



Universidade de Brasília
Programa de Pós-Graduação em Administração
Finanças e Métodos Quantitativos

**Estudo Comparativo entre Técnicas de
Aprendizado de Máquina para
Estimação de Risco de Crédito**

Maísa Cardoso Aniceto

140071598

Brasília

2016

Maísa Cardoso Aniceto

140071598

**Estudo Comparativo entre Técnicas de
Aprendizado de Máquina para
Estimação de Risco de Crédito**

Dissertação Apresentada no Curso de Mestrado Acadêmico do Programa de Pós-Graduação em Administração, na Área de Concentração de Finanças e Métodos Quantitativos, da Universidade de Brasília como Parte do Requisito Necessário para o Grau de Mestre.

Orientador: Prof. Dr. Herbert Kimura

Brasília

Março de 2016

Resumo

A avaliação do risco de crédito tem um papel relevante para as instituições financeiras por estar associada a possíveis perdas e gerar grande impacto nas atividades do mercado financeiro. Embora existam várias pesquisas sobre risco de crédito e aprendizado de máquina ainda falta um estudo que integra e sistematiza o conhecimento disponível. Assim, o objetivo deste trabalho é apresentar uma revisão sistemática de literatura sobre aprendizado de máquina em risco de crédito. Os principais artigos dessa área foram classificados e codificados e uma agenda de pesquisa e recomendações são apresentadas para auxiliar na realização de pesquisas futuras. A segunda parte da dissertação envolve uma análise empírica de dados de uma instituição financeira e aplicação de algoritmos de Support Vector Machine e Decision Tree, além de Bagging, AdaBoost e Random Forest. Os resultados foram comparados por meio da Curva ROC, Sensibilidade, Especificidade e AUC.

Palavras-chaves: Aprendizado de Máquina, Risco de Crédito, SVM, Árvore de decisão, Classificadores Ensemble.

Abstract

Despite several pieces of research on credit risk and machine learning, a study that integrates and systematizes available knowledge is still lacking. Credit risk evaluation has a relevant role to financial institutions for it is connected to possible losses. Also, it can have a great impact in the financial market. Thus, the aim of this study is to carry out a systematic literature review on credit risk machine learning. The main articles of this field have been classified and coded. Recommendations are presented to assist future research projects. Apart from that, this thesis focuses on an empirical data analysis of a financial institution, as well as the application of Support Vector Machine and Decision Tree algorithms; Bagging, AdaBoost and Random Forest, ensemble classifiers. The results were measured on the use of the ROC Curve, Sensitivity, Specificity and AUC.

Keywords: Machine Learning, Credit Risk, Support Vector Machine, Decision Tree, Classifier Ensemble.

Lista de ilustrações

Figura 1 – Classificações do risco.	17
Figura 2 – Artigos publicados por ano.	27
Figura 3 – Países com o maior número de pesquisadores.	27
Figura 4 – Word Cloud das palavras-chave dos artigos.	30
Figura 5 – Técnicas de aprendizado de máquina mais utilizadas.	31
Figura 6 – Rede de Autores.	35
Figura 7 – Curva ROC - Decision Tree C5.0.	61
Figura 8 – Erro X Número de árvores - Random Forest.	62
Figura 9 – Curva ROC - Random Forest.	63
Figura 10 – Curva ROC - SVM Linear.	65
Figura 11 – Curva ROC - SVM Radial.	66
Figura 12 – Curva ROC - Bagging.	68
Figura 13 – Curva ROC - AdaBoost.	69
Figura 14 – Curva ROC - Regressão Logística.	70
Figura 15 – Especificidade x Sensibilidade.	72
Figura 16 – Especificidade x Sensibilidade.	73
Figura 17 – Comparação das medidas de desempenho em diferentes amostras para Decision Tree.	74
Figura 18 – Comparação das medidas de desempenho em diferentes amostras para Random Forest.	74
Figura 19 – Comparação das medidas de desempenho em diferentes amostras para Bagging.	75

Figura 20 – Comparação das medidas de desempenho em diferentes amostras para Ada-Boost.	75
Figura 21 – Comparação das medidas de desempenho em diferentes amostras para SVM Linear.	76
Figura 22 – Comparação das medidas de desempenho em diferentes amostras para SVM Radial.	76

Lista de tabelas

Tabela 1 – Quadro de classificação e codificação dos estudos analisados.	25
Tabela 2 – Distribuição dos artigos nos periódicos.	28
Tabela 3 – Palavras-Chaves mais citadas.	29
Tabela 4 – Escopo dos artigos. <i>A: Application of a traditional technique. B: Enhancement on a technique. D: Development of a new technique. E: Comparison of techniques. F: Other.</i>	33
Tabela 5 – Número de citações no <i>Google Scholar</i> dos principais artigos em 18/10/2015.	35
Tabela 6 – Número de citações dos artigos mais recentes no <i>Google Scholar</i> dos principais artigos em 18/10/2015.	36
Tabela 7 – Variáveis utilizadas na análise.	50
Tabela 8 – Matriz de confusão para a pontuação de crédito.	59
Tabela 9 – Matriz de Confusão - Decision Tree.	60
Tabela 10 – Medidas de Desempenho - Decision Tree.	61
Tabela 11 – Matriz de Confusão - Random Forest.	62
Tabela 12 – Medidas de Desempenho - Random Forest.	63
Tabela 13 – Matriz de Confusão - SVM Linear.	64
Tabela 14 – Medidas de Desempenho - SVM Linear.	64
Tabela 15 – Matriz de Confusão - SVM Radial.	65
Tabela 16 – Medidas de Desempenho - SVM Radial.	66
Tabela 17 – Matriz de Confusão - Bagging.	67
Tabela 18 – Medidas de Desempenho - Bagging.	67

Tabela 19 – Matriz de Confusão - AdaBoost.	67
Tabela 20 – Medidas de Desempenho - AdaBoost.	68
Tabela 21 – Matriz de Confusão - Regressão Logística.	69
Tabela 22 – Medidas de Desempenho - Regressão Logística.	70
Tabela 23 – Resultados AUC.	71
Tabela 24 – Comparação das Medidas de Desempenho.	71
Tabela 25 – Comparação das medidas de desempenho em diferentes amostras.	77
Tabela 26 – Classificação e codificação dos estudos analisados.	90
Tabela 27 – Breve descrição dos objetivos e resultados de cada artigo analisado.	91

Lista de Abreviaturas

Abreviatura	Significado
AI	Artificial Intelligence
ANN	Artificial Neural Network
BNN	Backpropagation Neural Network
BP	Backpropagation
CART	Classification And Regression Tree
DT	Decision Tree
ELM	Extreme Learning Machine
GA	Genetic Algorithm
GP	Genetic Programming
I-ELM	Incremental Extreme Learning Machine
ML	Machine Learning
MLP	Multilayer Perceptron
SLFN	Single Hidden Layer Feedforward Networks
RF	Random Forest
RPA	Algoritmo de Particionamento Recursivo
SVM	Support Vector Machine
VBDTM	Vertical Bagging Decision Tree Model

Sumário

1	INTRODUÇÃO	13
1.1	Objetivo e Estrutura da Dissertação	15
2	REFERENCIAL TEÓRICO	17
2.1	Risco de Mercado	18
2.2	Risco de Liquidez	18
2.3	Risco Operacional	19
2.4	Risco de Crédito	19
3	REVISÃO DA LITERATURA	22
3.1	Classificação e Codificação	23
3.2	Análise Descritiva	26
3.3	Técnicas de Pesquisa	30
3.4	Fonte dos Dados	31
3.5	Tipos de Estudo	33
3.6	Principais Autores e Citações	34
4	APRENDIZADO DE MÁQUINA NA AVALIAÇÃO DE RISCO DE CRÉ- DITO	37
4.1	Artificial Neural Networks	40
4.2	Decision Trees	42
4.3	Support Vector Machines	43

4.4	Classificadores <i>Ensemble</i>	45
5	ESTUDO EMPÍRICO	48
5.1	Base de Dados	48
5.1.1	Seleção das Amostras	50
5.2	Metodologia	51
5.2.1	Decision Trees	51
5.2.2	Support Vector Machines	53
5.2.3	Bagging	55
5.2.4	Boosting e AdaBoost	56
5.2.5	Procedimento	58
5.3	Resultados	59
5.3.1	Decision Trees	59
5.3.1.1	Random Forest	61
5.3.2	SVM	63
5.3.2.1	Kernel Linear	64
5.3.2.2	Kernel Radial	65
5.3.3	Bagging	66
5.3.4	AdaBoost	67
5.3.5	Regressão Logística	69
5.4	Resultados e Discussão	70
5.4.1	Diferentes Tamanhos de Amostra	72
6	CONSIDERAÇÕES FINAIS	78

Referências	81
APÊNDICES	88
APÊNDICE A – CLASSIFICAÇÃO E CODIFICAÇÃO DOS ESTUDOS ANALISADOS	89
APÊNDICE B – BREVE DESCRIÇÃO DOS OBJETIVOS E RESULTADOS DE CADA ARTIGO ANALISADO	91

1 Introdução

A busca por crédito no Brasil aumentou 1% em 2015 comparado a 2014, segundo o *Indicador Serasa Experian da Demanda do Consumidor por Crédito* divulgado no dia 18 de janeiro de 2016. De acordo com a mesma pesquisa, aproximadamente 59 milhões de brasileiros começaram o ano como inadimplentes. A inadimplência ficou em 6,2% em janeiro de 2016 segundo dados divulgados pelo Banco Central no dia 24 de fevereiro de 2016.

Com o aumento do endividamento e da procura por empréstimos bancários, pode haver um crescimento na concessão de crédito. Ao oferecer crédito, geralmente as empresas se baseiam em uma análise para avaliar se o possível tomador merece ou não o crédito. A quantidade de informação necessária para a análise exigida aumenta de acordo com o volume de crédito oferecido para o cliente (DAMODARAN, 2010).

Toda concessão de crédito está relacionada a um risco. O risco de crédito é o risco da contraparte não ser capaz ou ser relutante em fazer o pagamento de suas obrigações contratuais (GREGORY, 2012). Esse risco pode ser caracterizado em termos de um *default* real ou, menos severamente, pela deterioração da qualidade de crédito da contraparte. No primeiro caso, pode resultar em uma perda real e imediata, no último caso, as perdas futuras tornam-se mais prováveis (GREGORY, 2012).

Durante muitos anos as instituições financeiras se baseavam somente numa análise subjetiva para avaliar a concessão de crédito, segundo Altman e Saunders (1998), somente após a década de 70 as técnicas estatísticas começaram a ser utilizadas por apresentarem maior precisão e conseqüentemente menores perdas. Assim, os modelos estatísticos começaram a ser uma

ferramenta importante para ajudar as instituições financeiras a tomarem decisões corretas.

Segundo [Altman e Saunders \(1998\)](#), nesse período as técnicas estatísticas dominantes eram a Regressão Logística e a Análise Discriminante. Porém, a maioria dos modelos estatísticos de classificação tem uma performance favorável apenas quando os pressupostos essenciais de identidade distribucional, normalidade dos erros e homoscedasticidade são satisfeitos.

Para superar os métodos estatísticos, na década de 90, tecnologias de inteligência artificial começaram a ser estudadas para uso em risco de crédito. Mais especificamente o Aprendizado de Máquina, um sub-campo da Inteligência Artificial, que foi definido por [Mitchell \(1997\)](#) como ‘Um programa de computador *aprende* com a experiência E com relação a alguma classe de tarefas T e medida de desempenho P , se o seu desempenho em tarefas em T , medida pelo P , melhora com a experiência E ’.

Artificial Neural Networks (ANN) e *Support Vector Machine (SVM)* são consideradas as duas abordagens de ML mais utilizadas ([ZHONG et al., 2014](#)). Ao contrario das técnicas estatísticas tradicionais, os métodos de aprendizado de máquina não exigem o conhecimento das relações entre as variáveis de entrada e de saída dos modelos. Estudos recentes tem mostrado que métodos de classificadores *ensemble* possuem performance melhor do que técnicas de inteligência artificial sozinhas ([WANG et al., 2011](#); [TSAI](#); [HSU](#); [YEN, 2014](#)).

Por envolver grandes quantias de dinheiro, o problema de classificação no contexto da concessão de crédito é importante e deve ser cada vez mais estudado. Atualmente, toda melhoria sobre os métodos já utilizados gera uma enorme economia para as instituições envolvidas.

1.1 Objetivo e Estrutura da Dissertação

Esta dissertação foi elaborada no formato de dois estudos. O primeiro estudo refere-se a uma revisão de literatura envolvendo análise de risco de crédito e técnicas de aprendizado de máquina. Foram analisados 80 artigos publicados em periódicos importantes. Os artigos foram classificados e codificados, tendo sido identificado o estado-da-arte sobre o tema, incluindo os artigos mais citados, os temas mais estudados e mais relevantes.

O segundo estudo envolve uma análise empírica, com a aplicação prática das técnicas de aprendizado de máquina mais utilizadas em um banco de dados real de operações de crédito.

Este trabalho é relevante devido a importância da análise de risco de crédito, para as empresas, principalmente as financeiras, e a falta de uma revisão e estudo unindo aprendizado de máquina e risco de crédito no Brasil. Dessa forma, o objetivo principal consiste em realizar uma revisão sistemática de literatura envolvendo análise de risco de crédito e técnicas de aprendizado de máquina e analisar o desempenho de quatro algoritmos de aprendizado de máquina *Decision Trees*, *Support Vector Machine*, *Bagging* e *AdaBoost*, no campo de classificação de devedores. Em termos de precisão média, sensibilidade e especificidade em um conjunto de dados de um banco brasileiro de grande porte. Tendo como objetivos secundários:

- Desenvolver cinco modelos de *credit scoring* mediante o uso de técnicas computacionais, *Decision Trees* e *Support Vector Machine*, *Random Forest*, *Bagging* e *AdaBoost* a produtos de crédito.
- Comparar os modelos desenvolvidos em termos de indicadores de qualidade e previsão.
- Comparar os modelos desenvolvidos com um modelo estatístico de Regressão Logística.
- Propor um modelo para classificação de clientes.

O referencial teórico sobre risco financeiro, especialmente sobre risco de crédito, está contido no capítulo dois. O capítulo três apresenta a revisão de sistemática de literatura realizada. No capítulo quatro é traçada a evolução das principais técnicas e classificadores de aprendizado de máquina utilizadas na avaliação do risco de crédito. O estudo empírico é apresentado no capítulo cinco. As conclusões e sugestões de trabalhos futuros são discutidos no capítulo seis.

2 Referencial Teórico

Segundo [Jorion \(2007\)](#) risco pode ser definido como a volatilidade de resultados inesperados. Empresas são expostas a vários tipos de riscos que se enquadram em duas categorias, o *business* risco e o risco financeiro. O risco financeiro está associado a possíveis perdas decorrentes das atividades do mercado financeiro.

As empresas precisam gerenciar o risco com cuidado, isto pode ser conseguido passivamente de forma simples tentando evitar a exposição a fatores de risco que podem ser potencialmente prejudiciais ([GREGORY, 2012](#)). As firmas podem também ter habilidade de entender os riscos e assumir a exposição a áreas específicas oferecendo uma forte vantagem competitiva.

O risco financeiro pode ser classificado em quatro categorias, [Figura 1](#): Risco de mercado, risco de liquidez, risco operacional e risco de crédito ([JORION, 2007](#)).

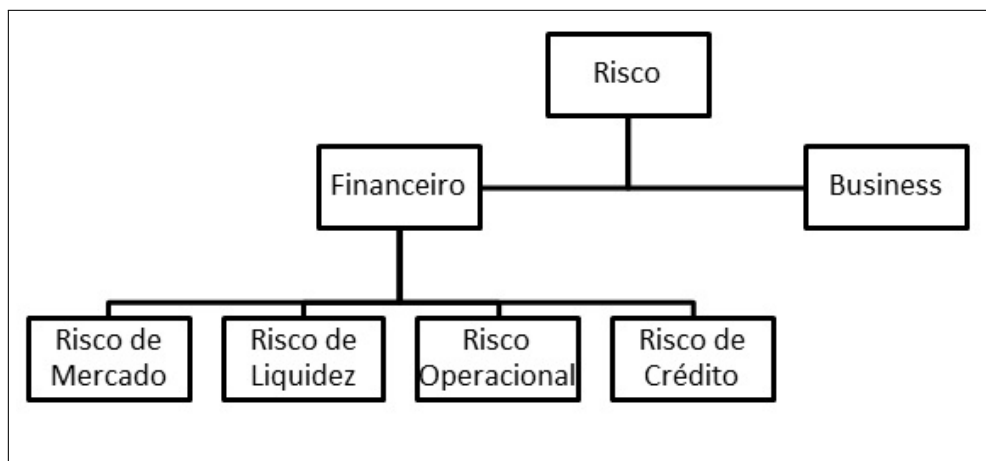


Figura 1 – Classificações do risco.

As quatro categorias de classificação serão agora brevemente explicadas, dando ênfase ao risco de crédito, que será o risco abordado nesta dissertação.

2.1 Risco de Mercado

O risco de mercado surge quando uma transação não pode ser conduzida a preços de mercado predominantes, devido ao tamanho da posição relativa dos lotes de negociação normais (JORION, 2007). Ele pode ser um risco linear, decorrente da exposição a variáveis subjacentes, como os preços das ações, taxas de juros, taxas de câmbio estrangeiras, preços das *commodities* ou *spreads* de crédito. Ou pode ser não linear surgindo da exposição a volatilidade do mercado. O risco de mercado tem sido o risco financeiro mais estudado nas últimas duas décadas (GREGORY, 2012). O acordo de Basileia 1, em 1995, permitiu que as instituições financeiras usassem modelos matemáticos proprietários para calcular suas exigências de capital risco de mercado.

2.2 Risco de Liquidez

O risco de liquidez é normalmente caracterizado de duas formas. O risco de liquidez de ativos, que representa o risco de uma transação não poder ser executada a preços de mercado (GREGORY, 2012).

E o risco de liquidez de financiamento ou risco de fluxo de caixa refere-se à incapacidade de cumprir as obrigações de pagamentos, que se refere a incapacidade de financiar pagamentos forçando uma liquidação antecipada de ativos e cristalização de perdas (JORION, 2007). Uma vez que tais perdas podem levar a outras questões de financiamento, o risco de liquidez de financiamento pode manifestar-se através do *feedback* negativo entre perdas e necessidades de caixa.

2.3 Risco Operacional

Segundo o Acordo de Basileia de 2001, o risco operacional é o risco de perda, direta ou indireta, resultante de processos internos inadequados ou deficientes, pessoas e sistemas, ou de eventos externos. Isso inclui erro humano, falha em processos, modelo de risco descalibrados, fraudes e risco legal. Enquanto algumas perdas por risco operacional podem ser comuns e moderadas, as perdas mais significativas são prováveis de ser um resultados de cenários altamente improváveis ou de uma combinação de eventos. Embora técnicas quantitativas estejam sendo cada vez mais aplicadas, o risco operacional é extremamente difícil de quantificar (GREGORY, 2012).

2.4 Risco de Crédito

O risco de crédito é a perda de capital em caso de falha do tomador de crédito para reembolsar o montante total da dívida para recuperar o passivo (GONEN; GONEN; GARGEN, 2012), a análise de risco de crédito é um tema importante na área de gestão financeira e usado por muitas organizações financeiras como bancos e varejistas.

Segundo Jorion (2007) o risco de crédito também pode ser definido como sendo o risco de perdas devido ao fato que contrapartes podem ser relutantes ou incapazes de cumprir as suas obrigações contratuais. Isso pode ser caracterizado em termos de falha real ou pela deterioração da qualidade do crédito da contraparte. No primeiro caso, pode resultar em uma perda real e imediata, no último caso, as perdas futuras se tornam mais prováveis o que conduz a um impacto mercado a mercado (JORION, 2007). Quando é caracterizado risco de crédito, a probabilidade da contraparte falhar é um aspecto chave. Entretanto, a potencial exposição a falha e o valor de recuperação associado são importantes quantidades a serem consideradas.

Segundo [Altman e Saunders \(1998\)](#), na década de 70 a maioria das instituições financeiras se baseavam quase exclusivamente na análise subjetiva ou na chamada experiência do banqueiro para avaliar o risco de crédito em empréstimos. Essencialmente, os banqueiros usavam informações sobre várias características do tomador como a sua reputação, o capital e a capacidade, para chegar a um julgamento basicamente subjetivo quanto a possibilidade de conceder crédito.

Entretanto, além dos sistemas multivariados de classificação de crédito tenderem a superar tais sistemas especialistas, os banqueiros tendem a ser excessivamente pessimista sobre o risco de crédito em países menos desenvolvidos ([ALTMAN; SAUNDERS, 1998](#)). Dessa forma, no final do século passado, as instituições financeiras se afastaram de sistemas subjetivos e de especialistas para usarem sistemas que possuem base mais objetiva.

Os modelos de escoragem, desenvolvidos por meio de técnicas estatísticas, são uma forma de mensurar o risco de crédito de forma objetiva. *Credit scoring* é o termo usado para descrever métodos estatísticos formais usados para classificar requerentes de crédito em categorias de ‘bom’ ou ‘mau’ tomadores ([HAND; ADAMS, 2000](#)).

Nessa análise de dados tem-se o interesse em conhecer ou explicar como fatores relacionados ao tomador podem estar correlacionados com a inadimplência. Para estudar essa possível relação existente entre a variável resposta e as covariáveis abordagens estatísticas são amplamente utilizadas ([ALTMAN; SAUNDERS, 1998](#)).

Até a década de 90 as quatro abordagens metodológicas para o desenvolvimento de sistemas multivariados de classificação de crédito mais utilizadas eram: (i) modelo linear de probabilidade, (ii) modelo logit, (iii) modelo probit e (iv) modelo de Análise Discriminante. As metodologias dominantes, em termos de publicações, era a Análise Discriminante seguido por

Análise Logit ([ALTMAN; SAUNDERS, 1998](#)).

Embora muitos desses modelos de classificação de crédito executassem muito bem, eles estavam sujeitos a algumas críticas ([ALTMAN; SAUNDERS, 1998](#)). Os modelos que foram predominantemente baseados em dados contábeis, medidos a intervalos discretos, falhavam em perceber mudanças mais sutis das condições do tomador. E a Análise Discriminante linear e os modelos de probabilidade lineares falhavam em deixar a previsão mais exata e abrandar os pressupostos de linearidade entre as variáveis explicativas.

Além disso, segundo [Huang et al. \(2004\)](#) apesar dos modelos estatísticos serem sucintos e fáceis de explicar, o problema com a aplicação desses métodos ao problema de previsões em finanças é que as suposições de normalidade multivariadas para variáveis independentes são frequentemente violados em conjuntos de dados financeiros, o que torna estes métodos muitas vezes inválidos para amostras finitas.

