,83	R\$ 5.451,90
(-) Custo das Mercadorias Vendidas	-R\$ 209,13	-R\$ 2.488,33	-R\$ 4.313,79
(=) Lucro Bruto	R\$ 40,24	R\$ 656,50	R\$ 1.138,11
(-) Despesas com Vendas	-R\$ 5,44	-R\$ 64,70	-R\$ 110,84
(-) Despesas Administrativas	-R\$ 94,80	-R\$ 410,40	-R\$ 434,40
(-) Despesas Financeiras	-R\$ 9,48	-R\$ 150,57	-R\$ 126,74
(-) Despesas com PCLD	-R\$ 4,38	-R\$ 98,18	-R\$ 376,64
(=) Resultado Líquido	-R\$ 73,87	-R\$ 67,35	R\$ 89,49

Fonte: Elaborado pelo autor. Referências: PCLD – Provisão com Créditos de Liquidação Duvidosa.

Nos DREs apresentados na tabela 2 é possível verificar a evolução da Receita Bruta bem como, o crescimento com Despesas Financeiras. O incremento na Receita Bruta levou a realização de um Resultado Líquido positivo em 2021 (R\$ 89,49 mil), conforme os dados consolidados até setembro de 2021. Nessa lógica, note que o impacto da inadimplência no resultado com a PCLD chega a consumir 33,1% do Lucro Bruto.

5.2 Cadastro de Clientes

A empresa possui um total de 1.963 clientes cadastrados, considerando dados até setembro de 2021. As vendas formalizadas no balcão, com retirada da mercadoria pelo cliente, são contabilizadas no cliente código “00001 – consumidor”. Entretanto, o cadastro de um cliente só é realizado nos casos de vendas de maior volume, vendas que demandem a realização de entrega pela loja e vendas realizadas com contratação de crediários. Estes clientes são

cadastrados em um sistema fornecido por uma empresa de software e apresenta os seguintes campos para preenchimento:

- Código Cliente (Gerado pelo sistema)
- Nome (Fornecido pelo cliente)
- Cliente Desde (Gerado pelo sistema)
- Endereço (Fornecido pelo cliente)
- Bairro (Fornecido pelo cliente)
- Cidade (Fornecido pelo cliente)
- Estado Civil (Fornecido pelo cliente)
- Mora com os pais (Fornecido pelo cliente)
- Casa própria (Fornecido pelo cliente)
- CPF/CNPJ (Fornecido pelo cliente)
- RG (Fornecido pelo cliente)
- Carteira de Trabalho (Fornecido pelo cliente)
- Profissão (Fornecido pelo cliente)
- Empresa (Fornecido pelo cliente)
- Endereço da Empresa (Fornecido pelo cliente)
- Outros dados (Fornecido pelo cliente)
- Salário (Fornecido pelo cliente)
- Referência 1 (Fornecido pelo cliente)
- Referência 2 (Fornecido pelo cliente)
- Referência 3 (Fornecido pelo cliente)
- SPC (Consulta Sistema)
- Habilitado (Fornecido pelo cliente)
- Data da Habilitação (Fornecido pelo cliente)
- Dados do Cônjuge (Fornecido pelo cliente)

Entretanto, em consultas realizadas nas bases de dado verificou-se que somente são preenchidos os campos “Nome”, “CPF/CNPJ”, “Endereço”, “Bairro” e “Cidade”. É importante destacar que, os campos “Estado Civil” e “Mora com os Pais” são preenchidos com as mesmas informações para todos os clientes, ou seja, “*pro forma*”. Portanto, os demais campos não são preenchidos, ou seja, informações acerca das rendas, trabalhos ou faturamento dos credores não estão disponíveis. Dessa maneira, foram extraídas as informações de 3 variáveis do cadastro do cliente: i) Código Cliente; ii) Valor da Compra e; iii) CPF/CNPJ (Tipo de Pessoa – Pessoa Física ou Pessoa Jurídica). Além disso, não há um histórico de relacionamento com os clientes, já que, foi excluído o campo “Cliente Desde”. Assim sendo, a base de dados é de baixa qualidade e, portanto, apresenta uma visão restrita das variáveis que podem influenciar tanto na classificação do risco de crédito quanto na formulação do MCS.

5.3 Procedimento das Vendas

Na realização das vendas, além do pagamento em Dinheiro (Espécie à Vista), há diferentes formas de pagamento disponibilizadas aos clientes. Estas opções abrangem o pagamento com cartões emitidos por instituições financeiras nas modalidades, débito, crédito à vista, crédito parcelado e crediário. No crediário do cartão o banco emitente financia a venda em até 60 parcelas. Nos casos das vendas no crediário e no cartão de crédito parcelado, a empresa recebe o valor descontado das taxas e tarifas em 2 dias. A empresa aceita também o pagamento com a utilização de cheques negociados como cheque à vista e cheque pré-datado. Para a opção de pagamento com cheques o empresário financia a venda e assume o risco de crédito do emitente, sendo estas vendas classificadas como crediário. Nesta modalidade, também podem ser aceitos como forma de pagamento, a depender do negócio e contexto,

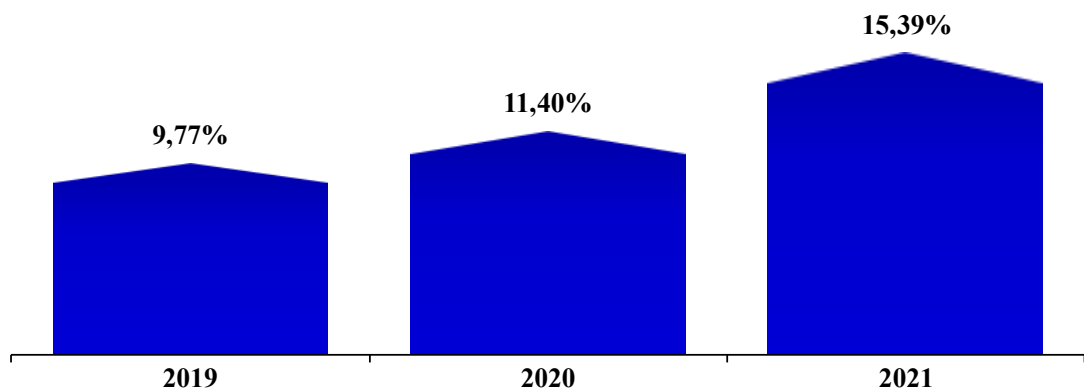
cheques emitidos por terceiros. Cabe destacar que o crediário tem participação significativa nas receitas da empresa, sendo a composição das vendas (em milhares de reais) apresentadas pela tabela 3. Na sequência, o histórico da participação do crediário nas vendas da empresa é apresentado pela tabela 3 e pelo gráfico da figura 11.

Tabela 3 – Histórico da Participação do Crediário nas Vendas

Tipo de Venda	2019	2020	2021
Crediário	R\$ 24,80	R\$ 359,88	R\$ 852,75
Cartão de Crédito	R\$ 31,65	R\$ 338,78	R\$ 472,66
Cartão de Débito	R\$ 52,39	R\$ 390,81	R\$ 327,12
Dinheiro (À Vista)	R\$ 145,01	R\$ 2.066,55	R\$ 3.888,50
Total Geral das Vendas	R\$ 253,85	R\$ 3.156,00	R\$ 5.541,03

Fonte: Elaborado pelo autor.