Outra forma de abordagem, além das técnicas estatísticas é baseada em Inteligência Artificial. Enquanto os métodos estatísticos tendem a ter os dados ajustados ao modelo e normalmente precisam que os pesquisadores coloquem as estruturas do modelo, [Huang et al. \(2004\)](#), as técnicas de aprendizado de máquina extraem automaticamente conhecimento a partir dos dados, gerando modelos complexos que tendem a se ajustar aos dados.

Vários estudos ([HUANG; TZENG; ONG, 2006](#); [ONG; HUANG; TZENG, 2005](#); [AHN, 2011](#)) têm comparado técnicas de Aprendizado de Máquina a técnicas estatísticas tradicionais e mostrado a superioridade delas.

Nos próximos capítulo essas técnicas serão abordadas de forma mais aprofundada.

3 Revisão da Literatura

Neste estudo, foi realizada uma revisão da literatura, seguindo o método proposto por Lage e Godinho (2010) que permite uma análise sistematizada dos artigos publicados na área, incluindo a identificação de principais temas, abordagens metodológicas, escopo, etc. A revisão possibilita ainda a identificação de *gaps* que podem vir a ser explorados em estudos futuros.

É importante destacar que, embora Lage e Godinho (2010) tenham feito uma revisão específica sobre variações do sistema *kanban*, vários pesquisadores já utilizaram esse método de revisão para outras áreas como, por exemplo, treinamento ambiental em organizações (JABBOUR, 2013) e *Data Envelopment Analysis* em desenvolvimento humano (MARIANO; SOBREIRO; REBELATTO, 2015). Porém, desconhecesse estudo que tenha usado o procedimento para a revisão da literatura sobre aprendizagem de máquina aplicada a risco de crédito.

De maneira mais específica, Lage e Godinho (2010) propuseram que, para a realização de uma revisão de literaturara que possibilite não somente um estudo abrangente da área como também indicações de *gaps* é necessário o desenvolvimento de cinco etapas:

1. Conduzir um levantamento amplo e uma avaliação detalhada dos artigos disponíveis sobre o tema em bases de dados acadêmicas;
2. Desenvolver um sistema de classificação com uma codificação lógica que permita agregar informações;
3. Utilizar o sistema de classificação para estruturar o conhecimento existente sobre o tema;
4. Estabelecer um perfil da produção científica publicada a partir do estudo das principais

características dos artigos, utilizando-se o no sistema de classificação; e

5. Identificar as lacunas, bem como as oportunidades e desafios, para estudos futuros.

A primeira etapa do presente estudo foi realizado entre dezembro de 2014 e fevereiro de 2015. Essa etapa envolveu uma pesquisa sistemática de artigos relacionados com aprendizado de máquina aplicado a risco de crédito em dois bancos de dados de artigos, Science Direct e Web of Science. As palavras usadas na pesquisa foram *Machine Learning and Credit Risk*, *Machine Learning and Credit Scoring*, *Machine Learning and Credit Card Borrowing* e *Artificial Intelligence and Credit Risk*. A pesquisa foi realizada nas bases de dados cobrindo título, resumo e palavras-chaves.

Após a exclusão de artigos repetidos, de estudos que não investigavam necessariamente aprendizagem de máquina aplicada a crédito e de trabalhos que não estavam disponíveis para *download*, foi realizada uma análise preliminar, buscando identificar os artigos que recebiam várias citações e que não estavam ainda no conjunto levantado. Com a inclusão dos artigos mais citados, a base final foi composta por 80 trabalhos que foram utilizados na etapa seguinte.

3.1 Classificação e Codificação

Após a coleta e a seleção dos principais estudos relacionados com aprendizado de máquina aplicada a risco de crédito, os trabalhos foram analisados para que fosse construída uma estrutura que permitisse classificar e codificar características dos artigos.

Considerando os trabalhos levantados, foi definido um sistema de classificação baseado em 12 itens principais como, por exemplo, técnica utilizada, base de dados, tipo de estudo, contribuição, etc. Para cada um dos itens de classificação numerados de 1 a 12, foram utilizadas codificações usando as letras do alfabeto (A, B, C, e assim por diante), para especificar as

características do artigo. Tendo em vista que trabalhos podem apresentar diversas perspectivas de análise, um artigo pode receber mais do que um código para cada item.

A Tabela 1, obtida a partir das características principais dos artigos levantados, apresenta a estrutura de classificação e códigos utilizados neste trabalho.

Classification	Meaning	Codes for alternatives
1	Type of study	A - Theoretical B - Empirical C - Review D - Outro
2	Main technique	A - Bagging B - Boosting C - SVM D - Artificial Neural Networks E - Decision Trees F - Bayesian Statistics G - Genetic Algorithm H - Random Forests I - Nearest Neighbor J - Fuzzy Logic K - ELM L - Other M - Not applicable
3	Secondary technique	A - Discriminant Analysis B - Logistic Regression C - Decision Trees D - Artificial Neural Networks E - Other F - Not applicable
4	Scope	A - Application of a traditional technique B - Enhancement on a technique C - Development of a new technique D - Comparison of techniques E - Other
5	Source of data	A - Test data B - Financial databases C - Proprietary data D - Not applicable
6	Type of credit	A - Credit card B - Personal loans and financing C - Corporate loans or bonds D - Sovereign bond E - Other F - Not applicable
7	Type of scale	A - Nominal/Categorical B - Ordinal C - Interval D - Ratio E - Other F - Not applicable
8	Number of explanatory variables	A - 1 to 10 B - 11 to 20 C - More than 20 D - Not applicable
9	Time horizon	A - Less than 3 years B - Between 4 to 5 years C - Between 5 to 10 years D - More than 10 years E - Not specified F - Not applicable
10	Type of sample	A - Small sample B - Medium sample C - Large sample D - Big data E - Not specified F - Not applicable
11	Country of origin of data	A - USA / Canada B - Europe C - Latin America D - Asia/Oceania E - Africa F - Not applicable
12	Results	A - Generate new conclusions B - Corroborate erliers studies C - Other

Tabela 1 – Quadro de classificação e codificação dos estudos analisados.

3.2 Análise Descritiva

Conforme já indicado, o levantamento de artigos, conduzido segundo a metodologia descrita anteriormente, identificou 80 estudos ligados à aplicação de técnicas de aprendizado de máquina em risco de crédito. Com base na análise dos artigos, foram estabelecidas as classificações e os códigos descritos na Tabela 26, no Apêndice.

Inicialmente são descritas as características básicas dos artigos revisados. Essas características incluem a temporalidade da publicação, os periódicos em que os artigos são publicados, os autores dos artigos, os países e as instituições de afiliação dos autores.

Os artigos analisados foram publicados entre 1992 e 2014. A Figura 2 mostra a evolução do número de artigos publicados ao longo dos anos. É possível notar que houve um crescimento considerável de artigos publicados após 2006, que pode ser explicado pela popularização recente das técnicas de aprendizado de máquina.

Nota-se, portanto, a relevância e a atualidade de trabalhos que exploram a aplicação de mecanismos de aprendizagem de máquina em risco de crédito. Os anos com maior número de publicações foram 2009 e 2012, com 12 publicações cada. Os anos mais recentes sugerem uma diminuição do número de artigos que pode ser explicada pela menor novidade do tema e maior maturidade da literatura.

A base de artigos levantada contempla autores de 26 países diferentes. Autores de Taiwan, China e Estados Unidos destacam-se no estudo do tema. A Figura 3 apresenta os países da afiliação dos pesquisadores citados na revisão de literatura.

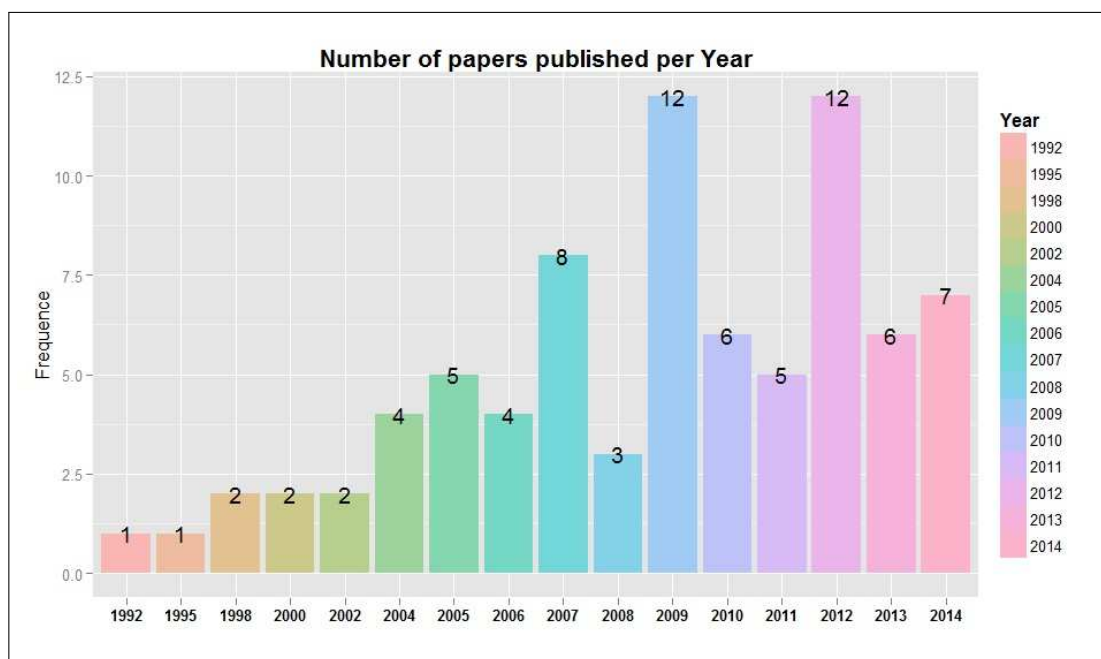


Figura 2 – Artigos publicados por ano.

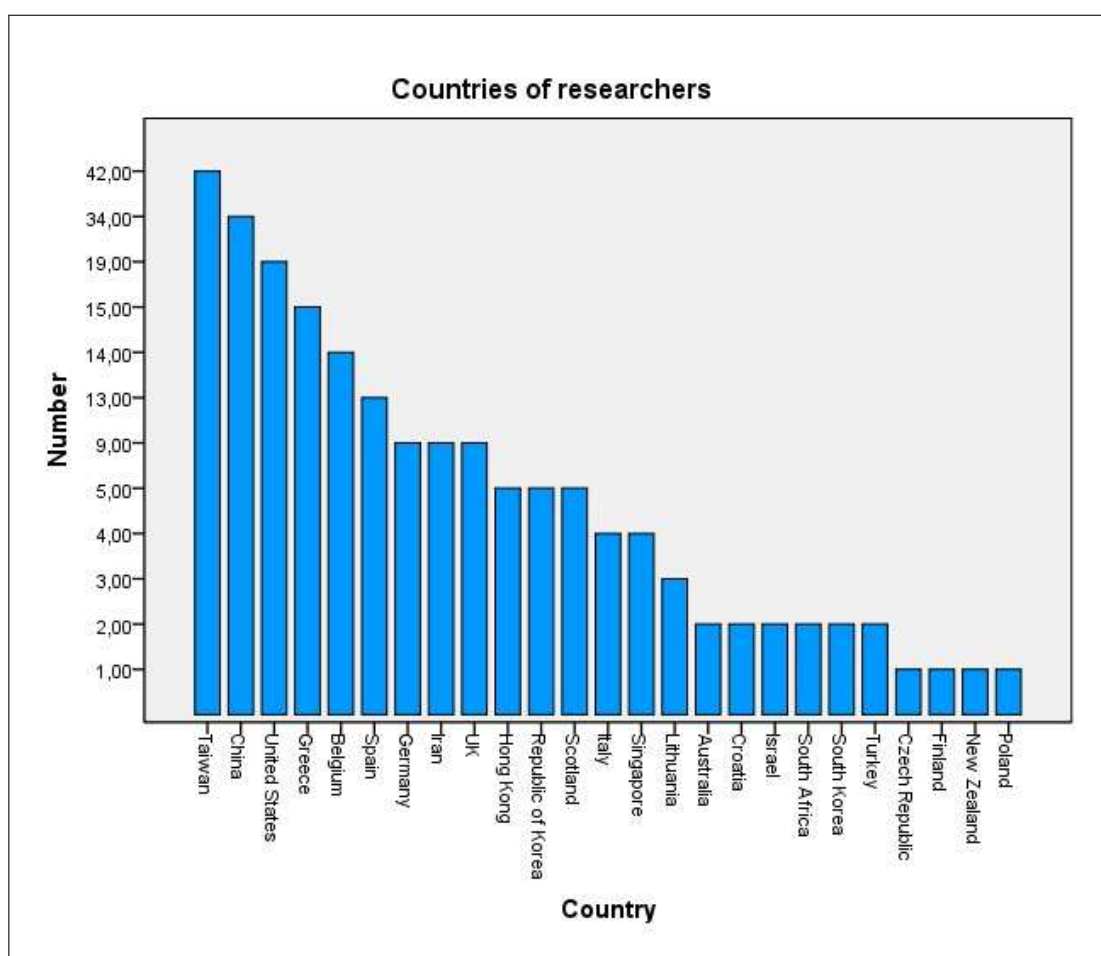


Figura 3 – Países com o maior número de pesquisadores.

É interessante destacar que a revisão da literatura não identificou autor ligado a universidades latino americanas, indicando atuação menos influente de pesquisadores dessa região na exploração de técnicas computacionais de aprendizagem de máquina em risco de crédito.

Considerando a produtividade individual, o autor mais proffícuo é Chih-Fong Tsai, da National Central University, de Taiwan, que publicou sete artigos, tendo como foco previsão de falência utilizando classificadores híbridos, incluindo SVM *Support Vector Machines*, DT *decision Trees* e ANN *Artificial Neural Networks*. Constantin Zopounidis e Michael Doumpos, da Technical University of Crete, na Grécia, também se sobressaem, tendo publicado quatro artigos cada, explorando aprendizado de máquina e risco de crédito.

Os *journals* em que os estudos foram veiculados são apresentados na Tabela 2. Considerando os artigos levantados no estudo, os principais *journals* que publicam artigos de aprendizagem de máquina em crédito são *Expert Systems with Applications* e *European Journal of Operational Research*, que congregam 49%, quase a metade, dos trabalhos analisados.

Journal	# of studies	Percent
Expert Systems with Applications	29	36
European Journal of Operational Research	10	13
Knowledge-Based Systems	4	5
Applied Mathematics and Computation	2	3
Applied Soft Computing	2	3
Computers and Operations Research	2	3
Decision Support Systems	2	3
Journal of Banking and Finance	2	3
Neurocomputing	2	3
Procedia Computer Science	2	3
Others	23	29

Tabela 2 – Distribuição dos artigos nos periódicos.

O estudo também identificou as palavras-chaves mais usadas, apresentadas na Tabela 3, permitindo identificar sub-temas ou sub-áreas que são mais explorados. Em particular, percebe-se

que o uso mais disseminado de técnicas que levam em consideração *Support Vector Machines*.

Palavra-Chave	Número de Citações
Credit Scoring	33
Support Vector Machine(s)	14
Classification	13
Machine Learning	11
Neural Networks	9
Credit Risk	8
Data Mining	8
Feature Selection	8
Bankruptcy Prediction	6

Tabela 3 – Palavras-Chaves mais citadas.

A Figura 4 apresenta um *word cloud* das palavras-chave dos estudos analisados, sendo possível identificar que as palavras *credit*, *vector*, *scoring* e *risk* se destacam. Obviamente, pelo próprio critério de busca utilizado, *credit* e *risk* seriam palavras já esperadas. Todavia, a Figura 4 sugere que muitos estudos exploram o problema mais específico de *credit scoring* usando *Support Vector Machines*.

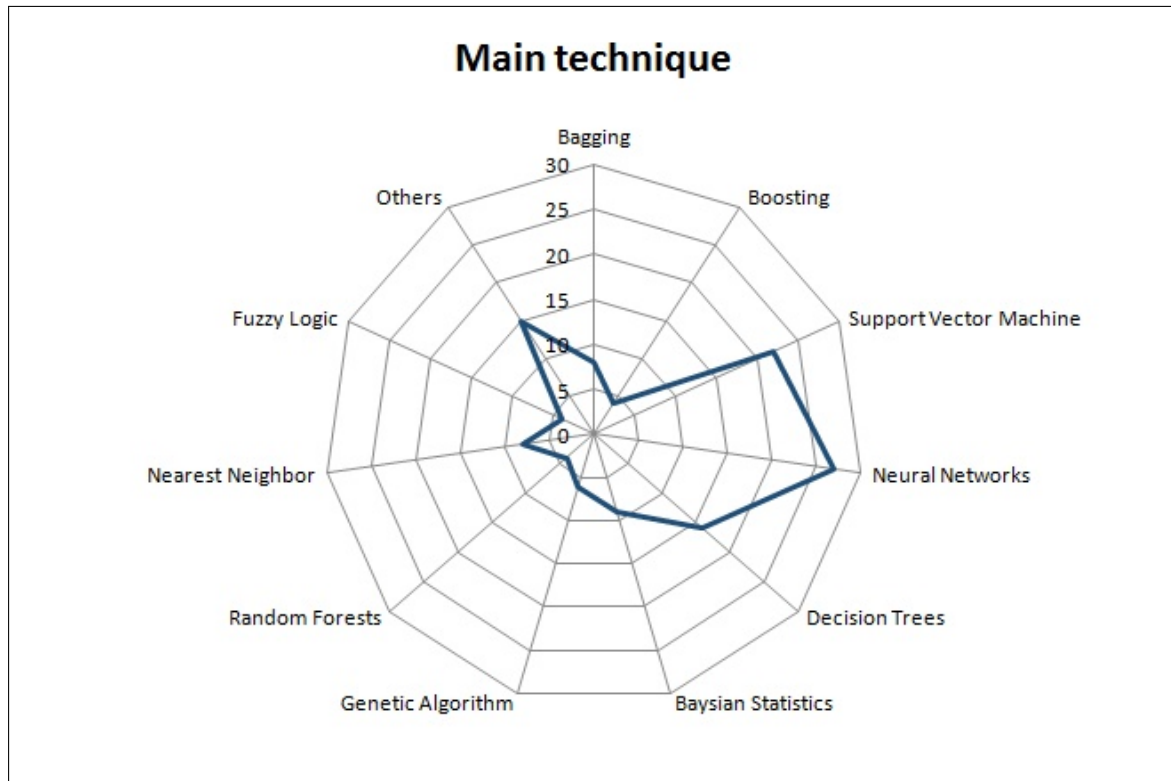


Figura 5 – Técnicas de aprendizado de máquina mais utilizadas.

secundárias mais utilizadas nos estudos. A análise da literatura indica que *Logistic Regression*, encontrada em 21% dos artigos, seguida da *Discriminant Analysis* com 17%, são técnicas secundárias mais usuais. *Artificial Neural Networks* associadas a outras técnicas de aprendizagem de máquina constam em 16% dos artigos.

3.4 Fonte dos Dados

Para o teste empírico dos modelos de aprendizagem de máquina propostos, cerca de 67% dos artigos utilizaram dados de empréstimos e financiamentos pessoais e/ou de cartões de crédito.

A maior parte dos artigos, 51%, utilizou dados de repositórios disponíveis *online* como, por exemplo, de arquivos referentes a Aprovação de Crédito da Austrália, Pagamento de Cartões de Crédito da Alemanha e Triagem de Crédito do Japão. Esses bancos de dados são de livre

acesso e disponibilizado no *Machine Learning Repository* pela *University of California Irvine*.

O repositório traz uma coleção de bancos e de geradores de dados, de diversas áreas do conhecimento, que é amplamente utilizada pela comunidade de pesquisadores de aprendizado de máquina para a análise empírica dos algoritmos. Adicionalmente, por constituírem bases padrão, os dados podem ser usados para comparar acurácia e velocidade de diferentes modelos de predição.

Destaca-se ainda, conforme já indicado, que diversos artigos aplicam técnicas de aprendizagem de máquina em outras bases de dados de instituições financeiras e empresas de cartão de crédito, permitindo identificar a adequação de modelos em outras aplicações do mundo real.

Para a construção de um modelo de previsão, no qual se busca classificar agentes em "bom pagador" ou "mau pagador", são necessárias variáveis que representem características de tomadores. Considerando a quantidade, o tipo e o horizonte de tempo das variáveis e o número de observações utilizadas nos estudos empíricos, o levantamento bibliográfico identificou que:

- A maioria dos artigos, 53%, utilizou bases de dados contemplando entre 11 e 20 variáveis explicativas.
- 46% dos artigos utilizaram variáveis de diversas escalas: *Nominal/Categorical*, *Ordinal*, *Interval e Ratio*.
- 60% dos artigos não especificaram qual o horizonte de tempo dos dados utilizados.
- 62% dos artigos utilizaram uma amostra contemplando entre 1000 e 5000 observações.

3.5 Tipos de Estudo

Com relação à abordagem metodológica, dos 80 artigos analisados, a grande maioria (78) apresenta estudo empírico, havendo apenas um artigo de revisão de literatura e um artigo somente teórico. É importante destacar, portanto, que a pesquisa na área tem se preocupado com o desenvolvimento e teste de técnicas de aprendizagem de máquina para aprimoramento prático do risco de crédito. Assim, o foco dos estudos tem se distanciado da discussão teórica da fundamentação da aplicabilidade da técnica e se direcionado para a identificação pragmática de mecanismos que melhorem a classificação de tomadores de recursos.

Como mostrado na Tabela 4, 49% dos estudos utilizam métodos para aperfeiçoamento de técnicas tradicionais de aprendizagem de máquina e somente 12% dos estudos têm como foco o desenvolvimento de novas técnicas.

Assim, dentro do contexto de aplicações de modelos de aprendizagem de máquina para risco de crédito, ainda há interesse dos *journals* em publicar trabalhos que explorem evoluções marginais de técnicas tradicionais.

Scope	# of studies	Percent
B	36	45.0
D	25	31.3
C	7	8.8
A	3	3.8
B & D	3	3.8
C & D	3	3.8
A & C	1	1.3
A & D	1	1.3
E	1	1.3

Tabela 4 – Escopo dos artigos. *A: Application of a traditional technique. B: Enhancement on a technique. D: Development of a new technique. E: Comparison of techniques. F: Other.*

Tendo em vista que ainda não foi identificado um algoritmo de aprendizagem de máquina

que seja consistentemente superior, a área do conhecimento ainda demanda trabalhos que analisem diferentes bases de dados. Considerando os artigos analisados, 83% possuem resultados que corroboram estudos anteriores, somente 15% apresentam novos resultados.

3.6 Principais Autores e Citações

Foi utilizado o *software* Pajek para elaborar uma rede de citações que indicasse os principais artigos na área. A Figura 6 apresenta os artigos que possuem mais citações entre os estudados. A seta indica a direção de citação dos autores, ou seja, ela aponta para o autor que cita aquele que está no outro extremo da linha. E o tamanho do círculo mostra o peso das citações de cada artigo. Quanto maior o círculo, mais artigos da amostra citaram esse autor.

Desses 10, o primeiro artigo do caminho e também o mais citados é o artigo de [Davis, Edelman e Gammerman \(1992\)](#).

O artigo de [West \(2000\)](#) também recebeu um grande número de citações.

O artigo *Credit rating analysis with support vector machines and neural networks: a market comparative study* de [Huang et al. \(2004\)](#) é o artigo com mais citações, 641. Seguido pelo artigo *Neural network credit scoring models*, de [West \(2000\)](#), com 622. Ambos os artigos envolvem modelos utilizando Artificial Neural Networks.

Dos 10 artigos mais citados, somente 3 não abordam técnicas ligadas a ANN. A tabela 5 apresenta os artigos que possuem mais de 100 citações.

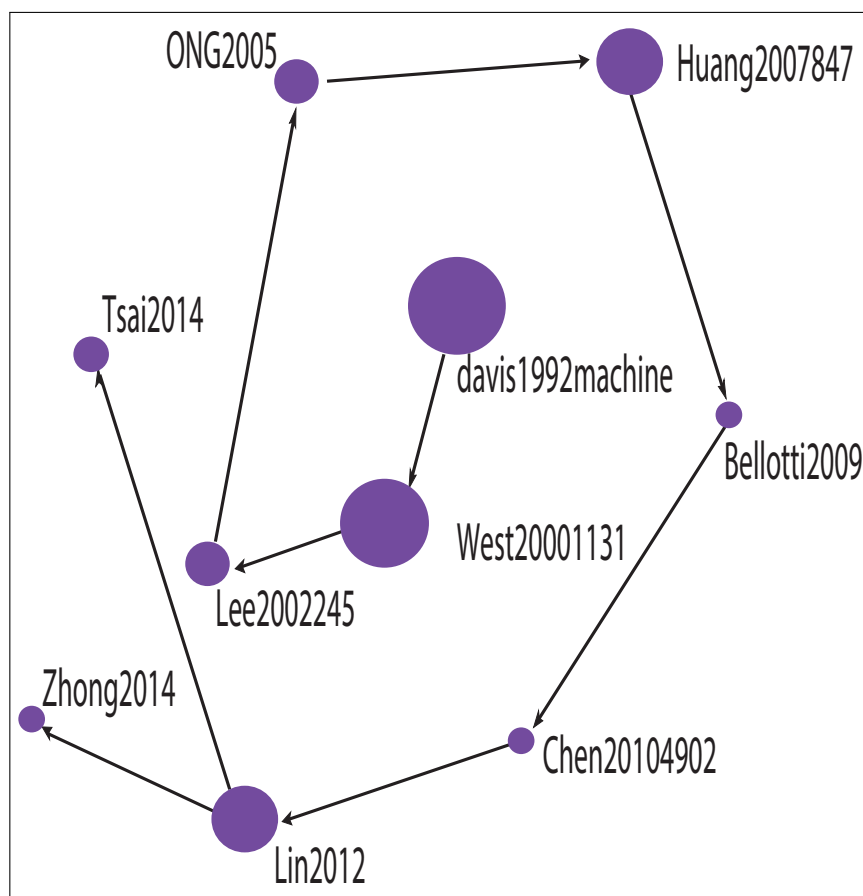


Figura 6 – Rede de Autores.

Título do Artigo	Autores	N. Citações
Credit rating analysis with support vector machines and neural networks: a market comparative study	Huang et al. (2004)	641
Neural network credit scoring models	West (2000)	622
Credit scoring with a data mining approach based on support vector machines	Huang, Chen e Wang (2007)	381
Credit scoring using the hybrid neural discriminant technique	Lee et al. (2002)	271
A New Fuzzy Support Vector Machine to Evaluate Credit Risk	Wang, Wang e Lai (2005)	231
Using neural network ensembles for bankruptcy prediction and credit scoring	TSAI e WU (2008)	229
Comprehensible credit scoring models using rule extraction from support vector machines	Martens et al. (2007)	228
Building credit scoring models using genetic programming	ONG, HUANG e TZENG (2005)	219
Genetic algorithms applications in the analysis of insolvency risk	Varetto (1998)	212
Hybrid mining approach in the design of credit scoring models	Hsieh (2005)	147
Application of support vector machines to corporate credit rating prediction	Lee (2007)	136
An intelligent-agent-based fuzzy group decision making model for financial multicriteria decision support: The case of credit scoring	Yu, Wang e Lai (2009)	134
Feature selection in bankruptcy prediction	Tsai (2008a)	126
Credit Risk Assessment Using Statistical and Machine Learning: Basic Methodology and Risk Modeling Applications	Galindo e Tamayo (2000)	123
A neural network approach for credit risk evaluation	Angelini, Tollo e Roli (2008)	123
Two-stage genetic programming (2SGP) for the credit scoring model	Huang, Tzeng e Ong (2006)	121
Support vector machines for credit scoring and discovery of significant features	Bellotti e Crook (2009)	119
Monotonicity Maintenance in Information-Theoretic Machine Learning Algorithms	Ben-David (1995)	118
Evaluating feature selection methods for learning in data mining applications	Piramuthu (2004)	114

Tabela 5 – Número de citações no Google Scholar dos principais artigos em 18/10/2015.

Selecionando somente os artigos mais recentes, com data de publicação a partir de 2010, foram selecionados os 10 mais citados, Tabela 6.

Título do Artigo	N. Citações	Autores
A comparative assessment of ensemble learning for credit scoring	98	Wang et al. (2011)
Recent advances on support vector machines research	60	Tian, Shi e Liu (2012)
Two credit scoring models based on dual strategy ensemble trees	37	Wang et al. (2012)
Machine Learning in Financial Crisis Prediction: A Survey	36	Lin, Hu e Tsai (2012)
A corporate credit rating model using multi-class support vector machines with an ordinal par wise partitioning approach	34	Kim e Ahn (2012)
Genetic algorithm-based heuristic for feature selection in credit risk assessment	32	Oreski e Oreski (2014)
Hybrid models based on rough set classifiers for setting credit rating decision rules in the global banking industry	31	Chen e Cheng (2013)
Combining cluster analysis with classifier ensembles to predict financial distress	23	Tsai (2014)
Credit Risk Evaluation Model Development Using Support Vector Based Classifiers	19	Danenas, Garsva e Gudas (2011)
Two-level classifier ensembles for credit risk assessment	19	Marques, Garcia J.a e Sanchez (2012)
Exploring the behaviour of base classifiers in credit scoring ensembles	19	Marques, Garcia e Sanchez (2012)
Comparing the learning effectiveness of BP, ELM, I-ELM, and SVM for corporate credit ratings	17	Zhong et al. (2014)
Consumer credit risk: Individual probability estimates using machine learning	16	Kruppa et al. (2013)
Credit risk evaluation with kernel-based affine subspace nearest points learning method	14	Zhou et al. (2011)
Feature selection in corporate credit rating prediction	11	Hajek e Michalak (2013)
A granular computing-based approach to credit scoring modeling	10	Saberi et al. (2013)

Tabela 6 – Número de citações dos artigos mais recentes no *Google Scholar* dos principais artigos em 18/10/2015.

Enquanto que nos artigos mais antigos, ANN era a técnica mais estudada nos artigos mais citados. Atualmente, a técnica mais estudada nos artigos com maior número de citações são os métodos de classificação *ensemble*, presente em 6 dos 10 artigos.

Após a classificação de todos os artigos foi desenvolvido um breve resumo que inclui os objetivos e principais conclusões de cada artigo utilizado na revisão de literatura. A Tabela 27, no apêndice, apresenta as principais características dos artigos estudados.

Combinando as Tabelas 26 e a Tabela 27, apresentada no Apêndice 1, foi desenvolvido o conteúdo do próximo capítulo. Serão descritas as trajetórias das principais técnicas de aprendizado de máquina utilizadas na avaliação do risco de crédito nos 80 artigos analisados.

4 Aprendizado de Máquina na Avaliação de Risco de Crédito

Um dos primeiros trabalhos a aplicar técnicas de aprendizado de máquina no contexto do risco de crédito foi [Davis, Edelman e Gamberman \(1992\)](#). Nesse artigo, os autores aplicam uma série de algoritmos para avaliação de risco de inadimplência em cartão de crédito. São testados dois modelos: (1) um modelo computacional geral denominado de G&T, composto por duas partes principais, o processo de seleção e o procedimento de escolha, e (2) um modelo conectivo baseado em ANN, que se refere à rede pelo qual as células são unidas por arcos com pesos associados.

Apesar de os resultados desse estudo serem limitados pelo pequeno número de observações da base dados e das características das técnicas testadas, o modelo G&T com os algoritmos de ANN apresentou resultados que indicam a relevância do uso de mecanismos de aprendizagem de máquina.

Posteriormente, [Ben-David \(1995\)](#) propôs uma métrica de seleção de atributo que leva tanto o erro como a monotonicidade em consideração ao construir DT. A métrica mostrada é empiricamente capaz de reduzir o grau de não-monotonicidade das DT significativamente sem sacrificar a sua exatidão indutiva.

Em 1998, [Hashemi et al. \(1998\)](#) e [Varetto \(1998\)](#) publicaram artigos propondo o uso de duas outras técnicas de aprendizado de máquina. O primeiro artigo utilizou um composto de conjuntos de componentes neurais que obteve uma alta taxa de classificação enquanto o segundo artigo comparou Análise Discriminante linear (LDA) com o algoritmo genético (GA).

Varetto (1998) identifica que o uso de GA torna possível obter funções lineares otimizadas que não dependem de hipóteses estatísticas restritiva em relação a normalidade das distribuições dos escores e da igualdade das matrizes de variância/covariância.

Vários outros autores publicaram artigos utilizando técnicas de aprendizado de máquinas na avaliação do risco de crédito. Por exemplo, Galindo e Tamayo (2000) que introduz uma metodologia de modelagem específica com base no estudo de curvas de erro e West (2000) que investiga a precisão de classificação de cinco modelos de redes neurais. Adicionalmente, o trabalho de Zopounidis e Doumpos (2002) propõe um novo método para alcançar a discriminação multi-grupos com base em um processo de segmentação binária iterativo.

Shi, Zhang e Qiu (2013) apresenta um recurso de ponderação de modelo de *credit scoring* SVM para avaliação de risco de crédito em que um F-score é adotado para o *ranking* de importância do recurso. As duas versões ponderadas pela metragem do SVM são testadas contra o SVM tradicional em dois conjuntos de dados do mundo real e os resultados da pesquisa revelam a validade do método ponderado proposto.

Mais recentemente, Zhong et al. (2014) realizou um estudo comparativo experimental abrangente sobre a eficácia de quatro algoritmos de aprendizagem, *Backpropagation* (BP), *Extreme Learning Machine* (ELM), *Incremental Extreme Learning Machine* (I-ELM), e SVM.

Os resultados experimentais mostraram que abordagens baseadas em *Single-Hidden Layer Feed Forward Neural Network* (SLFN) superam em confiabilidade os SVMs, enquanto que SVMs são melhores em distribuições de saída do que os outros. O ELM e BP são as abordagens mais confiáveis e independentemente da distribuição de saída, têm desempenho superior a I-ELM e SVM considerando-se a problemática de *rating* de crédito.

Kamalloo e Abadeh (2014) propuseram um novo classificador usando os princípios do

sistema imunológico e regras *fuzzy* para prever inadimplência em bancos. No modelo proposto é utilizada memória imunológica para lembrar as células boas durante o processo de clonagem. Foi desenhada duas formas de memória: a memória simples e de memória de camada k. Os resultados indicaram que o sistema de classificação proposto, baseado em imunologia é preciso na detecção de riscos de crédito.

Outros trabalhos envolvendo vários temas agregados como o estudo de [Oreski e Oreski \(2014\)](#), no qual é apresentado um algoritmo genético com redes neurais (HGA-NN), usado para identificar um subconjunto de variáveis ideal para aumentar a precisão da classificação e a escalabilidade na avaliação de risco de crédito.

Os resultados experimentais que foram alcançados com o classificador proposto HGA-NN são promissores para a seleção de variáveis e classificação na avaliação de risco de crédito de varejo. O artigo também indica que o classificador HGA-NN é um promissor complemento às técnicas de mineração de dados existentes.

Há ainda uma linha de estudos que explora dados contábeis como, por exemplo, o trabalho de [Niklis, Doumpos e Zopounidis \(2014\)](#), cujo objetivo foi criar um modelo de classificação de risco de crédito, através de uma metodologia de aprendizado de máquina que combina dados de contabilidade com a abordagem baseada em precificação de opções de Black, Scholes e Merton.

Os resultados obtidos demonstraram que, mesmo em condições problemáticas do mercado de ações como o mercado da Grécia na última década, a previsibilidade dos modelos baseados no mercado é muito competitivo com modelos de *rating* de crédito tradicionais.

Finalmente, considerando a diversidade de métodos de aprendizagem de máquina, é importante destacar que, segundo [Mitchell \(1997\)](#) e [Galindo e Tamayo \(2000\)](#), algoritmos para análise de risco de crédito variam enormemente em sua estrutura básica, mas podem ser

grosseiramente classificados em alguns grupos:

- Técnicas estatísticas tradicionais: Análise Discriminante linear, quadrática e logística, análise de regressão, MANOVA;
- Técnicas estatísticas modernas: *k-Nearest-Neighbors*, busca de projeção, *Alternating Conditional Expectations* (ACE), *Statistical Methods and Applications for Research in Technology* SMART, *Multivariate Adaptive Regression Splines* (MARS);
- *Decision Trees* e métodos de indução baseados em regras: DT, sistemas especialistas;
- *Neural Networks* e máquinas afins: *Feedforward ANN*, mapas auto-organizados, funções de base radial, SVM.
- Inferência e Redes Bayesianas;
- Métodos de combinação de modelos: *Boosting e Bagging*;
- Algoritmos genéticos e de agentes inteligentes;
- Lógica fuzzy, amostragem fractal e abordagens híbridas

Serão agora apresentados os históricos das técnicas de aprendizado de máquina mais utilizadas nos últimos anos.

4.1 Artificial Neural Networks

As *Artificial Neural Networks* (ANN) constituem uma das primeiras técnicas de aprendizado de máquina a ser utilizada na avaliação de risco de crédito. Por exemplo, [West \(2000\)](#) investigou a precisão de classificação de crédito de cinco modelos de ANN: *Multilayer Perceptron*, *Mixture of Experts*, *Radial Basis Function*, *Learning Vector Quantization* e *Fuzzy Adaptive*

Resonance, comparando resultados com os dois modelos paramétricos mais utilizados na época, Análise Discriminante linear e Regressão Logística.

Os resultados indicam que a ANN *Mixture of Experts* é um pouco mais precisa do que os outros modelos de *credit scoring* para os dois conjuntos de dados investigados na pesquisa. De acordo com West (2000), uma possível fonte de vantagem da ANN é a capacidade de o modelo particionar o espaço de entrada fazendo com que a formação da rede convirja mais rapidamente ao mínimo global na superfície de erro.

Outro trabalho que comparou ANN com as técnicas tradicionais é o de Lee et al. (2002), no qual foi explorado o desempenho na classificação de crédito, integrando *Backpropagation Neural Network* (BNN) com a abordagem tradicional de Análise Discriminante. A abordagem híbrida proposta no artigo converge muito mais rapidamente do que o modelo ANN convencional. Além disso, o aumento de precisão de classificação de crédito na metodologia proposta supera a Análise Discriminante tradicional e métodos baseados em Regressão Logística.

Angelini, Tollo e Roli (2008) também descrevem o caso de uma aplicação bem sucedida de ANN para avaliação de risco de crédito. Nesse trabalho, foram desenvolvidos dois sistemas de ANN, uma rede com um padrão de alimentação de entrada e outro com uma arquitetura para fins especiais. Os autores concluíram que as ANN podem ser muito bem sucedidas na aprendizagem e na estimativa do *default* de um tomador, desde que a análise dos dados de pré processamento e de treinamento seja realizada de forma cuidadosa.

Diversos outros autores publicaram artigos aplicando técnicas de ANN no risco de crédito como, por exemplo, Lin (2009), Setiono, Baesens e Mues (2009) e Oreski e Oreski (2014), indicando que essa técnica tem aplicações relevantes em finanças.

4.2 Decision Trees

Um dos primeiros artigos a utilizar técnicas de *Decision Trees* na avaliação do risco de crédito foi o de [Ben-David \(1995\)](#).

[Galindo e Tamayo \(2000\)](#) introduziram uma metodologia de modelagem específica com base no estudo de curvas de erro. Usando técnicas de modelagem, foram construídos mais de 9.000 modelos como parte do estudo. Os resultados mostraram que os modelos de *Decision Trees* baseados em *Classification And Regression Trees (CART)* fornecem melhor estimativa para *default* superando as ANN.

No estudo de [Piramuthu \(2004\)](#), são avaliados vários métodos de seleção de características baseados em distância probabilística quanto a sua eficácia no pré-processamento de dados de entrada para a indução de DT.

Os autores mostraram que o aprendizado em DT induzida é sensível aos dados de entrada utilizados. Ao selecionar características adequadas, através de pré-processamento, o desempenho de DT induzida pode ser melhorada sem muito esforço, uma vez que a maioria destas técnicas de pré-processamento não são intensivas em relação ao tempo e recursos computacionais.

Ainda usando o conceito de árvores de decisão, [Zhang et al. \(2010\)](#) propôs um novo modelo de *credit scoring*, o *Vertical Bagging Decision Tree Model (VBDTM)*. Os resultados da análise mostram que o desempenho do método proposto é excelente com relação a precisão da classificação.

Outro estudo que utilizou DT é o de [Marques, Garcia e Sanchez \(2012\)](#), no qual os autores avaliam o desempenho de sete técnicas de previsão individuais, *1-Nearest Neighbour (1-NN)*, *Naive Bayes Classifier (NBC)*, *Logistic Regression (logR)*, *Multilayer Perceptron (MLP)*,

Radial Basis Function (RBF) Neural Networks, Support Vector Machine (SVM) with a linear kernel e C4.5 Decision Tree. Quando usadas como parte de cinco métodos de classificadores *ensemble*, *AdaBoost*, *Bagging*, *Random Subspace*, *DECORATE* e *Rotation Forest*. O objetivo final desse estudo é sugerir classificadores apropriados para cada abordagem conjunta no contexto de análise de crédito.

Segundo, [Marques, Garcia e Sanchez \(2012\)](#) um classificador *ensemble* (também chamado de comitê de *learners*, mistura de especialistas e sistema de classificadores múltiplo), consiste em um conjunto de classificadores treinados individualmente (classificadores de base), cujas decisões são de alguma forma combinadas.

Os resultados experimentais e os testes estatísticos mostraram que a DT constitui a melhor solução para métodos de classificadores *ensemble*, seguido de perto pelo ANN *Multilayer Perceptron* e pela Regressão Logística.

4.3 Support Vector Machines

Um dos primeiros estudos a utilizar *Support Vector Machines (SVM)* na análise de crédito foi [Huang et al. \(2004\)](#). Nesse artigo, os autores compararam *Backpropagation Neural Network (BNN)* com SVM. Os resultados mostraram que o SVM alcança precisão comparável a de BNN. A aplicação dos resultados dessa pesquisa na interpretação do modelo de ANN contribuiu para mostrar a importância relativa das variáveis de entrada na área financeira.

[Wang, Wang e Lai \(2005\)](#) propuseram um modelo SVM *fuzzy* com maior capacidade de generalização, enquanto preserva insensibilidade a *outliers*. Os resultados empíricos mostraram que o modelo proposto pode encontrar aplicação promissora na análise de crédito, alcançando melhor resultado que o SVM padrão. O SVM *fuzzy* propõe combinações de métodos de Kernel e

de geração de adesão. Os resultados também mostraram que o novo SVM *fuzzy* pode consistentemente atingir um desempenho comparável aos métodos tradicionais de avaliação de crédito gerados por Regressão Logística.

Huang, Chen e Wang (2007) investigaram três estratégias para construir os modelos de *scoring* de crédito baseadas em modelos SVM híbridos para avaliar a pontuação de crédito do requerente a partir das variáveis de entrada.

Os modelo SVM teve o seu desempenho comparado com ANN, GP e DT. Os autores chegaram às seguintes conclusões: O modelo de *credit scoring* com abordagem baseada em SVM pode indicar corretamente os tomadores como "aceitos" ou "rejeitados", minimizando assim o risco de crédito e gerando consideráveis economias futuras.

Segundo os autores, o modelo baseado no SVM é muito competitivo com BPN e GP em termos de precisão de classificação. De acordo com o estudo, um sistema híbrido SVM-GA é uma boa alternativa para otimizar os parâmetros e dar maior funcionalidade ao subconjunto de variáveis.

Assim, com um pequeno subconjunto de variáveis, um sistema híbrido SVM-GA pode obter um bom desempenho na classificação. No entanto, Huang, Chen e Wang (2007) indicam que, quando se usa estratégia SVM-GA, bem como GP e BPN, deve-se evitar o excesso de treinamento. A desvantagem do modelo de pontuação de crédito SVM-GA é o seu longo tempo de treinamento.

Em tais casos, os algoritmos que proporcionam uma precisão de classificação muito boa não podem ser aplicados devido ao custo da alta complexidade computacional. Felizmente, os sistemas baseados em GA são bem adequados para a arquitetura paralela (HUANG; CHEN; WANG, 2007).

Lee (2007) sugeriu um modelo com melhor poder explicativo e estabilidade aplicando SVM para o problema de *rating* de crédito corporativo. Os resultados do experimento mostram que SVM supera os outros métodos testados como, por exemplo, *Case-based reasoning*, *Back-propagation neural networks* e *Multiple discriminant analysis*.

Vários outros autores publicaram estudos sobre SVM como, por exemplo, Liu, Fu e Lin (2010), que apresenta um novo modelo de *credit scoring* quantitativo baseado em SVM com algoritmo genético adaptativo, gr-GA-SVM, Tian, Shi e Liu (2012), que revê vários modelos de otimização de representação em SVM e suas aplicações em economia.

4.4 Classificadores *Ensemble*

Além das técnicas discutidas anteriormente, dentro do contexto de aprendizagem de máquina, existem ainda diversos mecanismos que podem ser usados na análise de crédito como, por exemplo, os métodos *ensemble*. Dois algoritmos *ensemble* tradicionais são o *Bagging* e o *Boosting*.

O *Bagging* (*Bootstrap Aggregating*), um método proposto por Breiman (1996), é fundamentado na geração de amostras *bootstrap* dos dados, permitindo a obtenção de preditores individuais que, posteriormente, são agregados, ou combinados, para formar um preditor final melhor. Breiman (1996) mostrou que a variância do preditor combinado é menor ou igual a variância de qualquer um dos preditores individuais envolvidos na combinação.