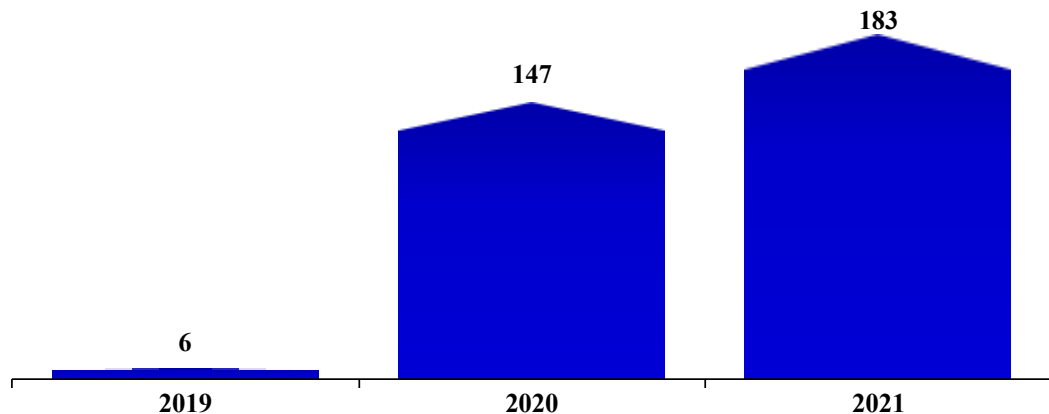
Figura 11 – Evolução das Vendas no Crediário



Fonte: Elaborado pelo autor.

A tabela 3 demonstra também que o progresso das vendas no Crediário, em relação as vendas totais da empresa, fechou em mais de 852 mil reais em setembro de 2021. A figura 11 revela que o crediário apresentou um acréscimo superior às vendas totais com um aumento de 57,54% em 2021. No processo de análise das vendas no crediário o empresário realiza consultas nos sistemas de proteção ao crédito como o SCPC da Associação Comercial de São Paulo e nas bases da Serasa. As consultas abrangem a verificação da existência de registro do documento do proponente nos órgãos de proteção ao crédito bem como, ocorrência do documento do proponente no Cadastro de Emitentes de Cheques Sem Fundo (CCF). Após análise do cliente e do negócio e, baseado na experiência e conhecimento empírico, o empresário decide sobre a aprovação da venda no crediário. Nessa prática de negócios, a empresa realizou 330 vendas no Crediário, conforme representado no gráfico da figura 12.

Figura 12 – Volume de Vendas no Crediário por Ano



Fonte: Elaborado pelo autor.

É possível verificar na figura 12 que o número de clientes que acessou o Crediário cresceu 2.950% de 2019 até setembro de 2021. Isto evidencia que no cotidiano do negócio é comum a realização de vendas no crediário com cheques pré-datados. As análises realizadas ainda identificaram uma grande dispersão referente aos valores comprados pelos clientes no crediário. A seguir, com base nas variáveis inseridas no cadastro da empresa, o perfil dos clientes classificados por faixa de valor que utilizaram o crediário são retratados pela tabela 4.

Tabela 4 – Perfil dos Clientes que Compram no Crediário

Faixa de Valor	Número de Clientes	Percentual (%)
Até 2 mil	195	59%
Até 5 mil	62	19%
Até 10 mil	38	12%
Até 20 mil	26	8%
Acima de 20 mil	9	3%

Fonte: Elaborado pelo autor.

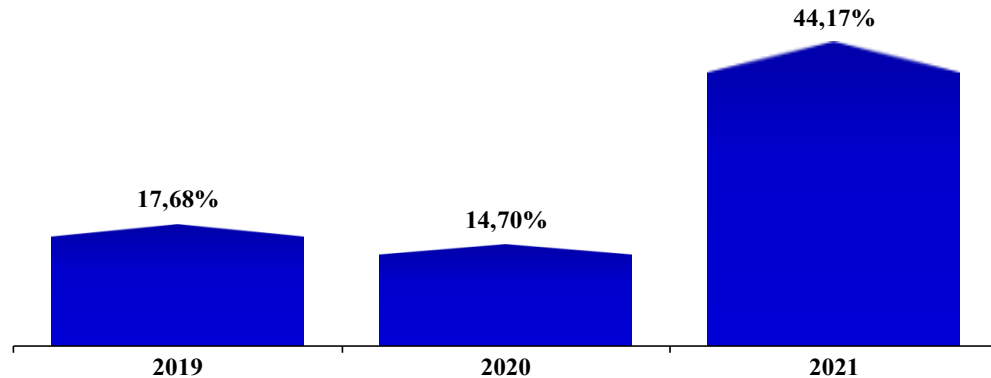
A tabela 4 demonstra a dispersão entre os valores das vendas para clientes que solicitaram o crediário, onde 195 clientes realizaram compras de até R\$ 2 mil e apenas 9 clientes efetuaram compras acima de R\$ 20 mil. Fica claro que 59% dos clientes realizaram compras de até R\$ 2 mil e apenas 3% compras acima de R\$ 20 mil. Ao que cuida da inadimplência, os valores em atraso contabilizaram, até setembro de 2021, o montante de R\$ 479,2 mil (5,3%). A totalidade dos valores não recebidos das vendas no Crediário (em milhares de reais) e seus impactos nos faturamentos totais da empresa é detalhada pela tabela 5. Em seguida, os percentuais referentes as inadimplências das vendas no Crediário são ilustradas pelo gráfico da figura 13.

Tabela 5 – Valores em Atraso das vendas no Crediário

Inadimplência	2019	2020	2021
Valores em Atraso	R\$ 4,38	R\$ 98,18	R\$ 376,64
Valor em Atraso/Faturamento Total	1,6%	3,0%	6,8%

Fonte: Elaborado pelo autor.

Figura 13 – Inadimplência das Vendas no Crediário



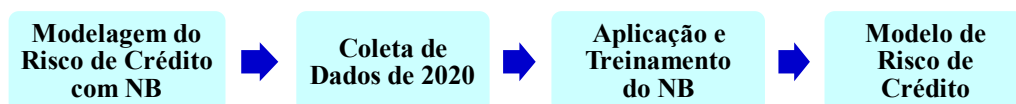
Fonte: Elaborado pelo autor.

A tabela 5 indica que, se comparado ao faturamento total, o saldo de inadimplência para ano de 2021 foi de 6,8%, ou seja, houve um crescimento de 322,4% em relação à 2019. E quando se compara as quantias em situação de inadimplência frente às vendas no Crediário o impacto é ainda maior, chegando a 44,2% no ano de 2021. A figura 13 evidencia o enorme problema de inadimplência no ano de 2021, pois a parcela de valores não pagos cresceu 149,8% em relação à 2019. Estas são as principais questões acerca da situação financeira da empresa em estudo desde 2019. Assim sendo, tanto a operacionalização das vendas em cada modalidade quanto os dados relativos à inadimplência constituem as bases de construção do MCS.

5.4 Modelo de *Credit Scoring*

A construção do MCS é composta de três etapas cuja primeira consiste no cálculo da probabilidade de adimplência e inadimplência de cada cliente. Esta etapa é denominada de Fluxo do Modelo de Risco de Crédito e, baseado nos dados cadastrais, é formulada por meio da aplicação do algoritmo NB, de acordo com a figura 14.

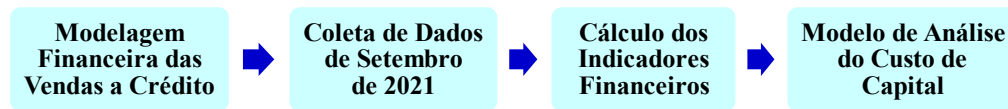
Figura 14 – Fluxo do Modelo de Risco de Crédito



Fonte: Elaborado pelo autor.

Em seguida, com base nos dados e históricos da empresa, são efetuados os cálculos das variáveis financeiras para obtenção e análise do Custo do Capital e do *Weighted Average Cost of Capital* (WACC), conforme demonstrado pela figura 15.

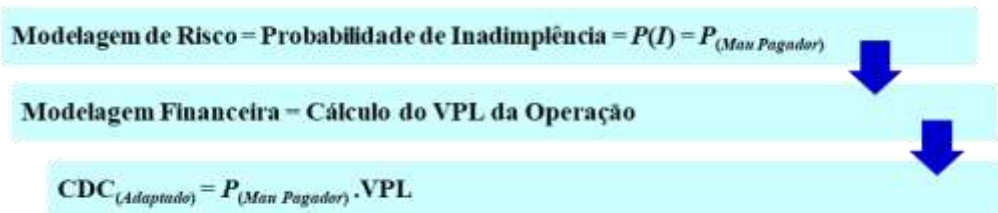
Figura 15 – Fluxo do Modelo de Análise do Custo de Capital



Fonte: Elaborado pelo autor.

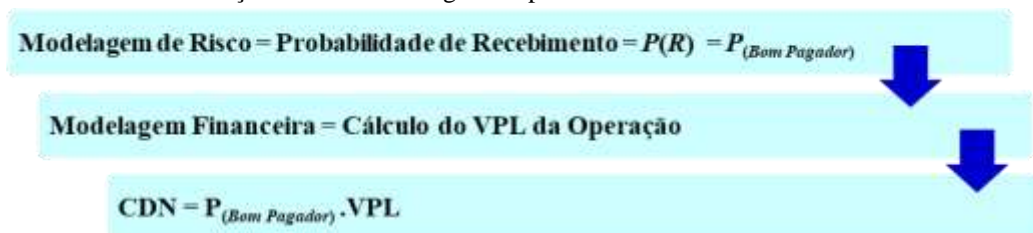
Por último, é formulada a adaptação do Custo de Negar (CDN) e do Custo de Conceder (CDC) para cálculo do escore das operações a serem configurados no MCS. Todo o fluxo do CDN e do CDC adaptados são respectivamente ilustrados pelas figuras 16 e 17.

Figura 16 – Fluxo de Elaboração do Custo de Conceder Adaptado



Fonte: Elaborado pelo autor.

Figura 17 – Fluxo de Elaboração do Custo de Negar Adaptado



Fonte: Elaborado pelo autor.

A estrutura apresentada pela figura 15 traz o modelo de análise do Custo de Capital, que trata do investimento na concessão de crédito. Logo, os resultados da modelagem financeira possibilitam o cálculo do Valor Presente Líquido (VPL) e a Taxa Interna de Retorno (TIR) aplicada ao cálculo consiste no WACC. A apuração do VPL juntamente com os resultados da modelagem de risco torna possível, conforme demonstrado pelas figuras 16 e 17, o cálculo da escoragem da operação por meio do $CDN_{(Adaptado)}$ e do $CDC_{(Adaptado)}$. As figuras 16 e 17 também reforçam a combinação dos modelos de risco e de avaliação de capital na formulação do CDC e CDN. As modelagens do Risco de Crédito e do Custo de Capital são detalhadas nas subseções 5.4.1 e 5.4.2.

5.4.1 Modelagem do Risco de Crédito

A modelagem do Risco de Crédito tem como primeira etapa a classificação dos clientes inadimplentes como “mau pagador” e adimplentes como “bom pagador”. Assim, clientes que apresentarem algum valor em atraso até dezembro de 2020 são classificados como “mau

pagador”, sendo todos os outros identificados como “bom pagador”. Após esta classificação, é calculado a frequência de cada atributo obtido dos dados cadastrais dos clientes e dos cheques selecionados. A distribuição das vendas no campo “Faixa de Valor” fora então dividida de acordo com as 99 ocorrências na faixa de até R\$ 2 mil e as 3 acima de R\$ 20 mil. Esta distribuição que justifica a aplicação do modelo considerando as diferentes faixas de valor das vendas é demonstrada pela tabela 6.

Tabela 6 – Distribuição das vendas por faixa de valor

Quantidade de Vendas por Faixa de Valor		Total	Faixa de Valor (%)
Faixa de Valores	Até 2 mil	99	67,4%
	Até 5 mil	23	15,7%
	Até 10 mil	17	11,6%
	Até 20 mil	5	3,4%
	Maior que 20 mil	3	2,0%

Fonte: Elaborado pelo autor.

Na tabela 6, é possível visualizar que 67,4% das vendas no Credíário foram realizadas em valores de até R\$ 2 mil e somente 2,0% para valores acima de R\$ 20 mil. Nesse sentido, a frequência de clientes classificados como “bom pagador” e “mau pagador” por campo selecionado é apresentada pela tabela 7.

Tabela 7 – Classificação de Clientes da Empresa

Atributos Cadastrais dos Clientes		Mau Pagador	Bom Pagador	Total
Clientes	Atributos	63	84	147
Tipo de Pessoa	PF	40	74	114
	PJ	23	10	33
Restrição Cadastral	Sim	25	7	32
	Não	38	77	115
Cheque Devolvido	Sim	20	11	31
	Não	43	73	116
Faixa de Valores	Até 2 mil	38	61	99
	Até 5 mil	12	11	23
	Até 10 mil	7	10	17
	Até 20 mil	4	1	5
	Maior que 20 mil	2	1	3

Fonte: Elaborado pelo autor. Referências às abreviações: PF – Pessoa Física; PJ – Pessoa Jurídica.

Após o levantamento das frequências trazidas pela tabela 7 é então formulada a matriz de cálculo da probabilidade (P) de um cliente ser “mau pagador” e “bom pagador”. A probabilidade (P) é obtida pela expressão $P = \frac{n(A)}{n(\Omega)}$, onde $n(A)$ é o número de casos que interessam ao evento A e $n(\Omega)$ é o número total de casos possíveis para “mau pagador” e “bom pagador”. A tabela 8 detalha a matriz de cálculo da probabilidade (P).

Tabela 8 – Matriz de Cálculo da Probabilidade (P)

Atributos Cadastrais dos Clientes		Mau Pagador		Bom Pagador	
Tipo	PF	40/114	35,09%	74/114	64,91%
	PJ	23/33	69,70%	10/33	30,30%
Restrição cadastral	Sim	25/32	78,13%	7/32	21,88%
	Não	38/115	33,04%	77/115	66,96%
Cheque devolvido	Sim	20/31	64,52%	11/31	35,48%
	Não	43/116	37,07%	73/116	62,93%
Faixa de valor	Até 2 mil	38/99	38,38%	61/99	61,62%
	Até 5 mil	12/23	52,17%	11/23	47,83%
	Até 10 mil	7/17	41,18%	10/17	58,82%
	Até 20 mil	4/5	80,00%	1/5	20,00%
	Maior que 20 mil	2/3	66,67%	1/3	33,33%

Fonte: Elaborado pelo autor.