Em 2010, Zhang et al. (2010) propôs um novo modelo de *credit scoring*, chamado VBDM. O novo modelo VBDM, diferente do Bagging tradicional, recebe uma agregação de classificadores por meio de uma combinação de variáveis preditivas. No modelo VBDM, todas as amostras de treino e apenas parte das variáveis participam na aprendizagem de cada

classificador. Em contraste, os classificadores são formados com os subconjuntos de amostra no método tradicional de Bagging e cada classificador tem as mesmas variáveis (ZHANG et al., 2010). De acordo com os autores, os resultados da análise mostraram que o desempenho do método proposto é superior na precisão da previsão, sendo mais robusto que outros tipo de modelos como ANN e SVM.

Outro artigo que mostrou a superioridade dos classificadores *ensemble* foi o de Wang et al. (2011). Os autores realizaram uma avaliação comparativa do desempenho de três métodos *ensemble*, Bagging, Boosting, e Stacking, a partir de quatro mecanismos-base de aprendizado, Regressão Logística, *Decision Trees*, *Artificial Neural Networks* e *Support Vector Machines*. Os resultados experimentais revelam que os três métodos *ensemble* podem melhorar substancialmente o aprendizado a partir das funções-base.

Em particular, Bagging tem um desempenho melhor do que Boosting. Além disso, Stacking e Bagging DT obtiveram melhores resultados em termos dos três indicadores de desempenho, a precisão média, erro tipo I e erro tipo II.

Marques, Garcia J.a e Sanchez (2012) mostraram que classificadores *ensemble* de dois níveis, no qual dois *ensembles* de diferentes naturezas são combinados, podem constituir uma solução adequada para os problemas de análise de crédito. Os resultados da pesquisa mostram desempenho superior em comparação com métodos *ensemble* tradicionais e muito significativamente melhor em relação a classificadores individuais.

Segundo os autores, os classificadores *ensemble* de dois níveis têm produzido os melhores resultados em termos de curva ROC, o que pode levar a economias de custo significativas em aplicações de classificação de crédito. Uma vez que a escolha do modelo de conjunto de dois níveis é importante, dentro da análise realizada, o uso conjunto de Bagging e Rotation Forest em

qualquer ordem tem um desempenho melhor do que outras combinações.

Os autores Tsai, Hsu e Yen (2014) realizaram um estudo de comparando os classificadores *ensemble* por três técnicas de classificação amplamente utilizados, MLP, SVM e DT. Para a análise, foi utilizado um conjunto de dados de falência de Taiwan, sendo que o resultado da pesquisa demonstra que o desempenho dos classificadores *ensemble* DT é superior aos outros *ensemble*.

O custo computacional médio de formação de *ensemble* DT por Boosting é relativamente baixo, sendo mais eficiente do que *ensemble* SVM por Bagging e que o *ensemble* MLP por Bagging e Boosting.

Os resultados experimentais mostraram que o método *ensemble* Boosting DT composto por 80-100 classificadores apresenta um melhor desempenho (TSAI; HSU; YEN, 2014). Portanto, *ensemble* Boosting DT pode ser considerado como a técnica de *ensemble* de partida em futuros estudos relacionados com classificadores.

Dessa forma, os métodos de classificadores *ensemble* com base DT serão aprofundados no estudo empírico deste trabalho.

5 Estudo Empírico

Neste capítulo são apresentados a base de dados utilizada, a análise empírica realizada, a metodologia e os resultados encontrados.

5.1 Base de Dados

Neste trabalho, foi utilizada uma base de dados com informações de uma instituição financeira brasileira de grande porte atuante no mercado de crédito. O produto analisado corresponde a uma linha de crédito a pessoas físicas, sem destinação específica, com limite pré-aprovado e disponibilizado automaticamente na conta do cliente. O empréstimo concedido apresenta taxa prefixada com pagamentos de prestações mensais.

Os dados de crédito coletados referem-se à base histórica do período de setembro de 2007 a janeiro de 2010 com prazos de contratação de 24 meses. Apesar de os dados não serem recentes, destaca-se que o foco do trabalho envolve o estudo da aplicabilidade dos modelos de aprendizagem de máquina na análise de risco de crédito, a partir de dados do tomador de recursos, sem a preocupação de se considerar situações conjunturais do mercado de crédito como um todo. Em particular, este estudo analisa uma carteira de crédito de elevado nível de risco de crédito, na qual somente 48% dos clientes permaneceram adimplentes até o fim do período contratado.

Os empréstimos contraídos nesse produto têm um valor médio de 1980 reais, sendo que o tomador possui uma idade média de 42 anos e renda bruta mensal média de 1975 reais. Os tomadores possuem, em média, conta corrente há 51 meses e conta de poupança há 63 meses. O saldo médio na conta corrente do tomador é de 652 reais. Dentre os clientes que se tornam

inadimplentes, metade entra em *default* até 386 dias, ou seja, aproximadamente 1 ano, após o início do contrato.

A Tabela 7 apresenta o nome das variáveis da base de dados utilizada no estudo. Vários autores, Garcia, Marques e Sanchez (2012), Kamaloo e Abadeh (2014), Siami, Gholamian e Basiri (2013), Shi, Zhang e Qiu (2013) e Saberi et al. (2013), utilizaram variáveis semelhantes como renda, histórico de dívida, saldo de poupança, emprego, estado civil, número de dependentes para avaliar o risco de crédito utilizando aprendizado de máquina.

Variável	Descrição do Item	Atributos da variável	
SITUAÇÃO RESIDÊNCIA	Retorna código referente à categoria que se enquadra o cliente avaliado.	Não declara	1
		Alugada	2
		Mora com parentes/pais	3
		Outros	4
		Cedida/Funcional	4
		Própria quitada	5
		Financiada por outras entidades	6
REGISTROS SERASA	Retorna código referente à categoria que se enquadra o cliente avaliado.	Não Tem SERASA Regularizado e nem SERASA em aberto	1
		Tem SERASA Regularizado	2
		Tem SERASA aberto	3
EMPREGO FORMAL PJ	Retorna código referente à categoria que se enquadra o cliente avaliado.	Se Não tem	1
		Se tem	2
VALOR - RECEN- TIDADE ADIANTA- MENTO DEPOSITANTE	Retorna valor referente ao intervalo de dias entre o último adiantamento a depositante e a data atual.	≥ 0	
VALOR - CHEQUE SEM FUNDO	Retorna 1 quando houve emissão de cheque sem fundos.	≥ 0	
VALOR - CONTA CORRENTE CONJUNTA	Retorna a quantidade de contas conjuntas na data da avaliação.	≥ 0	
VALOR CONTA POU- PANÇA CONJUNTA	Retorna a quantidade de contas poupança conjuntas na data da avaliação.	≥ 0	
RENDA FORMAL BRUTA	Retorna a soma do valor das rendas formais brutas informadas, provenientes de pessoa jurídica e pessoa física.	≥ 0	
RENDA BRUTA APU- RADA	Retorna o valor total da renda bruta apurada, ou seja a soma de todas as rendas brutas formais e informais, considerando o redutor para renda informal.	≥ 0	
RENDA LÍQUIDA APU- RADA	Retorna o valor total da renda formais líquidas apurada, ou seja a soma de todas as rendas formais líquidas e rendas informais, considerando o redutor para renda informal, se for o caso.	≥ 0	
CONTA SALÁRIO	Retorna código referente à categoria que se enquadra o cliente avaliado.	Conta de Crédito Salário	1
		Sem conta de Crédito Salário	2
VALOR IDADE	Retorna a idade do cliente na data da avaliação.	≥ 0	
TEMPO DE RENDA DE PJ	Retorna o tempo de renda proveniente de Pessoa Jurídica da renda mais antiga em dias.	≥ 0	

Continua na próxima página

Tabela 7 – Continuação da página anterior

Variável	Descrição do Item	Atributos da variável
TEMPO DE RENDA INFORMAL	Retorna o tempo de renda proveniente de Renda Informal / Não comprovada mais antiga em dias.	≥ 0
INSTRUÇÃO	Retorna o código da graduação referente ao grau de instrução de acordo com os atributos das variáveis.	468 Ensino fundamental incompleto
		469 Ensino fundamental completo
		470 Ensino Médio Incompleto
		471 Ensino Médio Completo
		472 Superior Incompleto
		473 Superior Completo
		474 Mestrado
		799 Especialização
		800 Doutorado
		801 Analfabeto
CEP RESIDENCIAL	Retorna as duas primeiras posições referentes ao CEP residencial do cliente.	≥ 0
ESTADO CIVIL	Retorna o estado civil do cliente.	475 Solteiro
		476 Casado(a) com comunhão de bens
		477 Casado(a) com comunhão parcial de bens
		478 Casado(a) com separação de bens
		479 Divorciado
		480 Separado Judicialmente
		481 Viúvo(a)
482 Outros		
QTDE DEPENDENTES COM RENDA	Retorna o número de dependentes com renda do cliente.	≥ 0
QTDE DEPENDENTES SEM RENDA	Retorna o número de dependentes sem renda do cliente.	≥ 0
RENDA LIQUIDA DOS DEPENDENTES	Retorna a soma do valor das rendas dos dependentes informadas.	≥ 0
CONTRATADO VALOR BRUTO	Retorna o valor bruto do empréstimo.	$260 \leq x \leq 10000$
MAX ATRASO	Retorna o número de dias máximo que o cliente atrasou o pagamento.	≥ 0

Tabela 7 – Variáveis utilizadas na análise.

5.1.1 Seleção das Amostras

Após a seleção das variáveis, o banco de dados foi dividido em duas amostras aleatórias, uma para o desenvolvimento do modelo, com 70% (87.237 observações) do banco de dados, e

outra para validação do modelo, com 30% (37.387 observações). As bases de desenvolvimento e de avaliação possuem características similares, apresentando 47,8% e 48,0% de inadimplência respectivamente.

5.2 Metodologia

Este trabalho envolve classificação de crédito entre devedores, não se preocupando exatamente em identificar explicações teóricas para justificar se uma variável afeta positiva ou negativamente a inadimplência. Assim, busca-se identificar modelos de previsão que possam ser gerados por algoritmos, a partir de dados reais, sem uma preocupação com uma eventual fundamentação teórica para a inclusão de uma variável explicativa sobre o potencial de pagamento.

A definição de inadimplência utilizada será referente a posições em atraso de 60 dias, uma vez que esse é o critério utilizado pela instituição financeira para classificar clientes do produto analisado.

A seguir, são apresentadas as técnicas estatísticas e os procedimentos adotados para a classificação de adimplentes e inadimplentes.

5.2.1 Decision Trees

As *Decision Trees* possuem a estrutura de árvore de cima para baixo, o qual divide os dados para criar ramos. O modelo em si compreende em uma série de decisões lógicas, semelhante a um fluxograma, com nós de decisão indicando uma decisão a ser tomada em um atributo. Os ramos indicam as escolhas das decisões (LANTZ, 2013).

Os nós em cada ramo representam tanto classes quanto distribuições de classe. O maior nó em uma árvore é o nó raiz com o maior ganho de informação (TSAI; HSU; YEN, 2014). Após o primeiro nó, um dos nós subsequentes com o maior ganho de informação é então escolhido como

o teste para o próximo nó. Esse processo continua até que todas as variáveis sejam comparadas ou não há variáveis restantes em que as amostras podem ser divididas. Então, a árvore termina em nós que mostram o resultado de seguir uma combinação de decisões, comparando-se classes ou distribuições de classes.

Segundo Tsai, Hsu e Yen (2014), selecionando um caso aleatório de um conjunto S de casos e estabelecendo que ele pertence a alguma classe C_j , a probabilidade de que uma amostra arbitrária pertence à classe C_j é estimada por:

$$P_i = \frac{\text{freq}(C_j, S)}{|S|} \quad (5.1)$$

onde $|S|$ é o número de amostras no conjunto S e assim as informações que transmite é $-\log_2 p_i$ bits.

Considerando uma distribuição de probabilidade discreta $P = p_1, p_2, \dots, p_n$ dada, então a informação transmitida por esta distribuição, também chamada de entropia de P , é conhecida como (TSAI; HSU; YEN, 2014):

$$\text{Info}(P) = \sum_{i=1}^n -p_i \log_2 p_i. \quad (5.2)$$

Se a partição de um conjunto de amostras T com base no valor de um atributo não-categorico X em conjuntos de T_1, T_2, \dots, T_m , então a informação necessária para identificar a classe de um elemento de T torna-se a média ponderada da informação necessária para identificar a classe de um elemento de T_i , isto é, a média ponderada de $\text{Info}(T_i)$.

$$\text{Info}(X, T) = \sum_{i=1}^m \frac{|T_i|}{T} \times \text{Info}(T_i). \quad (5.3)$$

O ganho de informação $Gain(X, T)$ é definido como:

$$Gain(X, T) = Info(T) - Info(X, T). \quad (5.4)$$

O ganho de informação representa a diferença entre a informação necessária para identificar um elemento de T e a informação necessária para identificar um elemento de T após o valor do atributo X ter sido avaliado. Dessa forma, $Gain(X, T)$ é o ganho de informação devido ao atributo X (TSAI; HSU; YEN, 2014).

Segundo (LANTZ, 2013), o método **Random Forest**, que se baseia em conjuntos de Decision Trees, combina versatilidade e potência em uma abordagem de aprendizado de máquina única. O método utiliza apenas uma pequena parte aleatória do conjunto completo de observações, podendo lidar com grandes conjuntos de dados, onde a chamada ‘maldição da dimensionalidade’ pode levar outros modelos a falhar.

Esse método foi defendido por Leo Breiman e Adele Cutler, e combina os princípios de base de Bagging com seleção aleatória de recursos para adicionar diversidade adicional para os modelos de Decision Tree. Após o Random Forest ser gerado, o modelo utiliza uma votação para combinar previsões das árvores.

Segundo Breiman (2001), Random Forest é um classificador que consiste em uma coleção de árvores classificadoras estruturadas $h(x, \Theta_k), k = 1, \dots$ onde os Θ_k são vetores aleatórios independentes e identicamente distribuídos e cada árvore lança um único voto para a classe mais popular a partir dos dados de entrada x .

5.2.2 Support Vector Machines

Support Vector Machine, tendo as suas raízes na estatística de Teoria da Aprendizagem e métodos de otimização, tornou-se uma ferramenta poderosa para resolver os problemas de

máquina de aprendizagem (TIAN; SHI; LIU, 2012).

O objetivo de um SVM é criar uma fronteira plana, chamado de hiperplano. O que leva a partições de dados em ambos os lados razoavelmente homogêneas (LANTZ, 2013).

O primeiro autor a introduzir Support Vector Machines para realizar classificação binária foi Vapnik (1998). No SVM separa-se um conjunto de vetores de treinamento em duas classes diferentes $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$, onde $x_i \in R^d$ denota vetores de características em um espaço d-dimensional e $y_i \in \{-1, 1\}$ é um marcador de classe, para produzir um classificador que tenha boa generalização.

Segundo Tsai, Hsu e Yen (2014), para gerar um modelo SVM, vetores de entrada são mapeados em um novo espaço de características de dimensão superior denotado como $\phi : R_d \rightarrow H^f$, onde, $d < f$.

Constrói-se um hiperplano de separação no novo espaço de características por uma função Kernel $K(x_i, x_j)$ que é o produto dos vetores de entrada x_i e x_j onde

$$K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j). \quad (5.5)$$

Segundo Chen, Ma e Ma (2009), a função Kernel pode estar associada a funções lineares, funções de base radial (RBF), polinômios ou sigmóides. Neste trabalho foram utilizadas funções Linear e RBF:

$$\text{Kernel Linear: } K(x_i, x_j) = x_i^T x_j$$

$$\text{RBF: } K(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / \sigma^2).$$

5.2.3 Bagging

Bagging é um método de classificadores *ensemble*. Segundo Breiman (1996) classificadores são treinados de forma independente por diferentes conjuntos de treinamento através do método de inicialização. Para construí-los é necessário montar k conjuntos de treinamento idênticos, replicar esses dados de treinamento de forma aleatória para construir k redes independentes por re-amostragem com reposição. Em seguida, deve-se agregar as k redes através de um método de combinação apropriada, tal como a maioria de votos.

Essa estratégia é simples, mas pode reduzir a variância quando combinado com as estratégias de geração de bases de aprendizado (WANG et al., 2011).

Bagging é particularmente atraente quando a informação disponível é de tamanho limitado. Segundo Wang et al. (2011), Para garantir que há amostras de treinamento suficientes em cada subconjunto, grandes porções de amostras (75-100%) são colocadas em cada subconjunto. Com isso, os subconjuntos individuais de formação se sobrepõem de forma significativa, com muitos casos fazendo parte da maioria dos subconjuntos e podendo até mesmo aparecer várias vezes num mesmo subconjunto.

A fim de assegurar a diversidade de situações, um *learner* de base relativamente instável é usado para que limites de decisão diferentes possam ser obtidos, considerando-se pequenas perturbações em diferentes amostras de treinamento (WANG et al., 2011).

O resumo do pseudo código do Bagging, proposto por Breiman (1996), é o seguinte:

Entrada: Dataset $D = \{(z_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$;

Número de rounds de aprendizagem T .

Processo:

Para $t = 1, 2, \dots, T$:

a. Forma conjuntos bootstrap de dados S_t selecionando m exemplos aleatórios do conjunto de treinamento com substituição

b. Deixa h_t ser o resultado da base de treinamento do algoritmo baseado em S_t

fim.

Saída:

Classificador combinado:

$$H(x) = \text{maioria}(h_1(x), \dots, h_T(x))$$

5.2.4 Boosting e AdaBoost

A abordagem por Boosting original foi proposta por [Schapire \(1990\)](#). No Boosting, de forma semelhante ao Bagging, cada classificador é treinado usando um conjunto de treinamento diferente.

Segundo [Lantz \(2013\)](#), a principal diferença em relação ao Bagging é que os conjuntos de dados re-amostrados em Boosting são construídos especificamente para gerar aprendizados complementares e a importância do voto é ponderado com base no desempenho de cada modelo, em vez de atribuição de mesmo peso para todos os votos.

Essencialmente, esse procedimento permite aumentar o desempenho de um limiar arbitrário simplesmente adicionando *learners* mais fracos. Dada a utilidade desse achado, Boosting é considerado uma das descobertas mais significativas em aprendizado de máquina ([LANTZ, 2013](#)).

Segundo [Tsai, Hsu e Yen \(2014\)](#), AdaBoost é uma combinação das ideias de Bagging e Boosting e não exige um grande conjunto de treinamento como os outros dois. Inicialmente, cada exemplo de formação de um determinado conjunto de treinamento tem o mesmo peso.

Para treinar o k – *esimo* classificador como um "modelo de aprendizagem fraca", n conjuntos de amostras de treinamento ($n < m$) entre S são usadas para treinar o k – *esimo* classificador. Em seguida, o classificador treinado é avaliado por S para identificar os exemplos de treinamento que não foram classificados corretamente. A rede $k + 1$ é então treinada por um conjunto de treinamento modificado que reforça a importância desses exemplos classificados incorretamente (TSAI; HSU; YEN, 2014).

Este procedimento de amostragem será repetido até que k amostras de treinamento sejam construídas para a construção da k – *esima* rede. Portanto, a decisão final baseia-se na votação ponderada dos classificadores individuais (TSAI; HSU; YEN, 2014).

Embora existam várias versões dos algoritmos de Boosting, o mais utilizado é o AdaBoost (WANG et al., 2011). Por isso, esse algoritmo será utilizado neste estudo.

O pseudo código do Adaboost, segundo Wang et al. (2011), é o seguinte:

Entrada: Dataset $D = \{(z_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$;

Algoritmo de base de aprendizagem L ;

Número de rounds de aprendizagem T .

Processo:

$$D_1(i) = 1/m.$$

% Inicializa a distribuição de pesos

Para $t = 1, 2, \dots, T$:

$$h_t = L(D, D_t);$$

% Treina a base de aprendizado h_t para D usando a distribuição D_t

$$\varepsilon_t = Pr_{i \in D_t}[h_t(x_i) \neq y_i];$$

% Mede o erro de h_t

$$\alpha_t = \frac{1}{2} \ln \frac{1 - \varepsilon_t}{\varepsilon_t}$$

% Determina o peso de h_t

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \begin{cases} \exp(-\alpha_t) & \text{se } h_t(x_i) = y_i \\ \exp(\alpha_t) & \text{se } h_t(x_i) \neq y_i \end{cases}$$

% Atualiza a distribuição

$$= \frac{D_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t}$$

% fator de normalização que permite D_{t+1} ser uma distribuição

fim.

Saída:

$$H(x) = \text{sign} \left(\sum_{t=1}^T \alpha_t h_t(x) \right)$$

5.2.5 Procedimento

Os critérios de avaliação do experimento adotados são as medidas-padrão estabelecidos nos domínios da classificação de crédito (WANG et al., 2011). Estas medidas incluem a precisão média, o erro tipo I e o erro tipo II.

A definição destas medidas pode ser explicada com respeito a uma matriz de confusão como mostrado na Tabela 8.

		Condição Atual	
		Positivo (Risco)	Negativo (Sem Risco)
Resultado do teste	Positivo (Risco)	Verdadeiro Positivo (VP)	Falso Positivo (FP)
	Negativo (Sem Risco)	Falso Negativo (FN)	Verdadeiro Negativo (VN)

Tabela 8 – Matriz de confusão para a pontuação de crédito.

As medidas são definidas da seguinte forma:

$$\text{Precisão média} = \frac{VP + VN}{VP + FP + FN + VN} \quad (5.6)$$

$$\text{Sensibilidade} = 1 - \text{Erro Tipo I} = 1 - \frac{FP}{VP + FP} \quad (5.7)$$

$$\text{Especificidade} = 1 - \text{Erro Tipo II} = 1 - \frac{FN}{FN + VN}. \quad (5.8)$$

A Sensibilidade tem valores próximos a 1 quando o Erro Tipo 1 é baixo e a Especificidade tem valores próximos a 1 quando o Erro do Tipo II é baixo. A Curva ROC (*Receiver Operating Characteristic*) também foi calculada para todos os modelos e, dessa forma, a medida AUC (*Area Under the Curve*), que fornece um critério de precisão para o conjunto de validação, foi utilizada para comparar os modelos encontrados.

5.3 Resultados

5.3.1 Decision Trees

Existem vários algoritmos de Decision Tree, como CART, C4.5, C5.0, ID3, etc. Neste trabalho será utilizado o algoritmo C5.0, desenvolvido por *J. Ross Quinlan* como um aperfeiçoa-

mento do algoritmo C4.5.

Segundo [Lantz \(2013\)](#), o algoritmo C5.0 tornou-se o padrão da indústria para a produção de Decision Trees por ter bons resultados para a maioria dos tipos de problemas quando comparado a outros modelos de aprendizagem de máquina avançados.