Posteriormente ao cálculo da probabilidade (P) do cliente ser um “mau pagador” e “bom pagador”, a probabilidade condicional calculada pelo algoritmo NB é dada por pela equação 8.

$$P(A/B) = \frac{P(A) * P(B/A)}{P(B)} \quad (8)$$

Assim, para a construção do NB seja a_1, \dots, a_n atributos do banco de dados e c uma classe a ser prevista, então, a previsão ótima é uma classe de valor c tal que $P(a_1, \dots, a_n/c)$. Logo, considerando a independência entre os atributos tem-se $\frac{P(a_1/c) * P(a_n/c) * P(c)}{P(a_1) * P(a_n)}$, conforme obtido pela equação 8. Aplicado ao caso em questão, o cálculo para se prever a probabilidade de um cliente ser “bom pagador”, considerando as características do cadastro e das pesquisas do cheque pré-datado, é dada pela equação 9.

$$P_{(bom\ pagador)} = \frac{P(tipo/bom) * P(restrição/bom) * P(cheque/bom) * P(valor/bom)}{P(total\ clientes/bom)} \quad (9)$$

sendo P a probabilidade para cada variável dada a condição de “bom pagador”. Já o cálculo para se prever a probabilidade de um cliente ser “mau pagador”, onde P é a probabilidade para cada variável dada a condição de “mau pagador”, é dada pela equação 10

$$P_{(\text{mau pagador})} = \frac{P(\text{tipo/mau}) * P(\text{restrição/mau}) * P(\text{cheque/mau}) * P(\text{valor/mau})}{P(\text{total clientes/mau})} \quad (10)$$

A partir dos resultados da matriz de probabilidades apresentada na Tabela 8 é possível a aplicação do NB e cálculo dos escores de “bom pagador” e de “mau pagador” para os clientes que tomaram crédito até o ano de 2020. Logo, os escores dos clientes para todas as combinações possíveis entre os atributos selecionados são demonstrados pela tabela 9.

Tabela 9 – *Escores* dos Clientes da Empresa

Restrição Cadastral	Cheque Devolvido	Faixa de Valor	Tipo PF		Tipo PJ	
			“Mau”	“Bom”	“Mau”	“Bom”
Sim	Nada consta	Até 2 mil	0,3470	0,6530	0,6933	0,3067
Sim	Consta	Até 2 mil	0,6212	0,3788	0,8747	0,1253
Não	Consta	Até 2 mil	0,1847	0,8153	0,4909	0,5091
Não	Nada consta	Até 2 mil	0,0684	0,9316	0,2380	0,7620
Sim	Nada consta	Até 5 mil	0,4820	0,5180	0,7983	0,2017
Sim	Consta	Até 5 mil	0,7417	0,2583	0,9244	0,0756
Não	Consta	Até 5 mil	0,2841	0,7159	0,6280	0,3720
Não	Nada consta	Até 5 mil	0,1139	0,8861	0,3536	0,6464
Sim	Nada consta	Até 20 mil	0,7733	0,2267	0,9355	0,0645
Sim	Consta	Até 20 mil	0,9133	0,0867	0,9782	0,0218
Não	Consta	Até 20 mil	0,5927	0,4073	0,8609	0,1391
Não	Nada consta	Até 20 mil	0,3204	0,6796	0,6673	0,3327
Sim	Nada consta	Maior que 20 mil	0,6304	0,3696	0,8789	0,1211
Sim	Consta	Maior que 20 mil	0,8404	0,1596	0,9573	0,0427
Não	Consta	Maior que 20 mil	0,4211	0,5789	0,7558	0,2442
Não	Nada consta	Maior que 20 mil	0,1907	0,8093	0,5007	0,4993

Fonte: Elaborado pelo autor.

A tabela 9 também detalha todas as combinações possíveis dos atributos cadastrais para os clientes tipo Pessoa Física (PF) e Pessoa Jurídica (PJ) bem como, as probabilidades de cada combinação em ser “bom pagador” e “mau pagador” segundo aplicação do NB. Logo, os escores detalhados na, tabelas 9 proporcionam uma visão do risco de inadimplência para tipo de cliente e a combinação de atributos cadastrais. Essa probabilidade é fruto da aplicação do

NB e revela a probabilidade condicional para cada combinação possível. Isto é, considera as probabilidades de cada atributo para chegar ao resultado da ocorrência de determinado evento no contexto da probabilidade de ocorrência de outros eventos.

5.4.2 Modelagem do Custo de Capital

Na modelagem do Custo de Capital, pelo fato do objeto de estudo ser uma pequena empresa, foi utilizado o conceito do Beta Realavancado. Assim sendo, o Beta Realavancado é dado pela expressão $\beta r = \beta d \cdot \left[1 + \left(\frac{D}{E} \right) \cdot (1 - T) \right]$, em que o βr é o Beta Realavancado, βd é o Beta Desalavancado, D é o Endividamento, E é o Valor de Mercado/Patrimônio Líquido e T = Alíquota de Imposto Sobre a Renda. Para esta modelagem, o βd corresponde ao Beta do último ano da ação LJQQ3 da Lojas Quero-Quero, que é uma grande varejista de materiais de construção com ações negociadas na B3. A tabela 10 demonstra o cálculo do Beta Realavancado e do *Capital Asset Pricing Model* (CAPM).

Tabela 10 – Beta Realavancado e *Capital Asset Pricing Model*

Beta Realavancado (βr)		CAPM	
βd	0,68	R_f	0,52%
D/E	4,44	βr	3,02
βr	3,02	R_m	0,70%
		CAPM	1,06%

Fonte: Elaborado pelo autor.

sendo R_m o Retorno de Mercado, baseado nos dividendos da Lojas Quero-Quero em 2021, e o R_f a Taxa Selic Meta. Por conseguinte, a tabela 11 apresenta, em milhares de reais, o resultado do *Weighted Average Cost of Capital* (WACC).

Tabela 11 – Cálculo do *Weighted Average Cost of Capital*

Cálculo do WACC	Principal	Taxa (a.m.)	Juros (a.m.)
Capital Próprio	R\$ 90	1,06%	R\$ 0,95
Capital de Terceiros	R\$ 310	2,70%	R\$ 8,37
Total	R\$ 400	2,33%	R\$ 9,32

Fonte: Elaborado pelo autor.

A tabela 11 demonstra que o custo financeiro da empresa referente ao financiamento das operações de concessão de crédito, dado pelo WACC. Para tanto foi calculado o custo com o Capital de Terceiros e o custo com o Capital Próprio por meio da aplicação do *Capital Asset*

Pricing Model (CAPM). É importante ressaltar que a modelagem financeira foi elaborada com base nos dados do ano de 2020.

5.4.3 Formulação do Modelo de *Credit Scoring*

Na modelagem proposta ao *Credit Scoring*, o modelo de risco de crédito e o modelo de análise do custo de capital são alicerces para a construção do MCS. Para construção do MCS foi calculado o VPL das operações, dado pela expressão $VPL = -I + \sum_{t=1}^n \frac{FC_t}{(1+i)^t} = 0$. Dessa maneira, o valor do empréstimo ou das prestações são trazidos para valor presente descontados com a taxa de juros (mensal em relação ao prazo das prestações) do cálculo do WACC.

À vista disso, foram realizadas adaptações com o emprego da probabilidade de ser “bom pagador” ou “mau pagador” apresentada no modelo de risco de crédito para o MCS. Logo, a equação de cálculo do CDN ($CDC = P(I) \cdot VO + CI + CB$) foi adaptada com a substituição do $VO + CI + CB$ pelo VPL. Já a equação de cálculo do CDC ($CDN = P(R) \cdot VO \cdot TJ$) foi adaptada com a substituição do $VO \cdot TJ$ pelo VPL. Assim, o CDC adaptado é obtido pela expressão 11.

$$CDC_{(Adaptado)} = P_{(mau)} \cdot VPL \quad (11)$$

onde $P_{(mau)}$ é a probabilidade de ser “bom pagador” obtida pelo modelo de risco de crédito multiplicado pelo VPL ponderado pelo WACC. Então, o CDN adaptado é expresso pela expressão 12.

$$CDN_{(Adaptado)} = P_{(bom)} \cdot VPL \quad (12)$$

onde $P_{(bom)}$ é a probabilidade de ser “bom pagador” multiplicado pelo VPL ponderado pelo WACC. Logo, o *escore* final da operação do cliente consiste na diferença entre o CDC e o CND, conforme demonstrado pela expressão 12.

$$ESCORE = CDN_{(Adaptado)} - CDC_{(Adaptado)} \quad (12)$$

No MCS proposto, quanto maior o *escore* maior é a chance de pagamento da operação, pois o $CDN_{(Adaptado)}$ será maior que o $CDC_{(Adaptado)}$. Portanto, o resultado do *escore* consiste na diferença da probabilidade de recebimento frente a probabilidade de inadimplência ponderados pelo VPL. Logo, o *escore* indica uma aprovação ou rejeição, ou seja, se o resultado for positivo

o crédito será aprovado, caso contrário, será negado. Note que o componente do custo de capital no financiamento das operações de venda à crédito emprega VPL calculado com base no WACC. Portanto, o MCS é sensível não somente às probabilidades de recebimento e de inadimplência como também em relação às mudanças na estrutura e no Custo de Capital.

6 EXPERIMENTAÇÕES E ANÁLISE DOS RESULTADOS

Esta seção expõe todas as experimentações e as análises dos resultados das aplicações do MCS. Além disso, descreve as contribuições práticas ao empresário com o uso do MCS.

6.1 Experimentações

As experimentações do MCS foram realizadas utilizando uma base de 1.013 pedidos de venda com cheques pré-datados no crediário para um de 183 clientes e abrangeu o período de janeiro a setembro de 2021. Todos os dados do problema referentes à consolidação das vendas por faixa de valor, os prazos de vencimento e o total de pedidos por cliente e são apresentados respectivamente pelas tabelas 12, 13, 14 e 15.

Tabela 12 – Distribuição de Pendências Cadastrais por Pedido

Atributos Cadastrais dos Clientes	Cheques Devolvidos		Restrições		Total Geral
	Consta	Nada consta	Não	Sim	
Pessoa Física	248	349	323	274	597
Até 2 mil	67	161	142	86	228
Até 5 mil	67	79	89	57	146
Até 10 mil	65	43	42	66	108
Até 20 mil	49	48	32	65	97
Maior que 20 mil		18	18		18
Pessoa Jurídica	255	161	307	109	416
Até 2 mil	57	35	53	39	92
Até 5 mil	48	37	46	39	85
Até 10 mil	67	14	63	18	81
Até 20 mil	64	19	73	10	83
Maior que 20 mil	19	56	72	3	75
Total Geral	503	510	630	383	1.013

Fonte: Elaborado pelo autor.

A tabela 12 apresenta a quantidade de pedidos, que apresentaram históricos de cheques devolvidos e existência de restrições, realizados pelos clientes do tipo Pessoa Física (PF) ou Pessoa Jurídica (PJ). É possível verificar que foram aprovados empréstimos para 274 pedidos de clientes “PF” com registros de restrições enquanto os históricos de cheques devolvidos totalizam 248. Já para os clientes “PJ” foram aprovadas vendas no crediário para 255 clientes com histórico de cheques sem fundos enquanto existência de restrições somam 109. Em relação ao número de pedidos por faixa de valor, a tabela 13 consolida dos dados de janeiro a setembro de 2021.

Tabela 13 – Distribuição de Pedidos por Faixa de Valor

Faixa de Valor	Quantidade de Pedidos	Percentual (%) de Pedidos
Até 2 mil	320	31,6%
Até 5 mil	231	22,8%
Até 10 mil	189	18,7%
Até 20 mil	180	17,8%
Maior que 20 mil	93	9,2%
Total Geral	1.013	100%

Fonte: Elaborado pelo autor.

A tabela 13 indica que os pedidos com valores de até 5 mil reais compreendem 54,4% do total de pedidos efetuados no período de janeiro até setembro de 2021. Já no que trata dos prazos de pagamento, 97,3% dos pedidos têm prazos de 1 mês, conforme demonstrado pela tabela 14.

Tabela 14 – Distribuição dos Prazos de Pagamentos dos Pedidos

Prazo dos Pedidos	Quantidade de Pedidos	Percentual (%) de Pedidos
1 mês	986	97,3%
2 meses	22	2,2%
3 meses	5	0,5%
Total Geral	1.013	100%

Fonte: Elaborado pelo autor.

Note que um aspecto de curto prazo para as operações de vendas no crediário também é evidenciado pela tabela 14. Todos os pedidos de vendas no crediário com prazos de pagamento de até dois meses (de janeiro até setembro de 2021), foram de 99,5%. A totalidade de clientes por faixa e quantidade de pedidos é apresentada pela tabela 15.

Tabela 15 – Totalidade de Clientes por Faixa e Quantidade de Pedidos

Total de Clientes	Faixa de Pedidos	Quantidade de Pedidos
76	1 pedido	76
84	de 2 até 10 pedidos	329
14	De 11 até 20 pedidos	214
2	De 21 até 30 pedidos	51
3	De 31 até 40 pedidos	107
1	De 41 até 50 pedidos	44
2	De 51 até 60 pedidos	102
1	90	90
183	1.013	1.013

Fonte: Elaborado pelo autor.

A tabela 15 também mostra que 107 clientes realizaram uma quantidade de pedidos que podem variar de 1 até 90. Isso ocorre devido a emissão de vários pedidos para obras de construção civil que são realizadas em etapas. Além disso, muitos pedidos para tais materiais são para obras realizadas por construtores parceiros. Por conseguinte, a construção e interface do MCS foi elaborada sob a plataforma de uma planilha do Microsoft Office Excel 2016. Assim, basta ao empresário apenas preencher os campos “Tipo”, “Valor e Prazo”, “Faixa de Valor”, “Restrição Cadastral” e “Cheque Devolvido” para que a planilha apresente o resultado calculado pelo MCS. A figura 18 ilustra a interface do MCS formulado na plataforma do Microsoft Office Excel 2016.

Figura 18 – Interface do Modelo de *Credit Scoring* Proposto

MODELO DE CREDIT SCORE			
Tipo	Restrição Cadastral	Cheque Devolvido	Faixa de Valor
PF	Sim	Consta	Maiores que
Tipo	Restrição Cadastral	Cheque Devolvido	Faixa de Valor
PF	Sim	Consta	Até 2 mil
Valor (Crédito)	Prazo (Dias)	Valor Presente (VP)	
R\$ 1.800,00	1	R\$ -	
Score do Cliente		Cálculo	
Mau Pagador	84,0%	R\$	-
Bom Pagador	16,0%	R\$	-
Resultado	R\$	-	APROVAR

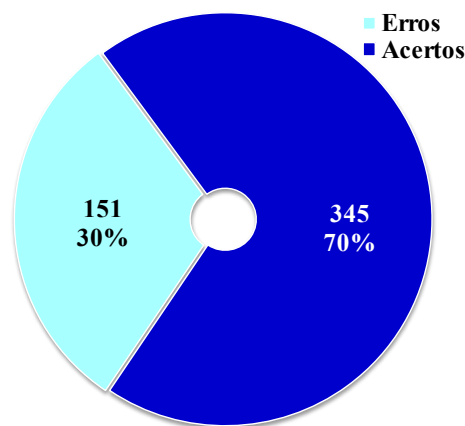
Fonte: Elaborado pelo autor.