O algoritmo C5.0 pode produzir mais de dois sub-grupos em cada divisão, oferecendo divisões não-binárias. A avaliação das separações possíveis é baseado na relação entre o ganho de informação que é uma medida da teoria da informação ([LANTZ, 2013](#)).

Considerando os dados de desenvolvimento do modelo, O algoritmo construiu uma árvore de tamanho igual a 1.974. O que indica que foram feitas 1.974 decisões. A Decision Tree foi então aplicada aos dados de validação e os resultados são apresentados na matriz de confusão,

Tabela 9.

Resposta real	Resposta predita		Total
	Mau	Bom	
Mau	13.758	5.667	19.425
Bom	8.026	9.936	17.962
Total	21.784	15.603	37.387

Tabela 9 – Matriz de Confusão - Decision Tree.

Dos 37.387 observações separadas para avaliação da árvore, o modelo previu corretamente que 13.758 são inadimplentes e 9.936 são adimplentes. Resultando em uma precisão de 63,4%. O resultado na amostra de validação teve pior desempenho do que na amostra de treinamento. Esse fato não é inesperado, uma vez que o desempenho de um modelo é muitas vezes pior em dados invisíveis.

A Tabela 10 apresenta as medidas de desempenho associadas ao modelo proposto.

A Figura 7 apresenta a curva ROC para o modelo proposto. A AUC para o algoritmo C5.0 é 65,7%.

Medida	Valor
Precisão Média	63,4%
Sensibilidade I	70,8%
Especificidade	55,3%

Tabela 10 – Medidas de Desempenho - Decision Tree.

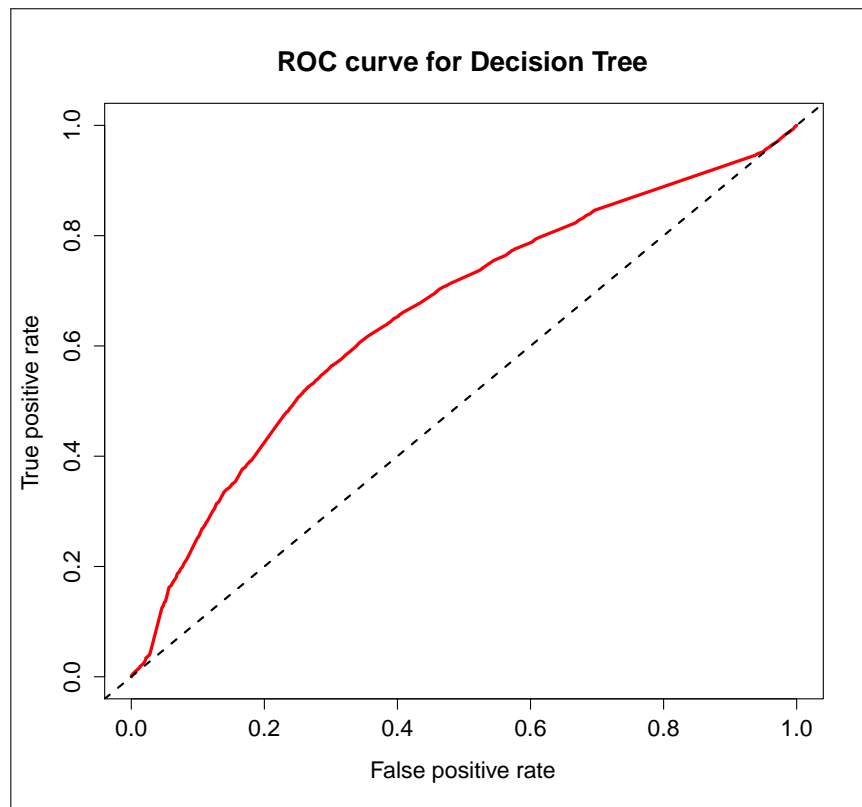


Figura 7 – Curva ROC - Decision Tree C5.0.

5.3.1.1 Random Forest

Foi também feito um modelo utilizando Random Forest. Segundo [Kruppa et al. \(2013\)](#), Random Forest é um conjunto de árvores de decisão, uma generalização de árvores de regressão e classificação. Foi utilizado o pacote *randomForest*, original de Leo Breiman and Adele Cutler, e código no R desenvolvido por Andy Liaw and Matthew Wiener.

A Figura 8 apresenta a variação no erro conforme é aumentado o número de árvores de decisão. É possível notar que há um decréscimo no erro, mas o erro tende a se estabilizar após aproximadamente 60 árvores serem incluídas no modelo.

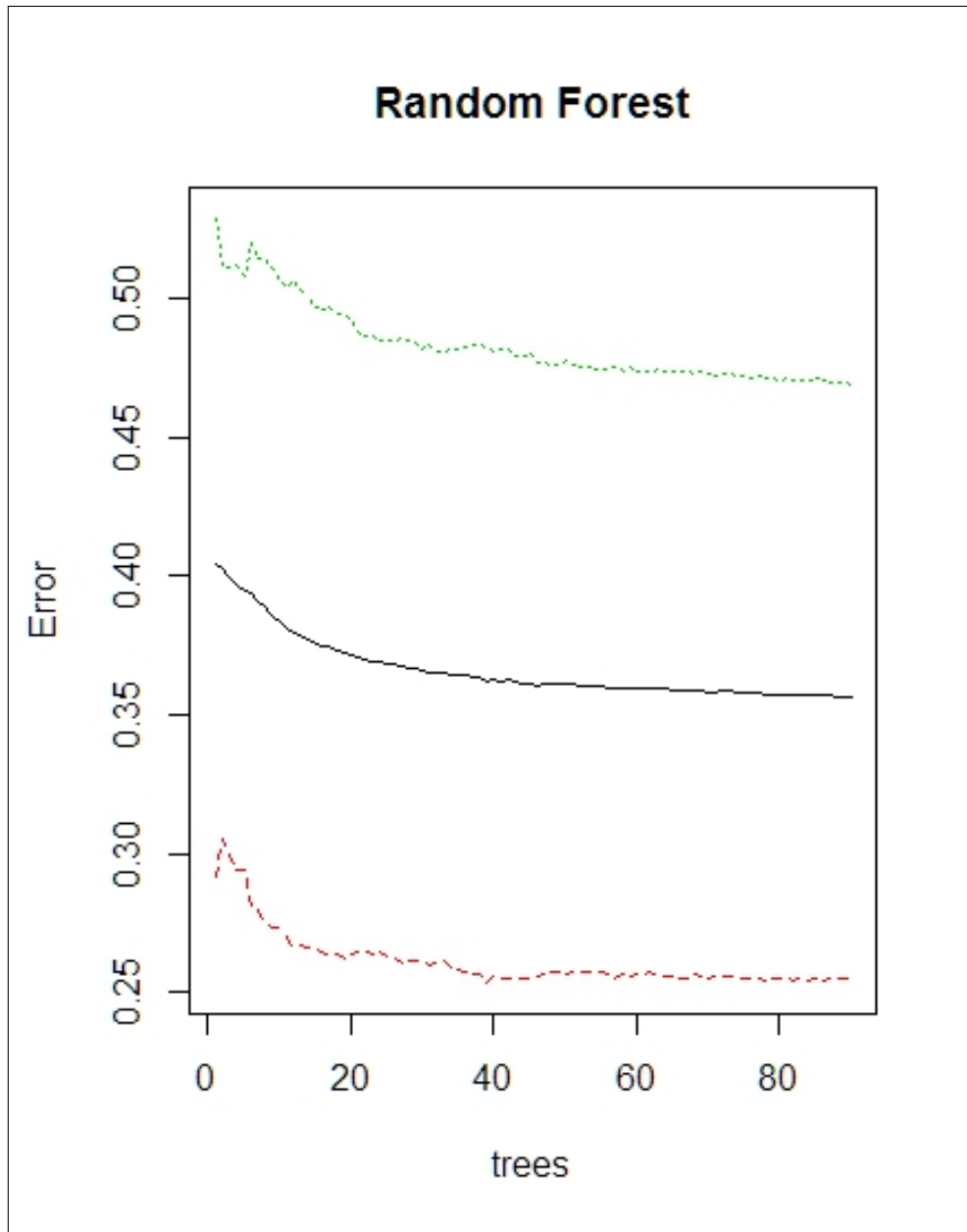


Figura 8 – Erro X Número de árvores - Random Forest.

A Tabela 11 apresenta matriz de confusão do modelo Random Forest aplicado aos dados de validação.

Resposta real	Resposta predita		Total
	Mau	Bom	
Mau	13.569	5.856	19.425
Bom	7.650	10.312	17.962
Total	21.219	16.168	37.387

Tabela 11 – Matriz de Confusão - Random Forest.

Dos 37.387 dados separados para avaliação, o modelo previu corretamente que 13.569 são inadimplentes e 10.312 são adimplentes, resultando em uma precisão de 63,9%.

A Tabela 12 apresenta as medidas de desempenho do modelo Random Forest.

Medida	Valor
Precisão Média	63,9%
Sensibilidade I	69,9%
Especificidade	57,4%

Tabela 12 – Medidas de Desempenho - Random Forest.

A Figura 9 apresenta a curva ROC para o modelo proposto. A AUC para o algoritmo é 68,9%.

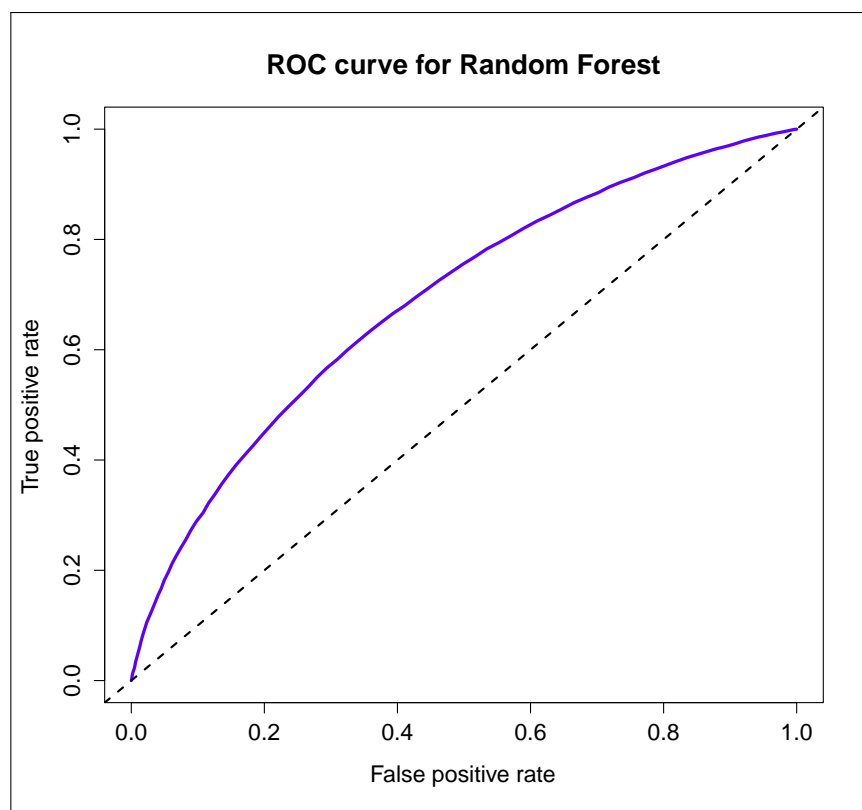


Figura 9 – Curva ROC - Random Forest.

5.3.2 SVM

Foram elaborados dois modelos SVM: um com função Kernel linear e outro com uma função Kernel de Base Radial. O pacote do R utilizado foi o e1071 que oferece uma interface do

C++ implementado por Chih-Chung Chang and Chih-Jen Lin.

Um Kernel K é uma função que recebe dois pontos x_i e x_j do espaço de entradas e computa o produto escalar desses dados no espaço de características.

A escolha correta dos parâmetros do Kernel é crucial para a obtenção de bons resultados, dessa forma foi utilizada a função `tune.svm` para encontrar os melhores parâmetros para o algoritmo.

5.3.2.1 Kernel Linear

O SVM com função Kernel Linear apresenta fronteiras lineares para a separação de dados pertencentes a duas classes.

A Tabela 13 apresenta matriz de confusão do modelo SVM-linear aplicado aos dados de validação.

Resposta real	Resposta predita		
	Mau	Bom	Total
Mau	14.356	5.069	19.425
Bom	8.444	9.518	17.962
Total	22.800	14.587	37.387

Tabela 13 – Matriz de Confusão - SVM Linear.

Dos 37.387 dados separados para avaliação, o modelo previu corretamente que 14.356 são inadimplentes e 9.518 são adimplentes. Resultando em uma precisão de 63,9%.

A Tabela 14 apresenta as medidas de desempenho do modelo SVM com kernel Linear.

Medida	Valor
Precisão Média	63,9%
Sensibilidade I	73,9%
Especificidade	53,0%

Tabela 14 – Medidas de Desempenho - SVM Linear.

A Figura 10 apresenta a curva ROC para o modelo proposto. A AUC para o algoritmo é 31,67%.

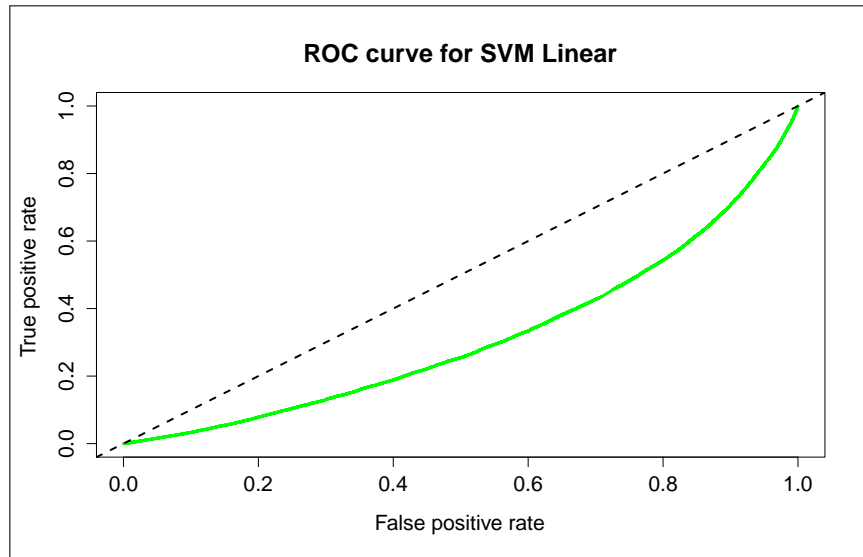


Figura 10 – Curva ROC - SVM Linear.

5.3.2.2 Kernel Radial

Da mesma forma, foi utilizado uma função Kernel de Base Radial. A Tabela 15 apresenta matriz de confusão do modelo SVM-Radial aplicado aos dados de validação.

Resposta real	Resposta predita		Total
	Mau	Bom	
Mau	14.468	4.957	19.425
Bom	8.550	9.412	17.962
Total	23.018	14.369	37.387

Tabela 15 – Matriz de Confusão - SVM Radial.

Das 37.387 observações para avaliação, o modelo previu corretamente que 14.468 são inadimplentes e 9.412 são adimplentes. Resultando em uma precisão média de 63,9%.

A Tabela 16 apresenta as medidas de desempenho do modelo SVM com Kernel Radial.

A Figura 11 apresenta a curva ROC para o modelo proposto. A AUC para o algoritmo é 31,49%.

Medida	Valor
Precisão Média	63,9%
Sensibilidade I	74,5%
Especificidade	52,4%

Tabela 16 – Medidas de Desempenho - SVM Radial.

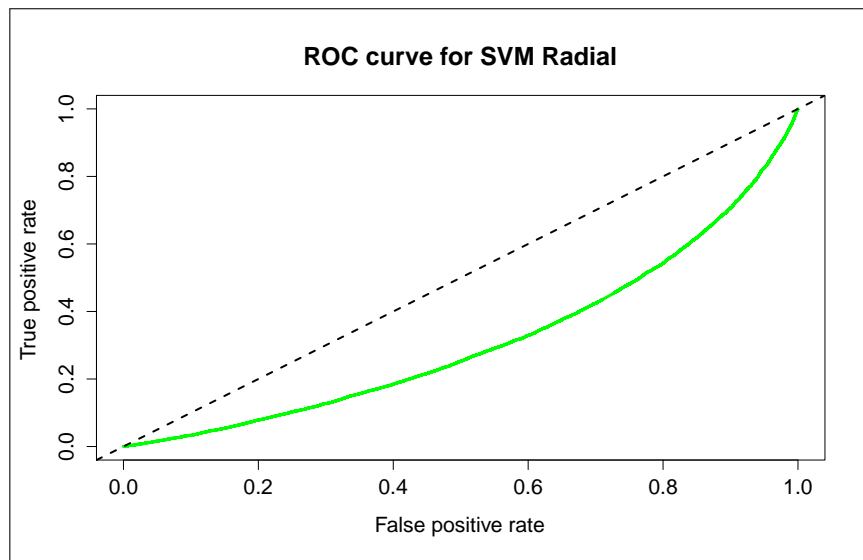


Figura 11 – Curva ROC - SVM Radial.

5.3.3 Bagging

Outro método de classificadores *ensemble* utilizado é o Bagging, esse método gera um conjunto de dados por amostragem *bootstrap* dos dados originais. O conjunto de dados gera um conjunto de modelos utilizando um algoritmo de aprendizagem simples que são combinados por votação simples para classificação.

O pacote *ipred* do *software* R, oferece uma aplicação clássica Bagging em Decision Tree. Para treinar o modelo é utilizada a função `bagging()` (LANTZ, 2013).

O modelo encontrado utilizando Bagging a partir da amostra de desenvolvimento foi aplicado aos dados de validação. Os resultados são apresentados na matriz de confusão, Tabela 17.

Dos 37.387 dados separados para avaliação, o modelo previu corretamente que 13.057

Resposta real	Resposta predita		Total
	Mau	Bom	
Mau	13.057	6.368	19.425
Bom	7.663	10.299	17.962
Total	20.720	16.667	37.387

Tabela 17 – Matriz de Confusão - Bagging.

são inadimplentes e 10.299 são adimplentes. Resultando em uma precisão de 62,5%.

A Tabela 18 apresenta as medidas de desempenho do algoritmo.

Medida	Valor
Precisão Média	62,5%
Sensibilidade I	67,2%
Especificidade	57,3%

Tabela 18 – Medidas de Desempenho - Bagging.

A Figura 12 apresenta a curva ROC para o modelo proposto. A AUC para o algoritmo é 67,1%.

5.3.4 AdaBoost

No algoritmo AdaBoost são construídas várias Decision Trees e então a melhor classe para cada exemplo é escolhida. O melhor modelo encontrado utilizando o AdaBoost foi com 20 tentativas. Foi utilizado o pacote C50 do *software* R.

O modelo encontrado utilizando AdaBoost foi então aplicado aos dados de validação, os resultados são apresentados na matriz de confusão, Tabela 19.

Resposta real	Resposta predita		Total
	Mau	Bom	
Mau	13.359	6.066	19.425
Bom	7.240	10.722	17.962
Total	20.599	16.788	37.387

Tabela 19 – Matriz de Confusão - AdaBoost.

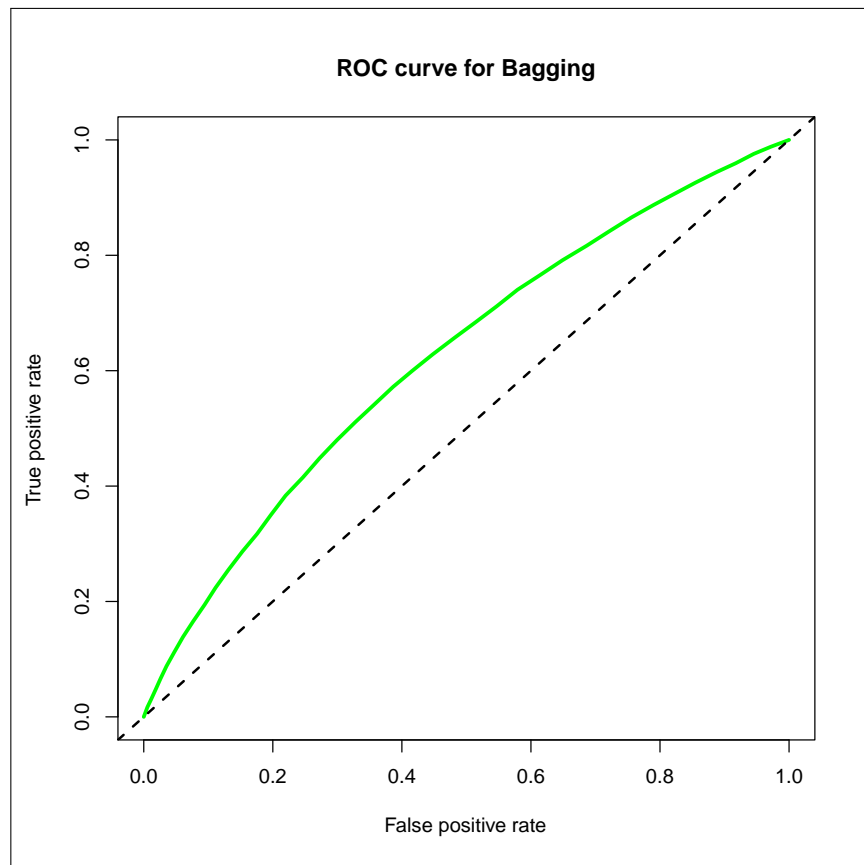


Figura 12 – Curva ROC - Bagging.

Dos 37.387 dados separados para avaliação da árvore, o modelo previu corretamente que 13.359 são inadimplentes e 10.722 são adimplentes. Resultando em uma precisão de 64,4%.

A Tabela 20 apresenta as medidas de desempenho associadas ao modelo proposto.

Medida	Valor
Precisão Média	64,4%
Sensibilidade I	68,8%
Especificidade	59,7%

Tabela 20 – Medidas de Desempenho - AdaBoost.

A Figura 13 apresenta a curva ROC para o modelo proposto. A AUC para o algoritmo é 69,7%.