A figura 18 demonstra que o MCS traz as probabilidades condicionais do cliente como “bom pagador” e “mau pagador”. As bases de dados de treinamento estão em planilhas ocultas no mesmo arquivo e permitem que sejam inseridas informações atualizadas para cada execução do MCS. Estas atualizações possibilitam uma melhor adaptação às novas realidades e contextos de risco da empresa, a serem realizadas por meio de simulações do MCS. A possibilidade de realizar diferentes simulações para diversas propostas de vendas no crediário com cheque pré-datado demonstram a flexibilidade do MCS.

6.2 Análise dos Resultados

Esta seção apresenta as análises dos resultados das experimentações realizadas para avaliação do MCS. As experimentações nos 1.013 pedidos do período de janeiro a setembro de 2021 demonstram que 496 pedidos apresentaram resultado positivo com a aplicação do MCS. Dentre os pedidos que seriam aprovados, 151 clientes ficaram inadimplentes após a aplicação do MCS. A mensuração dos percentuais de erros e acertos referente aos clientes que tiveram o crédito aprovado é ilustrada pelo gráfico da figura 19.

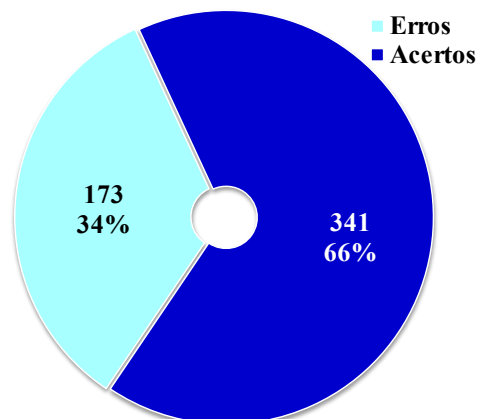
Figura 19 – Erros e Acertos de Pedidos com Crédito Aprovado



Fonte: Elaborado pelo autor.

A figura 19 demonstra que, do total de pedidos que tiveram o crédito aprovado, as quantidades inadimplentes são na ordem de 30% (151). Esta quantidade que corresponde a uma perda, em milhares de reais, de R\$ 171,25 mil. Já em relação aos 514 pedidos que seriam negados por apresentar *escore* negativo, o gráfico da figura 20 evidencia que os pedidos que teriam clientes inadimplentes totalizam 341.

Figura 20 – Erros e Acertos de Pedidos com Crédito Reprovado



Fonte: Elaborado pelo autor.

A inadimplência das vendas que tiveram *escore* negativo resultou, em milhares de reais, no valor R\$ 205,39. Nessa lógica, é possível obter uma economia com perdas por inadimplência, já que, os resultados de crédito negado foram validados pelas experimentações do MCS. A demonstração dos resultados, oriundos dos erros e acertos do MCS é exibida pela tabela 16.

Tabela 16 – Resultados do Modelo de *Credit Scoring* Proposto

Resultado das Vendas	Resultado do Crédito	Credciário	PCLD	Resultado
Positivo (> 0)	Crédito Aprovado	R\$ 425,46	R\$ 171,25	R\$ 254,21
Negativo (< 0)	Crédito Negado	R\$ 427,29	R\$ 205,39	-R\$ 221,91
Resultado Total	-	-	-	R\$ 32,30

Fonte: Elaborado pelo autor. Referências às abreviações: PCLD – Provisão com Créditos de Liquidação Duvidosa

A demonstração dos resultados exibidos pela tabela 16, em milhares de reais, é de R\$ 32,30. Este valor irá, portanto, impactar no valor da Provisão com Créditos de Liquidação Duvidosa (PCLD). A tabela 17 exhibe, em milhares de reais, os valores finais do Demonstrativo do Resultado do Exercício (DRE) ajustado aos resultados das experimentações do MCS.

Tabela 17 – Demonstrativo de Resultado Ajustado ao Modelo de *Credit Scoring*

Demonstrativo de Resultado	2021
Receita Bruta	R\$ 5.116,44
(-) Deduções e abatimentos	-R\$ 90,00
(=) Receita Líquida	R\$ 5.026,44
(-) Custo das Mercadorias Vendidas	-R\$ 3.982,61
(=) Lucro Bruto	R\$ 1.043,83
(-) Despesas com Vendas	-R\$ 110,84
(-) Despesas Administrativas	-R\$ 434,40
(-) Despesas Financeiras	-R\$ 126,74
(-) Despesas com PCLD	-R\$ 171,25
(=) Resultado Líquido	R\$ 200,60

Fonte: Elaborado pelo autor. Referências às abreviações: PCLD – Provisão com Créditos de Liquidação Duvidosa

Analisando o DRE ajustado verifica-se que, em razão da reprovação dos pedidos de venda no crediário, a aplicação do MCS leva a uma queda na Receita Bruta. Entretanto, há uma melhora da gestão do risco de crédito, visto que, houve um incremento no resultado financeiro devido a reprovação das vendas inadimplentes pelo MCS. As vendas negadas no crediário totalizam R\$ 427 mil e foram subtraídas da Receita Bruta. Já o valor da Custo das Mercadorias Vendidas foi regularizado proporcionalmente ao valor das mercadorias vendidas.

Na sequência, foi ajustado o valor da Provisão com Créditos de Liquidação Duvidosa com lançamento do valor atualizado, em milhares de reais, de R\$ 171,25 mil. A tabela 15 ratifica os impactos positivos no resultado da empresa, no qual o melhor controle do risco de crédito resultou em um Lucro Líquido de R\$ 200,60. Este valor corresponde a um aumento percentual de 124,2% no Lucro Líquido se comparado ao DRE sem aplicação do MCS. Estes resultados reforçam o foco na relação entre risco e retorno na realização dos negócios que implicam na análise do *Credit Scoring*.

Esta pesquisa demonstrou que o algoritmo probabilístico *Naive Bayes* é eficaz para avaliação de risco quando combinado com as ferramentas da Moderna Teoria das Finanças. Além disso, esforços apenas na mensuração do risco de vendas à crédito não foi o foco dado a construção e aplicação do MCS. A pesquisa se estendeu às decisões de investimento e financiamento em conjunto com os retornos financeiros gerados com a aplicação do MCS. É possível ao empresário realizar diferentes simulações com dados reais apenas alimentando as planilhas do MCS. Em termos gerais, a apresentação de resultados consistentes e que agilize a tomada de decisão do sócio administrador no processo de conceder ou não crédito a um cliente é o objetivo primeiro do MCS.