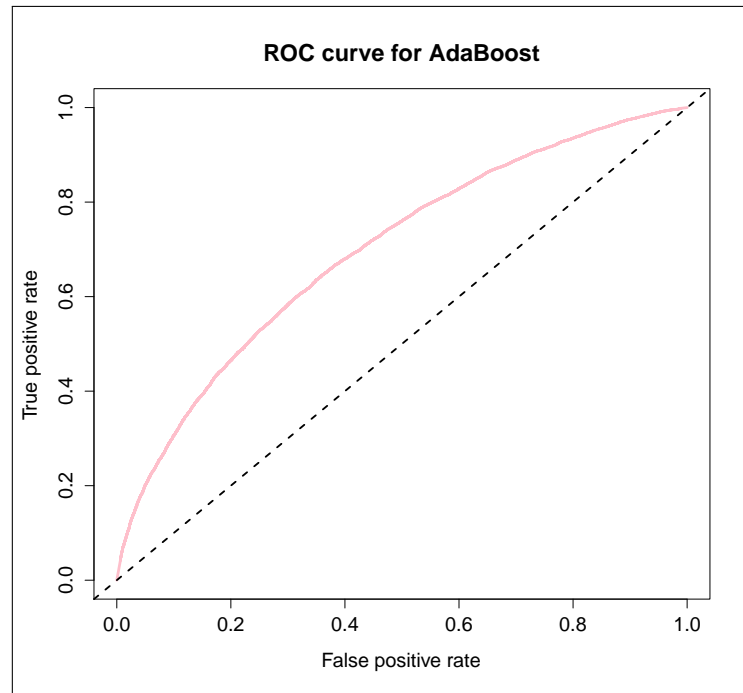


Figura 13 – Curva ROC - AdaBoost.

5.3.5 Regressão Logística

A Regressão Logística é a técnica tradicional mais utilizada atualmente para modelagem em risco de crédito. Dessa forma, foi criado um modelo utilizando Regressão Logística a fim de comparar dos resultados encontrados utilizando as técnicas de aprendizado de máquina.

O modelo encontrado utilizando Regressão Logística foi aplicado aos dados de validação, os resultados são apresentados na matriz de confusão, Tabela 21.

Resposta real	Resposta predita		
	Mau	Bom	Total
Mau	19.163	262	19.425
Bom	16.771	1.191	17.962
Total	35.934	1.453	37.387

Tabela 21 – Matriz de Confusão - Regressão Logística.

Dos 37.387 dados separados para avaliação, o modelo previu corretamente que 19.163 são inadimplentes e 1191 são adimplentes. Resultando em uma precisão de 54,4%.

A Tabela 22 apresenta as medidas de desempenho do algoritmo.

Medida	Valor
Precisão Média	54,4%
Sensibilidade I	98,7%
Especificidade	06,6%

Tabela 22 – Medidas de Desempenho - Regressão Logística.

A Figura 14 apresenta a curva ROC para o modelo proposto. A AUC para o modelo é 52,6%.

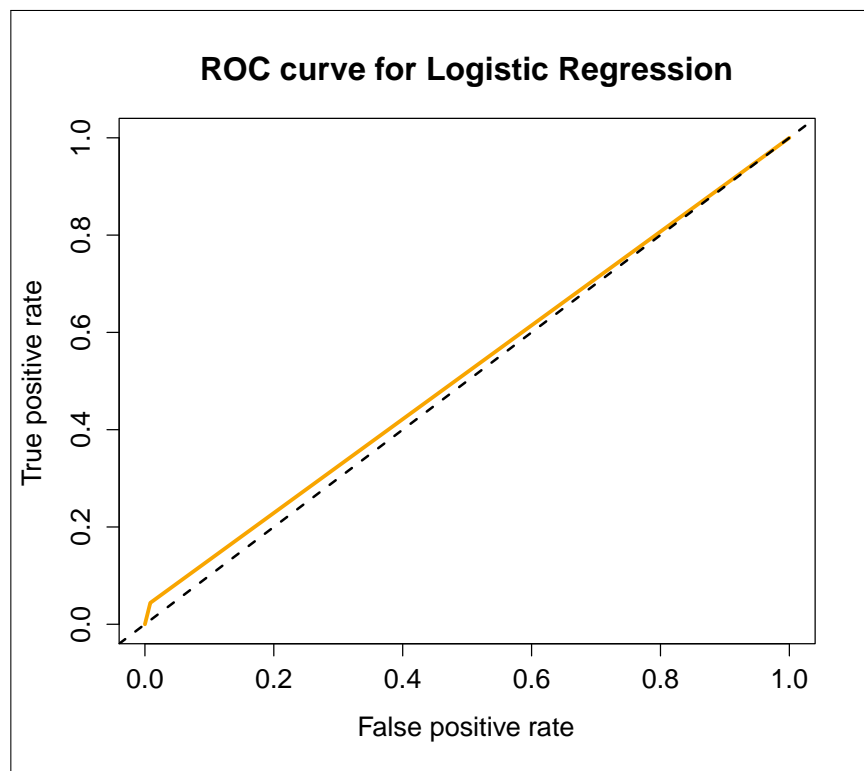


Figura 14 – Curva ROC - Regressão Logística.

5.4 Resultados e Discussão

A Tabela 23 apresenta os valores AUC de cada modelo proposto:

Como já discutido anteriormente, o AUC fornece um critério de precisão para os dados de avaliação. O método que conseguiu melhor AUC foi o AdaBoost, seguido do Random Forest e do Bagging. É interessante observar que os três métodos são baseados em classificadores *ensemble* de estimadores.

	AUC (%)
DT	65,77
RF	68,98
AdaBoost	69,70
Bagging	67,11
SVM L	31,67
SVM R	31,49
Reg. Logística	52,64

Tabela 23 – Resultados AUC.

Os SVMs tiveram um desempenho ruim, sendo piores que a Regressão Logística, técnica estatística tradicional em análise de risco de crédito, na precisão dos dados de avaliação.

A Tabela 24 apresenta os valores das medidas de desempenho, Precisão Média, Sensibilidade e Especificidade nos dados de teste.

	Precisão Média	Sensibilidade	Especificidade
DT	63,37 %	70,83 %	55,32 %
RF	63,88 %	69,85 %	57,41 %
AdaBoost	64,41 %	68,77 %	59,69 %
Bagging	62,47 %	67,22 %	57,34 %
SVM L	63,86 %	73,90 %	52,99 %
SVM R	63,87 %	74,48 %	52,40 %
Reg. Logística	54,44 %	98,65 %	06,63 %

Tabela 24 – Comparação das Medidas de Desempenho.

A Figura 15 compara os valores de Sensibilidade e Especificidade para cada modelo proposto. Apesar da Regressão Logística ter apresentado uma Sensibilidade de 98%, o modelo apresentou uma Especificidade de 6%. Por essa figura é possível notar a superioridade das técnicas de aprendizado de máquina no distanciamento entre os pontos no gráfico dos algoritmos de aprendizado de máquina e do modelo tradicional, Regressão Logística. Enquanto os algoritmos de ML apresentam precisão média de 63%, a Regressão Logística apresenta precisão de 54%.

A Figura 16 compara os valores de Sensibilidade e Especificidade somente para os algoritmos de aprendizado de máquina, mostrando a precisão média de cada um.

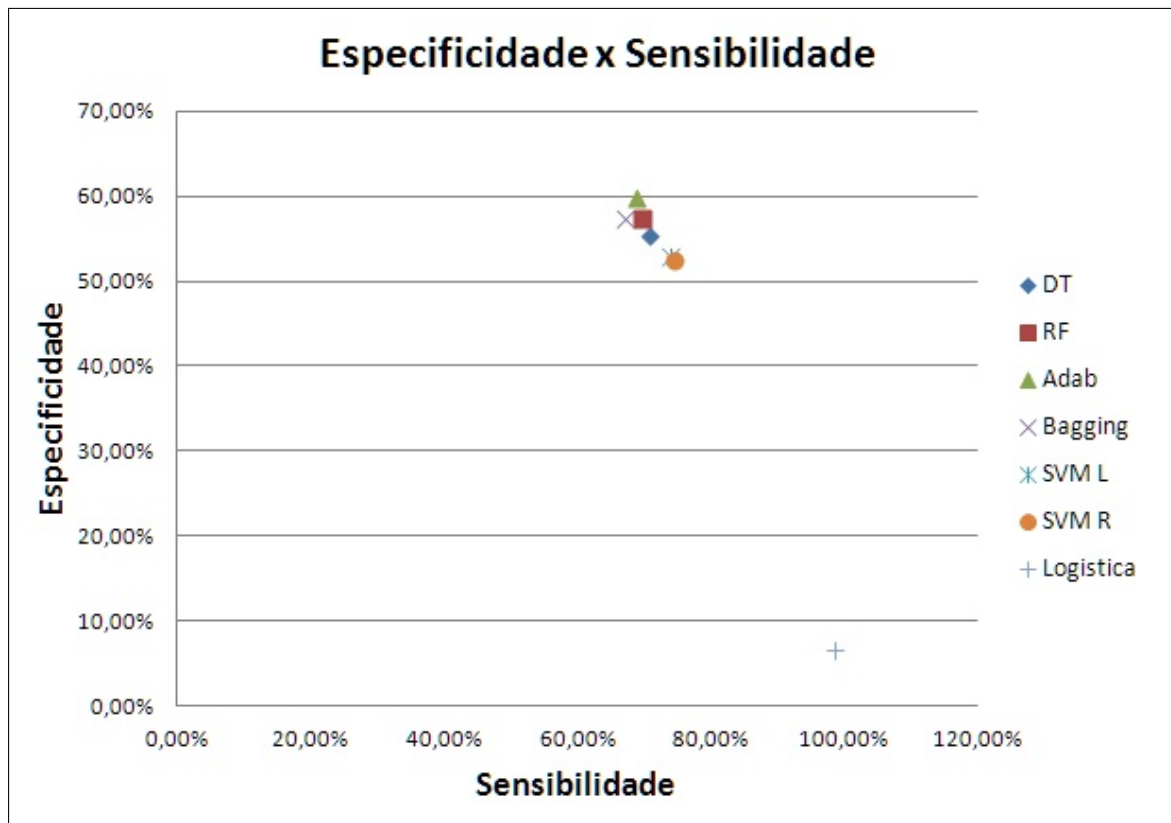


Figura 15 – Especificidade x Sensibilidade.

Os algoritmos SVM apresentaram melhor Sensibilidade, apresentando um Erro do Tipo 1 menor que os outros algoritmos, significando que os SVMs preveem melhor casos de mau tomadores. Porém a baixa Especificidade, mostra que o algoritmo não teve um bom desempenho ao identificar bons tomadores.

O algoritmo AdaBoost teve a melhor Especificidade, seguido de Bagging e Random Forest. Uma especificidade indica um baixo Erro do Tipo 2, e desta forma, o Adaboost foi melhor em identificar bons tomadores.

5.4.1 Diferentes Tamanhos de Amostra

Para verificar a sensibilidade dos modelos propostos, em relação ao tamanho da amostra, os modelos foram reproduzidos em amostras de 100 e 1000 observações.

Para a Decision Tree, o modelo com a base completa foi o de melhor desempenho,

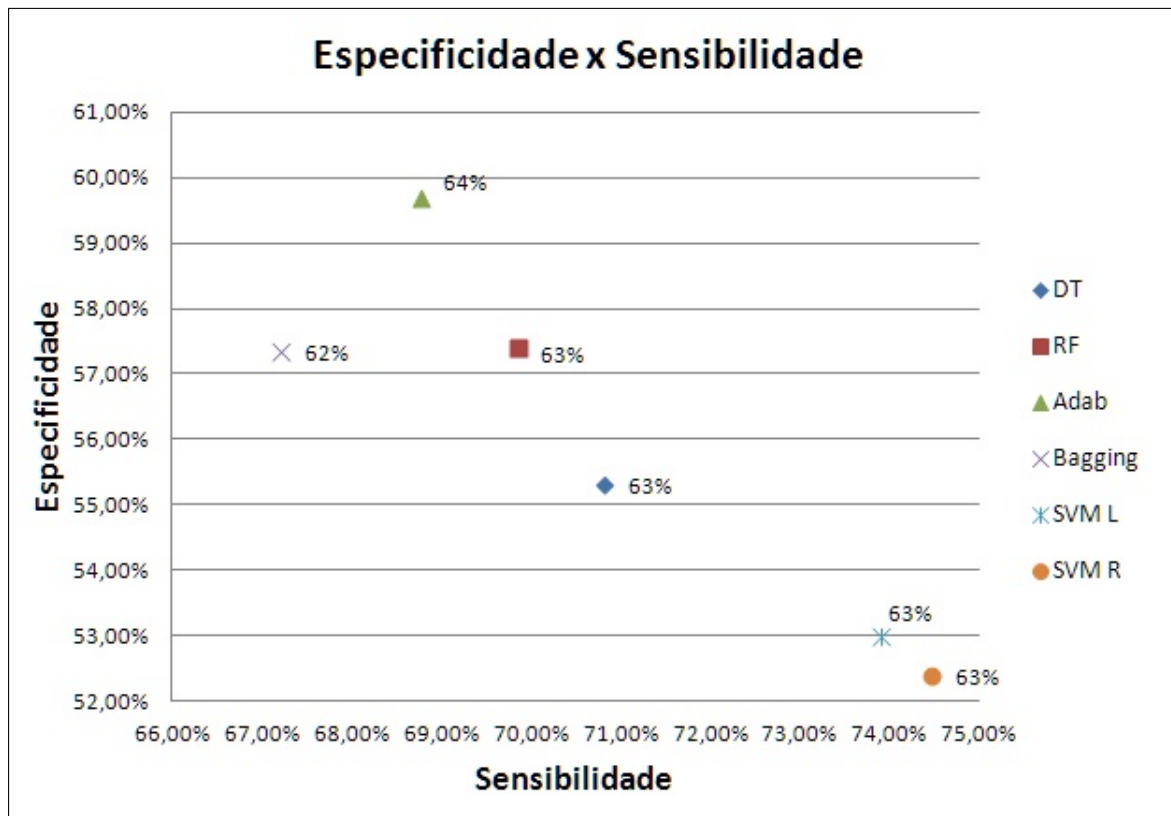


Figura 16 – Especificidade x Sensibilidade.

apresentando maior AUC e maior precisão média. Ambas as medidas decaem conforme o número de observações na amostra diminui. Os valores são apresentados na Figura 17.

Para o Random Forest os modelos com as diferentes amostras são equivalentes. O com a base completa apresentou AUC levemente superior. Os valores são apresentados na Figura 18.

O Bagging apresentou maior AUC conforme o tamanho da amostra aumenta. Porém, a precisão média foi levemente superior na menor amostra, com 100 observações. Os valores são apresentados na Figura 19.

No AdaBoost, o modelo com a base completa foi o de melhor desempenho, apresentando maior AUC e maior precisão média. Ambas as medidas decaem conforme o número de observações na amostra diminui. Os valores são apresentados na Figura 20.

O SVM Linear não apresentou bom desempenho. Comparando as medidas de desempe-

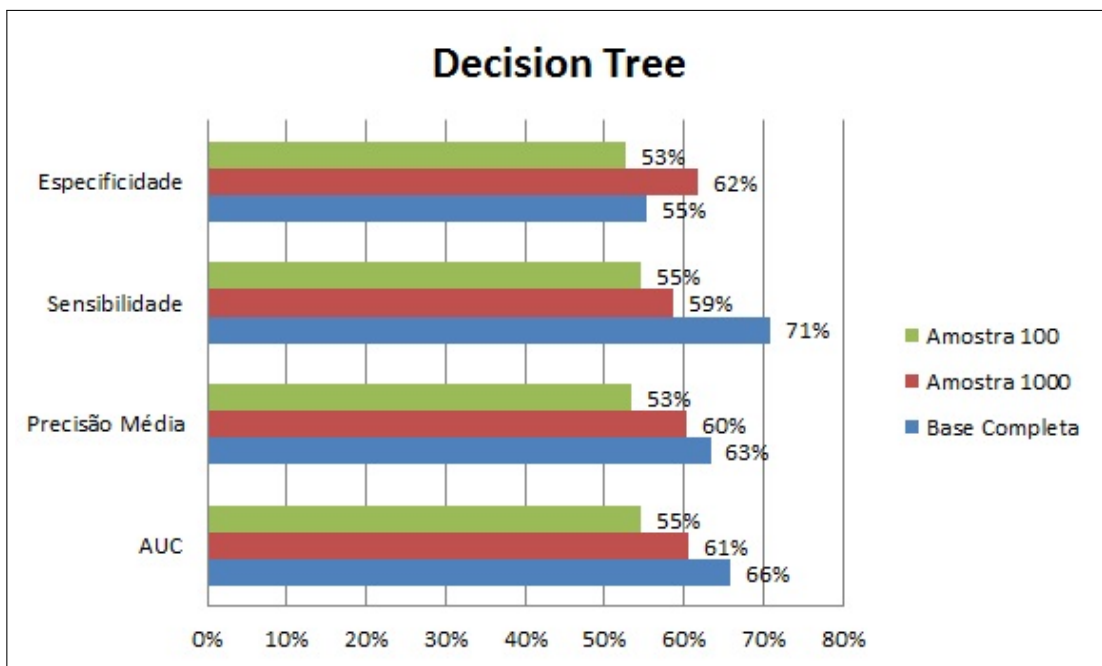


Figura 17 – Comparação das medidas de desempenho em diferentes amostras para Decision Tree.

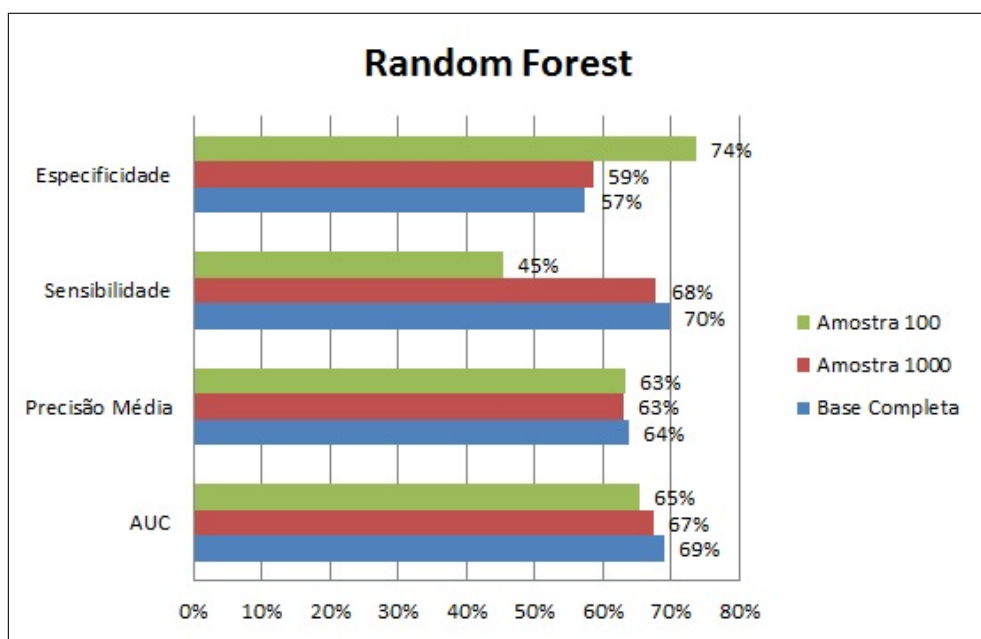


Figura 18 – Comparação das medidas de desempenho em diferentes amostras para Random Forest.

no das diferentes amostras, é possível notar que a menor amostra, com 100 observações, foi a de melhor AUC e com precisão média equivalente a amostra com a base completa.

O SVM Radial também não apresentou bom desempenho. Comparando as medidas de

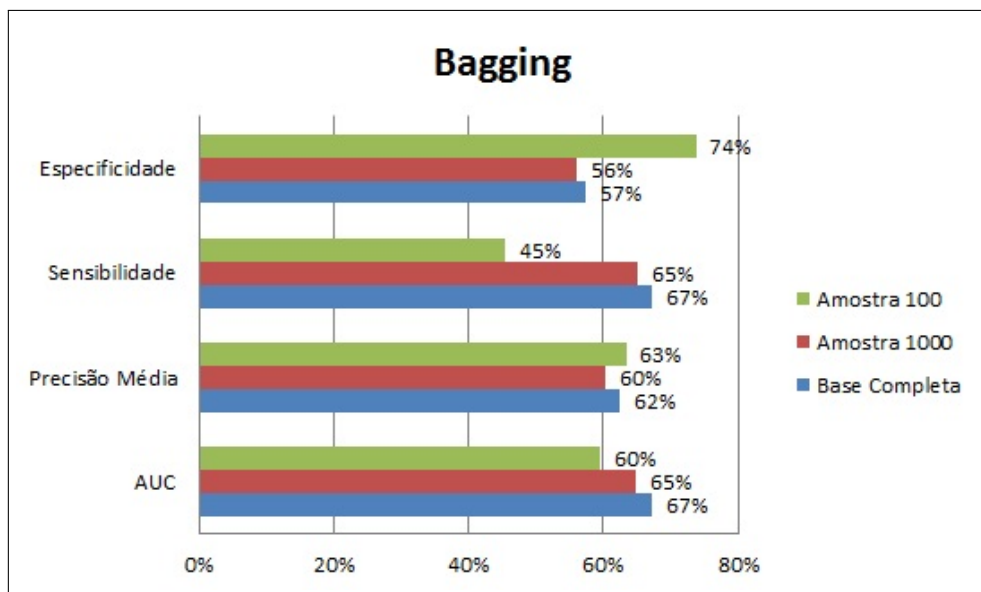


Figura 19 – Comparação das medidas de desempenho em diferentes amostras para Bagging.

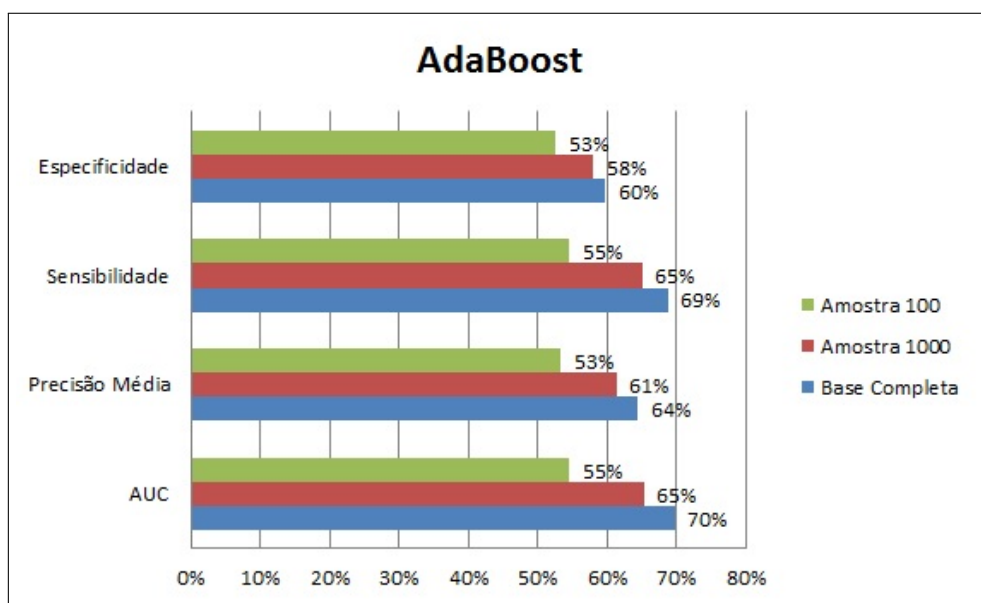


Figura 20 – Comparação das medidas de desempenho em diferentes amostras para AdaBoost.

desempenho das diferentes amostras, é possível notar que a menor amostra, com 100 observações, foi a de melhor AUC e precisão média, porém com sensibilidade igual a zero, isto é, o modelo classificou todos os tomadores como bons pagadores.

A Tabela 25 apresenta as medidas de desempenho AUC e precisão média para as todas as técnicas estudadas.

O AdaBoost foi novamente o algoritmo que apresentou o melhor AUC e melhor precisão

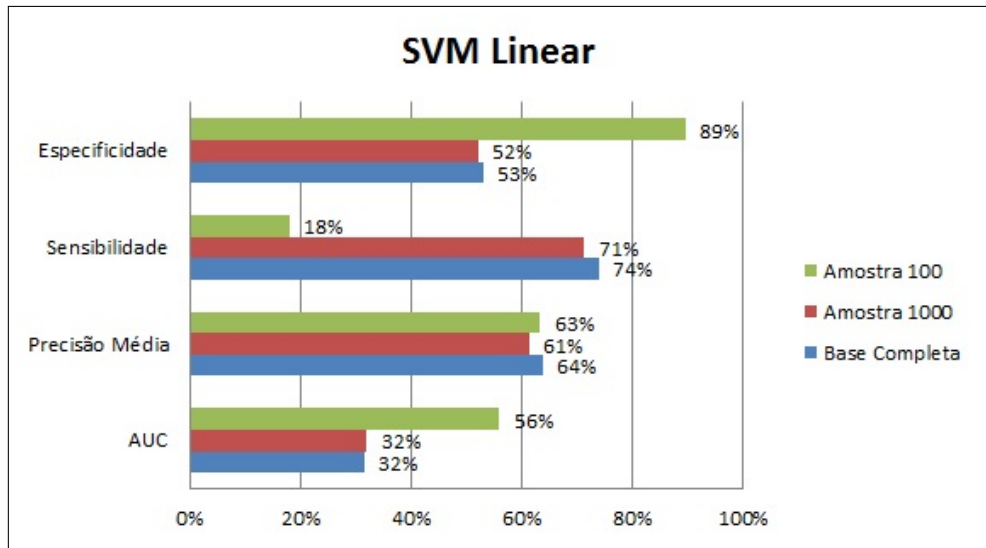


Figura 21 – Comparação das medidas de desempenho em diferentes amostras para SVM Linear.

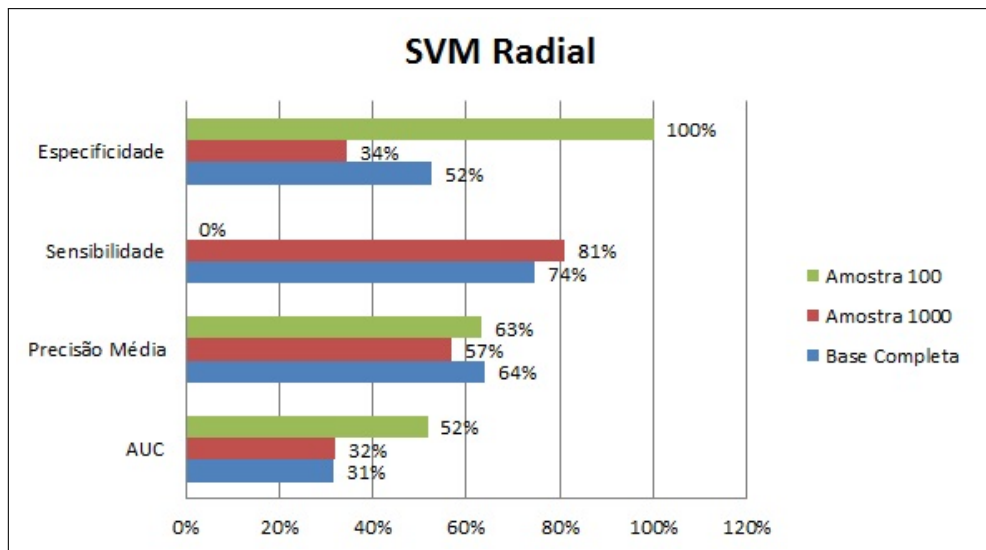


Figura 22 – Comparação das medidas de desempenho em diferentes amostras para SVM Radial.

média. Esses valores foram melhores na amostra considerando toda a base de dados. Considerando a amostra de 1000 observações, o Random Forest possui o melhor AUC, 67,4% e a maior precisão média, 63,3%. Para a amostra de 100 observações, o Random Forest também possui o melhor AUC, 65,3% e possui, juntamente com o Bagging e o SVM a melhor precisão média, 63,3%.

		Base Completa	Amostra 1000	Amostra 100
Decision	AUC	65,77%	60,63%	54,54%
Tree	Precisão Média	63,37%	60,33%	53,33%
Random Forest	AUC	68,98%	67,44%	65,31%
	Precisão Média	63,88%	63,00%	63,33%
AdaBoost	AUC	69,70%	65,31%	54,54%
	Precisão Média	64,41%	61,33%	53,33%
Bagging	AUC	67,11%	64,81%	59,56%
	Precisão Média	62,47%	60,33%	63,33%
SVM L	AUC	31,67%	32,01%	55,98%
	Precisão Média	63,86%	61,33%	63,33%
SVM R	AUC	31,49%	31,87%	51,67%
	Precisão Média	63,87%	56,67%	63,33%

Tabela 25 – Comparação das medidas de desempenho em diferentes amostras.

6 Considerações finais

Aprendizado de máquina, um sub-campo da Inteligência Artificial, tem sido bastante utilizado na avaliação do risco de crédito por apresentar resultados competitivos com os métodos tradicionalmente utilizados. Muitos pesquisadores têm obtido resultados promissores sobre previsão de notação de crédito utilizando diferentes técnicas estatísticas e Inteligência Artificial (HUANG et al., 2004).

Esta dissertação teve como primeiro propósito revisar, classificar e analisar artigos no campo de aprendizado de máquina e risco de crédito por meio de uma revisão sistemática de literatura.

Essa revisão foi conduzida utilizando o método proposto por Lage e Godinho (2010) que é particularmente útil pois permite uma integração do estado-da-arte sobre um tema, além de fornecer uma análise aprofundada dos principais estudos já publicados na área.

O segundo propósito da pesquisa foi realizar um estudo empírico, conduzindo uma aplicação prática das técnicas de aprendizado de máquina mais utilizadas em um banco de dados real. A modelagem foi construída para testar alguns modelos de aprendizagem de máquina no contexto do risco de crédito.

Dos 80 artigos identificados na revisão, 78 apresentam estudo empírico, evidenciando que pesquisadores da área têm se preocupado com o desenvolvimento e teste de técnicas para aprimoramento prático de mecanismos de análise do risco de crédito.

Assim, o foco dos estudos tem se direcionado para a identificação de mecanismos que melhorem o reconhecimento de padrões de crédito em situações reais.

O estudo da literatura indica que existe uma popularização recente das técnicas de aprendizado de máquina em aplicações de crédito, sendo que as técnicas mais utilizadas nos artigos estudados são Support Vector Machine, Decision Tree e Artificial Neural Networks.

Nesta dissertação, foi ainda realizado também um estudo empírico em uma base de dados com informações de uma instituição financeira brasileira utilizando 5 técnicas de aprendizado de máquina, Decision Trees, Random Forest, Support Vector Machines, Bagging e AdaBoost. Foi também utilizada Regressão Logística a fim de comparação.

Segundo o estudo empírico, as técnicas de aprendizado de máquina são mais eficientes que o modelo tradicional baseado em Regressão Logística. Enquanto os algoritmos de ML apresentam precisão média de 63%, a Regressão Logística apresenta precisão de 54%.

O método que conseguiu melhor AUC foi o AdaBoost, seguido do Random Forest e do Bagging, todos métodos de classificadores *ensemble*. Os algoritmos SVM apresentaram melhor Sensibilidade, porém com uma Especificidade baixa. O algoritmo AdaBoost teve a melhor Especificidade, seguido de Bagging e Random Forest. Em termos gerais pode-se dizer que o AdaBoost apresentou o melhor desempenho entre os modelos testados.

Foram também comparadas as medidas de desempenho considerando diferentes tamanhos de amostra para verificar a sensibilidade dos modelos propostos, em relação ao número de observações. Os modelos foram reproduzidos em amostras de 1000 e 100 observações. Nas duas amostras, o Random Forest superou os outros métodos propostos em AUC e precisão média, sendo seguido pelo AdaBoost.

Este trabalho teve como limitações a base utilizada que, como na maioria dos estudos, já possui uma amostra viesada, pois somente se tem dados dos empréstimos concedidos. Ou seja, já houve uma seleção inicial do banco. Os empréstimos analisados contem somente tomadores

que a instituição considerou adequados para recebimento dos empréstimos.

Como sugestão de trabalho futuro, indica-se a análise de diferentes custos de classificação errada. Afinal, classificar um mau tomador como bom é mais custoso do que classificar um bom tomador como ruim.

Outra sugestão envolve comparar os resultados das técnicas de aprendizado de máquina considerando diferentes definições de inadimplência como, por exemplo, 30, 90 e 120 dias de atraso.

Referências

- AHN, K.-J. K. H. [Corporate Credit Rating using Multiclass Classification Models with order Information](#). *World Academy of Science, Engineering and Technology*, 2011.
- ALTMAN, E. I.; SAUNDERS, A. [Credit risk measurement: Developments over the last 20 years](#). *Journal of Banking & Finance*, v. 21, p. 1721–1742, 1998.
- ANGELINI, E.; TOLLO, G. di; ROLI, A. [A neural network approach for credit risk evaluation](#). *The Quarterly Review of Economics and Finance*, Elsevier BV, v. 48, n. 4, p. 733–755, Nov 2008. ISSN 1062-9769.
- ANTONAKIS, A. C.; SFAKIANAKIS, M. E. [Assessing naive Bayes as a method for screening credit applicants](#). *Journal of Applied Statistics*, Informa UK Limited, v. 36, n. 5, p. 537–545, May 2009. ISSN 1360-0532.
- BELLOTTI, T.; CROOK, J. [Support vector machines for credit scoring and discovery of significant features](#). *Expert Systems with Applications*, Elsevier BV, v. 36, n. 2, p. 3302–3308, Mar 2009. ISSN 0957-4174.
- BEN-DAVID, A. [Monotonicity Maintenance in Information-Theoretic Machine Learning Algorithms](#). *Machine Learning*, Springer Science + Business Media, v. 19, n. 1, p. 29–43, 1995. ISSN 0885-6125.
- BEN-DAVID, A.; FRANK, E. [Accuracy of machine learning models versus hand crafted expert systems, A credit scoring case study](#). *Expert Systems with Applications*, Elsevier BV, v. 36, n. 3, p. 5264–5271, Apr 2009. ISSN 0957-4174.
- BREIMAN, L. [Bagging Predictors](#). *Machine Learning*, Springer Science + Business Media, v. 24, n. 2, p. 123–140, 1996. ISSN 0885-6125.
- BREIMAN, L. [Random Forests](#). *Machine Learning*, v. 45, n. 1, p. 5–32, 2001. ISSN 1573-0565.
- CHEN, F.-L.; LI, F.-C. [Combination of feature selection approaches with SVM in credit scoring](#). *Expert Systems with Applications*, v. 37, n. 7, p. 4902–4909, 2010. Cited By 31.
- CHEN, W.; MA, C.; MA, L. [Mining the customer credit using hybrid support vector machine technique](#). *Expert Systems with Applications*, v. 36, n. 4, p. 7611–7616, 2009. Cited By 30.
- CHEN, Y.-S.; CHENG, C.-H. [Hybrid models based on rough set classifiers for setting credit rating decision rules in the global banking industry](#). *Knowledge-Based Systems*, Elsevier BV, v. 39, p. 224–239, Feb 2013. ISSN 0950-7051. Disponível em: <<http://dx.doi.org/10.1016/j.knosys.2012.11.004>>.
- DAMODARAN, A. *Applied Corporate Finance*. 3. ed. [S.l.]: John Wiley & Sons, 2010, 2010.
- DANENAS, P.; GARSVA, G.; GUDAS, S. [Credit Risk Evaluation Model Development Using Support Vector Based Classifiers](#). *Procedia Computer Science*, Elsevier BV, v. 4, p. 1699–1707, 2011. ISSN 1877-0509.

- DAVIS, R.; EDELMAN, D.; GAMMERMAN, A. [Machine-learning algorithms for credit-card applications](#). *IMA Journal of Management Mathematics*, IMA, v. 4, n. 1, p. 43–51, 1992.
- DOUMPOS, M.; ZOPOUNIDIS, C. [Model combination for credit risk assessment: A stacked generalization approach](#). *Annals of Operations Research*, Springer Science + Business Media, v. 151, n. 1, p. 289–306, Nov 2006. ISSN 1572-9338.
- 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. Cited By 7.
- FLOREZ-LOPEZ, R. [Modelling of insurers rating determinants. An application of machine learning techniques and statistical models](#). *European Journal of Operational Research*, Elsevier BV, v. 183, n. 3, p. 1488–1512, Dec 2007. ISSN 0377-2217.
- GALINDO, J.; TAMAYO, P. [Credit Risk Assessment Using Statistical and Machine Learning: Basic Methodology and Risk Modeling Applications](#). *Computational Economics*, Springer Science + Business Media, v. 15, n. 1/2, p. 107–143, 2000. ISSN 0927-7099.
- GARCIA, V.; MARQUES, A.; SANCHEZ, J. [On the use of data filtering techniques for credit risk prediction with instance-based models](#). *Expert Systems with Applications*, Elsevier BV, v. 39, n. 18, p. 13267–13276, Dec 2012. ISSN 0957-4174.
- GESTEL, T. V. et al. [A process model to develop an internal rating system: Sovereign credit ratings](#). *Decision Support Systems*, Elsevier BV, v. 42, n. 2, p. 1131–1151, Nov 2006. ISSN 0167-9236.
- GONEN, G. B.; GONEN, M.; GARGEN, F. [Probabilistic and discriminative group-wise feature selection methods for credit risk analysis](#). *Expert Systems with Applications*, Elsevier BV, v. 39, n. 14, p. 11709–11717, Oct 2012. ISSN 0957-4174.
- GREGORY, J. *Counterparty Credit Risk and Credit Value Adjustment: A Continuing Challenge for Global Financial Markets*. 2. ed. [S.l.]: Wiley, 2012.
- GUNDUZ, Y.; UHRIG-HOMBURG, M. [Predicting credit default swap prices with financial and pure data-driven approaches](#). *Quantitative Finance*, Informa UK Limited, v. 11, n. 12, p. 1709–1727, Dec 2011. ISSN 1469-7696. Disponível em: <<http://dx.doi.org/10.1080/14697688.2010.531041>>.
- HAJEK, P.; MICHALAK, K. [Feature selection in corporate credit rating prediction](#). *Knowledge-Based Systems*, Elsevier BV, v. 51, p. 72–84, Oct 2013. ISSN 0950-7051. Disponível em: <<http://dx.doi.org/10.1016/j.knosys.2013.07.008>>.
- HAND, D. J.; ADAMS, N. M. [Defining attributes for scorecard construction in credit scoring](#). *Journal of Applied Statistics*, Informa UK Limited, v. 27, n. 5, p. 527–540, Jul 2000. ISSN 1360-0532. Disponível em: <<http://dx.doi.org/10.1080/02664760050076371>>.
- HASHEMI, R. et al. [A hybrid intelligent system for predicting bank holding structures](#). *European Journal of Operational Research*, Elsevier BV, v. 109, n. 2, p. 390–402, Sep 1998. ISSN 0377-2217. Disponível em: <[http://dx.doi.org/10.1016/S0377-2217\(98\)00065-4](http://dx.doi.org/10.1016/S0377-2217(98)00065-4)>.
- HOFFMANN, F. et al. [Inferring descriptive and approximate fuzzy rules for credit scoring using evolutionary algorithms](#). *European Journal of Operational Research*, v. 177, n. 1, p. 540–555, 2007. Cited By 34.

- HSIEH, N.-C. [Hybrid mining approach in the design of credit scoring models](#). *Expert Systems with Applications*, v. 28, n. 4, p. 655–665, 2005. Cited By 78.
- HUANG, C.-L.; CHEN, M.-C.; WANG, C.-J. [Credit scoring with a data mining approach based on support vector machines](#). *Expert Systems with Applications*, v. 33, n. 4, p. 847–856, 2007. Cited By 194.
- HUANG, J.-J.; TZENG, G.-H.; ONG, C.-S. [Two-stage genetic programming \(2SGP\) for the credit scoring model](#). *Applied Mathematics and Computation*, Elsevier BV, v. 174, n. 2, p. 1039–1053, Mar 2006. ISSN 0096-3003. Disponível em: <http://dx.doi.org/10.1016/j.amc.2005.05.027>.
- HUANG, Z. et al. [Credit rating analysis with support vector machines and neural networks: a market comparative study](#). *Decision Support Systems*, Elsevier BV, v. 37, n. 4, p. 543–558, Sep 2004. ISSN 0167-9236. Disponível em: [http://dx.doi.org/10.1016/S0167-9236\(03\)00086-1](http://dx.doi.org/10.1016/S0167-9236(03)00086-1).
- JABBOUR, C. J. C. [Environmental training in organisations: From a literature review to a framework for future research](#). *Resources, Conservation and Recycling*, Elsevier BV, v. 74, p. 144–155, May 2013. ISSN 0921-3449. Disponível em: <http://dx.doi.org/10.1016/j.resconrec.2012.12.017>.
- JORION, P. *Value at Risk - The new benchmark for managing financial risk*. [S.l.]: The McGraw-Hill Companies, 2007.
- KAMALLOO, E.; ABADEH, M. S. [Credit Risk Prediction Using Fuzzy Immune Learning](#). *Advances in Fuzzy Systems*, Hindawi Publishing Corporation, v. 2014, p. 1–11, 2014. ISSN 1687-711X. Disponível em: <http://dx.doi.org/10.1155/2014/651324>.
- KHANDANI, A. E.; KIM, A. J.; LO, A. W. [Consumer credit-risk models via machine-learning algorithms](#). *Journal of Banking & Finance*, Elsevier BV, v. 34, n. 11, p. 2767–2787, Nov 2010. ISSN 0378-4266. Disponível em: <http://dx.doi.org/10.1016/j.jbankfin.2010.06.001>.
- KIM, K.-j.; AHN, H. [A corporate credit rating model using multi-class support vector machines with an ordinal pairwise partitioning approach](#). *Computers & Operations Research*, Elsevier BV, v. 39, n. 8, p. 1800–1811, Aug 2012. ISSN 0305-0548. Disponível em: <http://dx.doi.org/10.1016/j.cor.2011.06.023>.
- KRUPPA, J. et al. [Consumer credit risk: Individual probability estimates using machine learning](#). *Expert Systems with Applications*, Elsevier BV, v. 40, n. 13, p. 5125–5131, Oct 2013. ISSN 0957-4174. Disponível em: <http://dx.doi.org/10.1016/j.eswa.2013.03.019>.
- LAGE, M. J.; GODINHO, M. F. [Variations of the kanban system: Literature review and classification](#). *International Journal of Production Economics*, Elsevier BV, v. 125, n. 1, p. 13–21, May 2010. ISSN 0925-5273. Disponível em: <http://dx.doi.org/10.1016/j.ijpe.2010.01.009>.
- LANTZ, B. *Machine Learning with R*. [S.l.]: Packt Publishing Ltd, 2013.
- LEE, T.-S. et al. [Credit scoring using the hybrid neural discriminant technique](#). *Expert Systems with Applications*, v. 23, n. 3, p. 245–254, 2002. Cited By 130.
- LEE, Y.-C. [Application of support vector machines to corporate credit rating prediction](#). *Expert Systems with Applications*, Elsevier BV, v. 33, n. 1, p. 67–74, Jul 2007. ISSN 0957-4174.

- LIN, S. [A new two-stage hybrid approach of credit risk in banking industry](#). *Expert Systems with Applications*, v. 36, n. 4, p. 8333–8341, 2009. Cited By 28.
- LIN, W.-Y.; HU, Y.-H.; TSAI, C.-F. [Machine Learning in Financial Crisis Prediction: A Survey](#). *IEEE Trans. Syst., Man, Cybern. C*, Institute of Electrical & Electronics Engineers (IEEE), v. 42, n. 4, p. 421–436, Jul 2012. ISSN 1558-2442.
- LIU, X.; FU, H.; LIN, W. [A Modified Support Vector Machine model for Credit Scoring](#). *International Journal of Computational Intelligence Systems*, Informa UK Limited, v. 3, n. 6, p. 797–804, Dec 2010. ISSN 1875-6883.
- LIU, Y.; SCHUMANN, M. [Data mining feature selection for credit scoring models](#). *J Oper Res Soc*, Nature Publishing Group, v. 56, n. 9, p. 1099–1108, Apr 2005. ISSN 1476-9360.
- LUO, S.-T.; CHENG, B.-W.; HSIEH, C.-H. [Prediction model building with clustering-launched classification and support vector machines in credit scoring](#). *Expert Systems with Applications*, Elsevier BV, v. 36, n. 4, p. 7562–7566, May 2009. ISSN 0957-4174.
- MARIANO, E. B.; SOBREIRO, V. A.; REBELATTO, D. A. d. N. [Human development and data envelopment analysis: A structured literature review](#). *Omega*, Elsevier BV, v. 54, p. 33–49, Jul 2015. ISSN 0305-0483.
- MARINAKIS, Y. et al. [Ant colony and particle swarm optimization for financial classification problems](#). *Expert Systems with Applications*, Elsevier BV, v. 36, n. 7, p. 10604–10611, Sep 2009. ISSN 0957-4174.
- MARQUES, A.; GARCIA J.A, V.; SANCHEZ, J. [Two-level classifier ensembles for credit risk assessment](#). *Expert Systems with Applications*, Elsevier BV, v. 39, n. 12, p. 10916–10922, Sep 2012. ISSN 0957-4174.
- MARQUES, A.; GARCIA, V.; SANCHEZ, J. [Exploring the behaviour of base classifiers in credit scoring ensembles](#). *Expert Systems with Applications*, Elsevier BV, v. 39, n. 11, p. 10244–10250, Sep 2012. ISSN 0957-4174.
- MARTENS, D. et al. [Comprehensible credit scoring models using rule extraction from support vector machines](#). *European Journal of Operational Research*, v. 183, n. 3, p. 1466–1476, 2007. Cited By 122.
- MITCHELL, T. M. *Machine Learning*. 1. ed. New York, NY, USA: McGraw-Hill, Inc., 1997. ISBN 0070428077, 9780070428072.
- MUES, C. et al. [Decision diagrams in machine learning: an empirical study on real-life credit-risk data](#). *Expert Systems with Applications*, Elsevier BV, v. 27, n. 2, p. 257–264, Aug 2004. ISSN 0957-4174.
- NIKLIS, D.; DOUMPOS, M.; ZOPOUNIDIS, C. [Combining market and accounting-based models for credit scoring using a classification scheme based on support vector machines](#). *Applied Mathematics and Computation*, Elsevier BV, v. 234, p. 69–81, May 2014. ISSN 0096-3003.
- ONG, C.; HUANG, J.; TZENG, G. [Building credit scoring models using genetic programming](#). *Expert Systems with Applications*, Elsevier BV, v. 29, n. 1, p. 41–47, Jul 2005. ISSN 0957-4174.