7 CONSIDERAÇÕES FINAIS

Esta dissertação apresentou um Modelo de *Credit Scoring* (MCS) para uma Micro e Pequena Empresa (MPE) do ramo de comércio de materiais de construção do interior do estado de São Paulo (SP). A pesquisa é aplicada e de natureza quantitativa, e caracterizada como um estudo de caso em uma MPE. A construção do MCS foi ponderada por decisões de investimento e financiamento e o cálculo da probabilidade de inadimplência e adimplência foi fundamentado no algoritmo *Naive Bayes*. A ponderação foi baseada nas decisões de investimentos e financiamento da Moderna Teoria de Finanças e empregou o Custo de Conceder (CDC) e o Custo de Não conceder (CDN). Tais custos foram alicerçados no *Capital Asset Pricing Model* (CAPM), *Weighted Average Cost of Capital* (WACC) e o Valor Presente Líquido (VPL). Logo, o MCS foi adaptado à realidade em questão e desenvolvido (entradas/saídas de dados) por meio de planilhas eletrônicas da plataforma do Microsoft Office Excel 2016. Experimentos utilizando os bases dados das vendas no crediário do ano de 2021 demonstraram um resultado financeiro positivo de R\$ 32,20 mil com a aplicação do MCS. Quando avaliado o impacto da aplicação do MCS no DRE da empresa, o lucro líquido foi maior em 124,2%. No caso dos clientes com crédito aprovado, o MCS apresentou acertos de 70% e para os clientes com crédito negado os acertos foram de 66%. As análises dos resultados demonstraram que o MCS é capaz de ponderar o risco nas decisões de investimento e financiamento na concessão de vendas a crédito da referida MPE. As principais contribuições desta pesquisa estão em prover um MCS de baixo custo, de fácil manuseio e implementação no Microsoft Office Excel 2016. Além disso, o MCS proposto traz a relação entre risco e retorno nos negócios das MPEs agrupando o *Credit Scoring*, a Teoria das Finanças e ao algoritmo *Naive Bayes*. A pesquisa também trouxe importantes compreensões ao empresário sobre o risco de inadimplência nas vendas no crediário e as decisões investimento e financiamento abalizados na Moderna Teoria de Finanças. Além disso, esta pesquisa também demonstrou que existe uma lacuna na aplicação de MCSs que focam na lucratividade das MPEs. Esta pesquisa é um passo inicial na aplicação de MCSs aos processos de concessão de crédito realizados por MPE. Interessante sugestão para futuras pesquisas é o desenvolvimento e comparação de MCSs com foco nos diferentes tipos, portes e peculiaridades das MPEs.

REFERÊNCIAS

- ABDOU, H. A.; POINTON, J. Credit scoring, statistical techniques and evaluation criteria: a review of the literature. **Intelligent systems in accounting, finance and management**, v. 176, n. January, p. 161–176, 2011.
- ABREU FILHO, J. C. **Finanças Corporativas**. 8. ed. Rio de Janeiro: Editora FGV, 2007.
- ALTMAN, E. I. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. **The Journal of Finance**, v. 29, n. 1, p. 312–312, 1974.
- AMORIM, G. Concessão de crédito e receita financeira: uma ferramenta de análise econômico-gerecncial. **Revista de Empreendedorismo e Gestão de Pequenas Empresas**, v. 8, n.2 p. 410–424, 2019.
- ANDRIOSOPOULOS, D.; DOUMPOS, M.; PARDALOS, P. M.; ZOPOUNIDIS, C. Computational approaches and data analytics in financial services: A literature review. **Journal of the Operational Research Society**, v. 70, n. 10, p. 1581–1599, 2019.
- ANTONAKIS, A. C.; SFAKIANAKIS, M. E. Assessing naïve Bayes as a method for screening credit applicants. **Journal of Applied Statistics**, v. 36, n. 5, p. 537–545, 2009.
- BAESENS, B.; SETIONO, R.; MUES, C.; VANTHIENEN, J. Using neural network rule extraction and decision tables for credit-risk evaluation. **Management Science**, v. 49, n. 3, p. 312–329, 2003.
- BASTANI, K.; ASGARI, E.; NAMAVARI, H. Wide and deep learning for peer-to-peer lending. **Expert Systems with Applications**, v. 134, p. 209–224, 2019.
- BERNASTEIN, P. L. A Primer on Money, Banking, and Gold. New Jersey: John Wiley & Sons, 2008.
- BERTRAND, J.W. M.; FRANSOO, J. C. Modelling and simulation: Operations management research methodologies using quantitative modeling. **International Journal of Operations and Production Management**, 22(2): 241–264, 2002.
- BRAVO, C.; MALDONADO, S.; WEBER, R. Granting and managing loans for micro-entrepreneurs: New developments and practical experiences. **European Journal of Operational Research**, v. 227, n. 2, p. 358–366, 2013.
- BRITO, G. A. S.; ASSAF NETO, A.; Modelo de Classificação de Risco de Crédito de Empresas. **Revista de Contabilidade e Finanças - USP**, v. 19, n. 46, p. 18-29, 2008X’.
- CAPON, N. Credit scoring systems: a critical analysis. **Journal of Marketing**, v. 46, n. Spring, p. 82–91, 1982.
- CASTRO JUNIOR, F. H. F.; FAMÁ, R. As novas finanças e a teoria comportamental no contexto da tomada de decisão sobre investimentos. **CadeRno de Pesquisa em Administração**. São Paulo, v. 09, n.2, abril/julho 2002.
- CHEN, N.; RIBEIRO, B.; CHEN, A. Financial credit risk assessment: a recent review. **Artificial Intelligence Review**, v. 45, n. 1, p. 1–23, 2016.

ÇİĞŞAR, B.; ÜNAL, D. Comparison of Data Mining Classification Algorithms Determining the Default Risk. **Scientific Programming**, v. 2019, 2019.

CNDL/SPC Brasil. Release Uso do Credito v.5, n.1, julho de 2019, Disponível em: <https://www.spcbrasil.org.br/wpimprensa/wp-content/uploads/2019/07/release_uso_do_credito_crediariorio_julho_2019_V5-1.pdf>. Acesso em: 07 dez. 2020.

DONTHU, N; KUMAR, S; MUKHERJEE, D; PANDEY, N; MARCLIM, W. How to conduct a bibliometric analysis: An overview and guidelines. **Journal of Business Research**, v.133, p. 285–296, 2021.

DOUMPOS, M.; LEMONAKIS, C.; NIKLIS, D.; ZOPOUNIDIS, C. **Analytical Techniques in the Assessment of Credit Risk: An Overview of Methodologies and Applications**. Switzerland: Springer, 2019.

EISENBEIS, R. A. Problems in applying discriminant analysis in credit scoring models. **Journal of Banking and Finance**, v. 2, n. 3, p. 205–219, 1978.

FIGUEIREDO, N. Da importância dos artigos de revisão da literatura. **Revista Brasileira de Biblioteconomia e Documentação**, 23(1), 131-135, 1990.

FINLAY, S. Are we modelling the right thing? The impact of incorrect problem specification in credit scoring. **Expert Systems with Applications**, v. 36, n. 5, p. 9065–9071, 2009.

FINLAY, S. Credit scoring for profitability objectives. **European Journal of Operational Research**, v. 202, n. 2, p. 528–537, 2010.

HAEGEMAN, K.; MARINELLI, E.; SCAPOLLO, F.; RICCI, A.; SOKOLOV, A. Quantitative and qualitative approaches in Future-oriented Technology Analysis (FTA): From combination to integration? **Technological Forecasting and Social Change**, v. 80, n. 3, p. 386–397, 2013.

HILL, T. and LEWICKI, P. **Statistics: Methods and Applications**. Tulsa: StatSoft, 2007.

INCE, H.; AKTAN, B. A comparison of data mining techniques for credit scoring in banking: A managerial perspective. **Journal of Business Economics and Management**, v. 10, n. 3, p. 233–240, 2009.

KOZENY, V. Genetic algorithms for credit scoring: Alternative fitness function performance comparison. **Expert Systems with Applications**, v. 42, n. 6, p. 2998–3004, 2015.

KOZODOI, N.; LESSMANN, S.; PAPAKONSTANTINO, K; GATSOULIS, Y. A multi-objective approach for profit-driven feature selection in credit scoring. **Decision Support Systems**, v. 120, n. January, p. 106–117, 2019.

KRICHENE, A. Using a naive Bayesian classifier methodology for loan risk assessment: Evidence from a Tunisian commercial bank. **Journal of Economics, Finance and Administrative Science**, v. 22, n. 42, p. 3–24, 2017.

KRUPPA, J.; SCHWARZ, A.; ARMINGER, G.; ZIEGLER, A. Consumer credit risk: Individual probability estimates using machine learning. **Expert Systems with Applications**, v. 40, n. 13, p. 5125–5131, 2013.

- LEONARD, K. J. Credit-scoring models for the evaluation of small-business loan applications. **IMA Journal of Management Mathematics**, v. 4, n. 1, p. 89–95, 1992.
- LESSMANN, S.; BAESENS, B.; SEOW, H.; THOMAS, L. C. Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. **European Journal of Operational Research**, v. 247, n. 1, p. 124–136, 2015.
- LI, Y.; CHEN, W. A comparative performance assessment of ensemble learning for credit scoring. **Mathematics**, v. 8, n. 10, p. 1–19, 2020.
- LINNENLUECKE, M. K.; MARRONE, M.; SINGH, A. K. Conducting systematic literature reviews and bibliometric analyses. **Australian Journal of Management**. v. 45, n. 2, 175–194, 2020.
- LIU, J.; BO, S. Naive Bayesian classifier based on genetic simulated annealing algorithm. **Procedia Engineering**, v. 23, p. 504–509, 2011.
- LOUZADA, F.; ARA, A.; FERNANDES, G. B. O impacto da estratégia de crédito na liquidez e rentabilidade: uma análise das lojas de departamento do setor de comércio varejista. **Surveys in Operations Research and Management Science**, v. 21, n. 2, p. 117–134, 2016.
- LOBATO, F. T. R.; SILVA, A. P.; RIBEIRO, K. C. S. Classification methods applied to credit scoring: Systematic review and overall comparison. **Revista Contemporânea de Contabilidade**, v. 1, n. 12, p. 11–34, 2009.
- MALDONADO, S.; BRAVO, C.; LOPES, J.; PEREZ, J. Integrated framework for profit-based feature selection and SVM classification in credit scoring. **Decision Support Systems**, v. 104, p. 113–121, 2017.
- MAIA, V. I.; PEREIRA, E. M.; SILVA, F. H.; COELHO, J. D.; DIAS, L.A. C. Gestão financeira de MPÉs: o setor varejista na região de Pará de Minas. **SynThesis Revista Digital FAPAM**, v.1, n.1 p. 261–273, 2009.
- MARQUÉS, A. I.; GARCÍA, V.; SÁNCHEZ, J. S. A literature review on the application of evolutionary computing to credit scoring. **Journal of the Operational Research Society**, v. 64, n. 9, p. 1384–1399, 2013.
- MIGUEL, P. A. C. Estudo de caso na engenharia de produção: estruturação e recomendações para sua condução. **Produção**, v. 17, n. 1, p. 216–229, 2007.
- NALIĆ, J.; MARTINOVIC, G. Building a Credit Scoring Model Based on Data Mining Approaches. **International Journal of Software Engineering and Knowledge Engineering**, v. 30, n. 2, p. 147–169, 2020.
- ORGLER, Y. E. A Credit Scoring Model for Commercial Loans. **Journal of Money, Credit and Banking**, v. 2, n. 4, p. 435–445, 2014.
- RENDE, R. **Um plano de cadastro, crédito e cobrança**. Belo Horizonte: SEBRAE/MG, 2013.
- ŘEZÁČ, M. ESIS2: Information Value Estimator for Credit Scoring Models. **Computational Economics**, v. 45, n. 2, p. 303–322, 2014.

ROSS, S. A.; WESTERFIELD, R. W.; JAFFE, J. F. **Administração financeira**. São Paulo: Editora Atlas, 1995.

SAMANEZ, C. P. **Engenharia Econômica**. 1. ed. São Paulo: Pearson Prentice Hall, 2009.

SANTOS, K. F.; CARMONA, C. U. M.; LEISMANN, E. L. Gerenciamento do risco de crédito em MPÉs: recortes teórico-empíricos. **Revista da Micro e Pequena Empresa**, v. 3, n. 3, p. 83–95, 2010.

SARIANNIDIS, N.; PAPADAKIS, S.; GAREFALAKIS, A.; LEMONAKIS, C.; KYRIAKI-ARGYRO, T. Default avoidance on credit card portfolios using accounting, demographical and exploratory factors: decision making based on machine learning (ML) techniques. **Annals of Operations Research**, n. 0123456789, 2019.

SEBRAE. Pequenos negócios na economia brasileira. Disponível em: <<https://www.sebrae.com.br/sites/PortalSebrae/ufs/mt/noticias/micro-e-pequenas-empresas-geram-27-do-pib-do-brasil.aspx>> Acesso em 30 Maio 2021.

SEBRAE. Pequenos negócios na economia brasileira. Disponível em: <<https://www.sebrae.com.br/sites/PortalSebrae/ufs/mg/sebraeaz/planilhas-de-gestao-financeira,5d4049fe44fe6610VgnVCM1000004c00210aRCRD>> Acesso em 30 Maio 2021.

SECURATO, J. R. **Credito: análise e avaliação do risco: pessoas físicas e jurídicas**. 2. ed. São Paulo: Saint Paul Editora, 2012.

SERRANO-CINCA, C.; GUTIÉRREZ-NIETO, B. The use of profit scoring as an alternative to credit scoring systems in peer-to-peer (P2P) lending. **Decision Support Systems**, v. 89, p. 113–122, 2016.

SICSÚ, A. L. **Credit Scoring: desenvolvimento, implantação, acompanhamento**. 1. ed. São Paulo: Blucher, 2010.

SILVA, J. P. **Gestão e análise de risco de crédito**. 9. ed. São Paulo: Cengage Learning, 2016.

SINHA, A. P.; ZHAO, H. Incorporating domain knowledge into data mining classifiers: An application in indirect lending. **Decision Support Systems**, v. 46, n. 1, p. 287–299, 2008.

ŠUŠTERŠIČ, M.; MRAMOR, D.; ZUPAN, J. Consumer credit scoring models with limited data. **Expert Systems with Applications**, v. 36, n. 3 PART 1, p. 4736–4744, 2009.

TRIVEDI, S. K. A study on credit scoring modeling with different feature selection and machine learning approaches. **Technology in Society**, v. 63, n. September, p. 101413, 2020.

VERBRAKEN, T.; BRAVO, C.; WEBER, R.; BAESSENS, B. Development and application of consumer credit scoring models using profit-based classification measures. **European Journal of Operational Research**, v. 238, n. 2, p. 505–513, 2014.

VUKOVIC, S.; DELIBASIC, B.; UZELAC, A.; SUKNOVIC, M. A case-based reasoning model that uses preference theory functions for credit scoring. **Expert Systems with Applications**, v. 39, n. 9, p. 8389–8395, 2012.

WU, X. Top 10 algorithms in data mining. **Knowl Inf Syst** 14, 1–37, 2008.