- ORESKI, S.; ORESKI, G. [Genetic algorithm-based heuristic for feature selection in credit risk assessment](#). *Expert Systems with Applications*, Elsevier BV, v. 41, n. 4, p. 2052–2064, Mar 2014. ISSN 0957-4174.
- PIRAMUTHU, S. [Evaluating feature selection methods for learning in data mining applications](#). *European Journal of Operational Research*, Elsevier BV, v. 156, n. 2, p. 483–494, Jul 2004. ISSN 0377-2217.
- PIRAMUTHU, S. [On preprocessing data for financial credit risk evaluation](#). *Expert Systems with Applications*, Elsevier BV, v. 30, n. 3, p. 489–497, Apr 2006. ISSN 0957-4174.
- SABERI, M. et al. [A granular computing-based approach to credit scoring modeling](#). *Neurocomputing*, Elsevier BV, v. 122, p. 100–115, Dec 2013. ISSN 0925-2312.
- SANCHEZ, C. P.; MONELLOS, P. de L.; LOPEZ, M. R. [Proporciona la auditoria evidencias para detectar y evaluar tensiones financieras latentes? Un diagnostico comparativo mediante tecnicas econométricas e inteligencia artificial](#). *Revista Europea de Direccion y Economia de la Empresa*, Elsevier BV, v. 22, n. 3, p. 115–130, Jul 2013. ISSN 1019-6838.
- SCHAPIRE, R. E. <http://dx.doi.org/10.1007/BF00116037>The strength of weak learnability. *Mach Learn*, Springer Science + Business Media, v. 5, n. 2, p. 197–227, Jun 1990. ISSN 1573-0565.
- SEOW, H.-V.; THOMAS, L. [Using adaptive learning in credit scoring to estimate take-up probability distribution](#). *European Journal of Operational Research*, v. 173, n. 3, p. 880–892, 2006. Cited By 7.
- SETIONO, R.; BAESSENS, B.; MUES, C. [A note on knowledge discovery using neural networks and its application to credit card screening](#). *European Journal of Operational Research*, v. 192, n. 1, p. 326–332, 2009. Cited By 17.
- SHI, J.; ZHANG, S.-y.; QIU, L.-m. [Credit scoring by feature-weighted support vector machines](#). *Journal of Zhejiang University SCIENCE C*, Journal of Zhejiang University SCIENCE, v. 14, n. 3, p. 197–204, Mar 2013. ISSN 1869-196X.
- SIAMI, M.; GHOLAMIAN, M. R.; BASIRI, J. [An application of locally linear model tree algorithm with combination of feature selection in credit scoring](#). *International Journal of Systems Science*, Informa UK Limited, v. 45, n. 10, p. 2213–2222, Jan 2013. ISSN 1464-5319.
- TANG, T.-C.; CHI, L.-C. [Predicting multilateral trade credit risks: Comparisons of Logit and Fuzzy Logic models using ROC curve analysis](#). *Expert Systems with Applications*, v. 28, n. 3, p. 547–556, 2005. Cited By 31.
- TIAN, Y.; SHI, Y.; LIU, X. [Recent advances on support vector machines research](#). *Technological and Economic Development of Economy*, Vilnius Gediminas Technical University, v. 18, n. 1, p. 5–33, Mar 2012. ISSN 2029-4921.
- TSAI, C.; WU, J. [Using neural network ensembles for bankruptcy prediction and credit scoring](#). *Expert Systems with Applications*, Elsevier BV, v. 34, n. 4, p. 2639–2649, May 2008. ISSN 0957-4174.
- TSAI, C.-F. [Financial decision support using neural networks and support vector machines](#). *Expert Systems*, Wiley-Blackwell, v. 25, n. 4, p. 380–393, Sep 2008. ISSN 1468-0394.

- TSAI, C.-F. [Financial decision support using neural networks and support vector machines](#). *Expert Systems*, Wiley-Blackwell, v. 25, n. 4, p. 380–393, Sep 2008. ISSN 1468-0394.
- TSAI, C.-F. [Combining cluster analysis with classifier ensembles to predict financial distress](#). *Information Fusion*, Elsevier BV, v. 16, p. 46–58, Mar 2014. ISSN 1566-2535.
- TSAI, C.-F.; CHEN, M.-L. [Credit rating by hybrid machine learning techniques](#). *Applied Soft Computing*, Elsevier BV, v. 10, n. 2, p. 374–380, Mar 2010. ISSN 1568-4946.
- TSAI, C.-F.; HSU, Y.-F.; YEN, D. C. [A comparative study of classifier ensembles for bankruptcy prediction](#). *Applied Soft Computing*, Elsevier BV, v. 24, p. 977–984, Nov 2014. ISSN 1568-4946.
- TSAKONAS, A.; DOUNIAS, G. [EVOLVING NEURAL-SYMBOLIC SYSTEMS GUIDED BY ADAPTIVE TRAINING SCHEMES: APPLICATIONS IN FINANCE](#). *Applied Artificial Intelligence*, Informa UK Limited, v. 21, n. 7, p. 681–706, Aug 2007. ISSN 1087-6545.
- TWALA, B. [Multiple classifier application to credit risk assessment](#). *Expert Systems with Applications*, Elsevier BV, v. 37, n. 4, p. 3326–3336, Apr 2010. ISSN 0957-4174.
- TWALA, B. [Impact of noise on credit risk prediction: Does data quality really matter?](#) *Intelligent Data Analysis*, IOS Press, v. 17, n. 6, p. 1115–1134, 2013.
- VAPNIK, V. N. *Statistical Learning Theory*. [S.l.]: Wiley, 1998.
- VARETTO, F. [Genetic algorithms applications in the analysis of insolvency risk](#). *Journal of Banking & Finance*, Elsevier BV, v. 22, n. 10-11, p. 1421–1439, Oct 1998. ISSN 0378-4266.
- WANG, G. et al. [A comparative assessment of ensemble learning for credit scoring](#). *Expert Systems with Applications*, Elsevier BV, v. 38, n. 1, p. 223–230, Jan 2011. ISSN 0957-4174.
- WANG, G. et al. [Two credit scoring models based on dual strategy ensemble trees](#). *Knowledge-Based Systems*, Elsevier BV, v. 26, p. 61–68, Feb 2012. ISSN 0950-7051.
- WANG, Y.; WANG, S.; LAI, K. [A new fuzzy support vector machine to evaluate credit risk](#). *IEEE Trans. Fuzzy Syst.*, Institute of Electrical & Electronics Engineers (IEEE), v. 13, n. 6, p. 820–831, Dec 2005. ISSN 1063-6706.
- WEST, D. [Neural network credit scoring models](#). *Computers and Operations Research*, v. 27, n. 11-12, p. 1131–1152, 2000. Cited By 298.
- WON, C.; KIM, J.; BAE, J. K. [Using genetic algorithm based knowledge refinement model for dividend policy forecasting](#). *Expert Systems with Applications*, Elsevier BV, v. 39, n. 18, p. 13472–13479, Dec 2012. ISSN 0957-4174.
- WU, H.-C.; HU, Y.-H.; HUANG, Y.-H. [Two-stage credit rating prediction using machine learning techniques](#). *Kybernetes*, Emerald, v. 43, n. 7, p. 1098–1113, Jul 2014. ISSN 0368-492X.
- WU, W.-W. [IMPROVING CLASSIFICATION ACCURACY AND CAUSAL KNOWLEDGE FOR BETTER CREDIT DECISIONS](#). *International Journal of Neural Systems*, World Scientific Pub Co Pte Lt, v. 21, n. 04, p. 297–309, Aug 2011. ISSN 1793-6462.
- XU, X.; ZHOU, C.; WANG, Z. [Credit scoring algorithm based on link analysis ranking with support vector machine](#). *Expert Systems with Applications*, v. 36, n. 2 PART 2, p. 2625–2632, 2009. Cited By 23.

- YANG, Y. [Adaptive credit scoring with kernel learning methods](#). *European Journal of Operational Research*, Elsevier BV, v. 183, n. 3, p. 1521–1536, Dec 2007. ISSN 0377-2217.
- YU, L.; WANG, S.; LAI, K. K. [An intelligent-agent-based fuzzy group decision making model for financial multicriteria decision support: The case of credit scoring](#). *European Journal of Operational Research*, Elsevier BV, v. 195, n. 3, p. 942–959, Jun 2009. ISSN 0377-2217.
- ZHANG, D. et al. [Vertical bagging decision trees model for credit scoring](#). *Expert Systems with Applications*, Elsevier BV, v. 37, n. 12, p. 7838–7843, Dec 2010. ISSN 0957-4174.
- ZHONG, H. et al. [Comparing the learning effectiveness of BP, ELM, I-ELM, and SVM for corporate credit ratings](#). *Neurocomputing*, Elsevier BV, v. 128, p. 285–295, Mar 2014. ISSN 0925-2312.
- ZHOU, X.; JIANG, W.; SHI, Y. [Credit risk evaluation by using nearest subspace method](#). *Procedia Computer Science*, Elsevier BV, v. 1, n. 1, p. 2449–2455, May 2010. ISSN 1877-0509.
- ZHOU, X. et al. [Credit risk evaluation with kernel-based affine subspace nearest points learning method](#). *Expert Systems with Applications*, Elsevier BV, v. 38, n. 4, p. 4272–4279, Apr 2011. ISSN 0957-4174.
- ZOPOUNIDIS, C.; DOUMPOS, M. [Multi-group discrimination using multi-criteria analysis: Illustrations from the field of finance](#). *European Journal of Operational Research*, Elsevier BV, v. 139, n. 2, p. 371–389, Jun 2002. ISSN 0377-2217.

Apêndices

APÊNDICE A – Classificação e codificação dos estudos analisados

Study	Country	Categories											
		1	2	3	4	5	6	7	8	9	10	11	12
Wang, Wang e Lai (2005)	China	B	C	E	C	AB	ABC	ABCD	C	E	AB	BD	A
Marinakos et al. (2009)	Greece	B	I	E	C	B	C	ABD	C	A	B	B	A
Niklis, Doumpos e Zopounidis (2014)	Greece	B	L	E	B	B	C	D	A	B	B	B	B
Zhong et al. (2014)	Singapore	B	CDKL	D	D	B	C	D	C	E	E	A	B
Kruppa et al. (2013)	Germany	B	AHI	B	D	C	E	AC	B	E	C	B	B
Khandani, Kim e Lo (2010)	United States	B	E	E	C	C	AB	ABD	C	B	E	A	A
Galindo e Tamayo (2000)	United States	B	DEHIL	E	D	C	B	ABCD	C	A	B	C	B
Zhou et al. (2011)	China	B	I	C	B	AB	AB	ABCD	C	E	B	ABD	B
Kamalloo e Abadeh (2014)	Iran	B	J	F	B	A	AB	ABCD	B	E	B	BD	B
Mues et al. (2004)	Belgium	B	D	C	B	AB	B	ABCD	C	E	B	B	A
Davis, Edelman e Gammerman (1992)	Scotland	B	D	F	B	C	A	ABC	D	E	B	B	A
Florez-Lopez (2007)	Spain	B	E	BC	B	B	C	D	C	A	A	B	B
Zopounidis e Doumpos (2002)	Greece	B	L	A	C	B	ABCD	ABD	B	E	A	ABD	A
Twala (2010)	South Africa	B	ADEFL	AE	D	AB	AB	ABCD	B	AB	AB	ABD	B
Gunduz e Uhrig-Homburg (2011)	Germany	B	C	E	D	B	C	F	D	E	F	F	B
Tian, Shi e Liu (2012)	China	A	C	F	B	D	F	F	D	F	F	F	B
Lin, Hu e Tsai (2012)	Taiwan	C	M	F	E	D	F	F	D	F	F	F	B
Huang et al. (2004)	United States	B	CD	D	D	B	C	D	C	C	A	AD	B
Kim e Ahn (2012)	SouthKorea	B	C	E	CD	B	C	AD	C	A	B	D	A
Piramuthu (2004)	United States	B	E	C	B	B	AC	AD	B	A	B	A	B
Chen e Cheng (2013)	Taiwan	B	L	E	D	B	AD	D	C	C	C	ABD	B
Hajek e Michalak (2013)	Czech Republic	B	CDFHIL	A	D	B	C	D	C	A	B	AB	B
Piramuthu (2006)	United States	AB	E	E	B	AB	ABC	ABCD	C	A	B	ABD	B
Lee (2007)	South Korea	B	C	ADE	A	B	C	D	C	B	B	D	B
Gonen, Gonen e Gargen (2012)	Turkey	B	L	E	B	A	AB	ABCD	B	E	B	BD	B
Tsai e Chen (2010)	Taiwan	B	EF	ACDE	B	C	B	ABC	A	A	C	D	B
Zhou, Jiang e Shi (2010)	China	B	I	E	B	C	A	F	C	E	C	A	B
Danenas, Garsva e Gudas (2011)	Lithuania	B	C	A	B	B	C	AD	B	A	B	A	B
Ben-David e Frank (2009)	Israel	B	ABDEFH	E	D	C	B	ABC	E	E	A	D	A
Tsai, Hsu e Yen (2014)	Taiwan	B	CDE	E	D	A	AB	ABCD	B	E	B	BD	B
Ahn (2011)	Republic of Korea	B	CD	B	B	A	C	D	B	A	B	D	B
Yang (2007)	Germany	B	L	E	C	A	B	F	C	E	C	B	B
Saberi et al. (2013)	Iran	B	D	E	AC	A	B	AB	B	E	B	B	B
Gestel et al. (2006)	Belgium	B	C	B	CD	B	D	ABCD	C	C	A	ABCDE	B
Doumpos e Zopounidis (2006)	Greece	B	L	E	D	AB	AC	ABD	A	E	B	B	B
ONG, HUANG e TZENG (2005)	Taiwan	B	G	BCD	AD	A	AB	ABCD	B	E	B	BD	B
Tsai (2014)	Taiwan	B	L	BCD	CD	A	AB	ABCD	B	E	B	BD	B
Shi, Zhang e Qiu (2013)	China	B	C	E	B	A	AB	ABCD	B	E	B	BD	B
Tsai (2008a)	Taiwan	B	D	E	D	A	AB	ABCD	B	E	B	BD	B
Tsai (2008b)	Taiwan	B	CD	F	A	A	AB	ABCD	B	E	B	BD	B
WU (2011)	Taiwan	B	F	F	B	AB	B	AB	B	E	B	B	B
Wu, Hu e Huang (2014)	Taiwan	BC	A	E	BD	B	C	ABCD	B	C	C	D	B
Ben-David (1995)	Israel	B	E	F	B	C	CE	F	A	E	A	D	A
Luo, Cheng e Hsieh (2009)	Taiwan	B	CL	E	B	A	AB	ABCD	B	E	B	BD	B
Wang et al. (2012)	China	B	AE	ABDE	B	A	AB	ABCD	B	E	B	BD	B
Huang, Tzeng e Ong (2006)	Taiwan	B	G	A	B	A	AB	ABCD	B	E	B	BD	B
Won, Kim e Bae (2012)	Republic of Korea	B	G	E	B	B	C	ACD	A	D	A	D	B
TSAI e WU (2008)	Taiwan	B	D	D	B	A	AB	ABCD	B	E	B	BD	B
Zhang et al. (2010)	China	B	AE	F	B	A	AB	ABCD	B	E	B	BD	B
Liu, Fu e Lin (2010)	China	B	C	E	B	A	AB	ABCD	B	E	B	BD	B
Siami, Gholamian e Basiri (2013)	Iran	B	E	F	B	A	AB	ABCD	B	E	B	BD	B
Angelini, Tollo e Roli (2008)	Italy	B	D	F	B	C	C	ABD	B	A	A	B	B
Antonakis e Sfakianakis (2009)	Greece	B	F	ABCDE	D	AB	AB	ABC	B	E	B	BD	C
Liu e Schumann (2005)	China	B	DEI	B	D	C	B	A	C	E	C	B	B
Tsakonas e Dounias (2007)	Greece	B	D	F	B	A	AB	ABCD	B	E	B	BD	B
Twala (2013)	South Africa	B	DEFI	A	D	A	ABC	ABCD	B	E	B	ABD	B
Oreski e Oreski (2014)	Croatia	B	DG	F	B	A	C	ABD	B	A	B	B	A
Hashemi et al. (1998)	USA	B	D	D	A	D	F	F	D	A	F	A	C
Garcia, Marques e Sanchez (2012)	Spain	B	L	L	D	AB	ABC	ABCD	B	E	B	BD	B
Marques, Garcia J.a e Sanchez (2012)	Spain	B	AB	F	B	AB	AB	ABCD	B	E	B	ABD	B
Yu, Wang e Lai (2009)	China	B	J	F	C	AB	B	ABCD	B	E	B	BD	A
Marques, Garcia e Sanchez (2012)	Spain	B	BDEF	E	D	AB	AB	ABCD	B	E	B	ABD	B
Sanchez, Monelos e Lopez (2013)	Spain	B	D	B	D	B	C	AB	B	C	A	B	B
Varetto (1998)	Italy	B	G	A	D	C	C	E	D	D	B	B	B
Wang et al. (2011)	China	B	ABL	BCDE	D	AC	ABC	ABCD	B	A	B	BD	B
Bellotti e Crook (2009)	Scotland	B	C	AB	D	C	A	AB	B	A	B	B	B
Chen e Li (2010)	Taiwan	B	C	E	D	A	AB	ABCD	B	E	B	BD	B
Chen, Ma e Ma (2009)	China	B	C	C	B	C	B	AB	C	E	E	D	B
Finlay (2009)	UK	B	G	BE	B	C	B	AB	C	A	C	B	B
Hoffmann et al. (2007)	Germany	B	J	E	BD	AB	AB	ABCD	B	E	B	BD	B
Hsieh (2005)	Taiwan	B	D	E	B	A	AB	ABCD	B	E	B	BD	B
Huang, Chen e Wang (2007)	Taiwan	B	C	BCD	BD	A	AB	ABCD	B	E	B	BD	B
Lee et al. (2002)	Taiwan	B	D	A	D	C	A	ABD	A	E	C	D	B
Lin (2009)	Taiwan	B	D	B	C	B	E	BD	C	A	A	D	B
Martens et al. (2007)	Belgium	B	C	F	B	AB	AB	ABCD	B	E	B	BD	B
Seow e Thomas (2006)	UK	B	F	E	B	C	E	E	A	F	E	F	B
Setiono, Baesens e Mues (2009)	Singapore	B	D	C	B	D	AB	F	C	E	B	F	B
Tang e Chi (2005)	Taiwan	B	J	B	D	B	C	CD	C	B	B	D	A
West (2000)	United States	B	D	ABDE	D	A	AB	ABCD	B	E	B	BD	B
Xu, Zhou e Wang (2009)	China	B	C	E	B	A	A	ABCD	B	E	A	D	B

Tabela 26 – Classificação e codificação dos estudos analisados.

APÊNDICE B – Breve Descrição dos Objetivos e Resultados de Cada Artigo Analisado

Tabela 27 – Breve descrição dos objetivos e resultados de cada artigo analisado.

Artigo	Breve Resumo
Wang, Wang e Lai (2005)	O artigo propõe um novo fuzzy SVM para discriminar credores. O novo fuzzy SVM tem maior capacidade de generalização enquanto preserva insensibilidade a outliers podendo encontrar aplicação promissora na análise de crédito e alcançando melhor resultado que SVM padrão.
Marinakis et al. (2009)	Nesse trabalho são apresentados dois algoritmos para a seleção de recursos baseado em colônias de formigas e otimização por enxame de partículas. Esses algoritmos são combinados com três classificadores baseados em <i>Nereast Neighbour</i> . Os resultados obtidos indicam o alto desempenho dos algoritmos propostos em um conjunto reduzido de recursos com alta precisão.
Niklis, Doumpos e Zopounidis (2014)	O objetivo desse artigo é criar um modelo através de uma metodologia de aprendizado de máquina que combina dados de contabilidade com a abordagem baseada em opções de Black, Scholes e Merton. Mesmo em condições problemáticas do mercado de ações (mercado grego na última década), a previsibilidade dos modelos baseados no mercado é muito competitivo com modelos de <i>credit rating</i> tradicionais.

continua na próxima página

continuação da Tabela 4

Zhong et al. (2014)	Esse trabalho realiza um estudo comparativo experimental abrangente sobre a eficácia de quatro algoritmos de aprendizagem, BP, ELM, I-ELM, e SVM. Os resultados experimentais mostram que as abordagens baseadas em SLFN superam as SVM em confiabilidade, enquanto as SVM são melhores em distribuições de saída do que as outras.
Kruppa et al. (2013)	O objetivo desse trabalho é aplicar métodos de aprendizado de máquina usando <i>random forests</i> , <i>k-nearest neighbors</i> e <i>bagged k-nearest neighbors</i> . Quando aplicado a um grande conjunto de dados de classificação de crédito, <i>random forests</i> superou uma regressão logística afinada.
Khandani, Kim e Lo (2010)	O objetivo deste artigo é aplicar técnicas de aprendizado de máquina em uma medida cardinal de risco de crédito. Combinando fatores de crédito tradicionais, como razões da dívida em relação ao rendimento, com as operações bancárias de consumo. As previsões de aprendizado de máquina são consideravelmente mais adaptáveis, e são capazes de pegar a dinâmica da mudança de ciclos de crédito tão bem quanto os níveis absolutos das taxas de falha .
Galindo e Tamayo (2000)	O artigo introduz uma metodologia de modelagem específica com base no estudo de curvas de erro. Usando o estado da arte de técnicas de modelagem, os autores constroem mais de 9.000 modelos como parte do estudo. Os resultados mostram que os modelos de árvore de decisão CART fornecem melhor estimativa para a falha. Redes Neurais possui o segundo melhor resultado.

continua na próxima página

continuação da Tabela 4

Zhou et al. (2011)	Nesse trabalho é proposto uma nova abordagem para a avaliação do risco de crédito, método de aprendizado baseado em kernel, o chamado kernel affine subspace nearest point (KASNP). Os resultados experimentais mostram que o KASNP proposto é um método de classificação para a avaliação de risco eficaz e é um método de aprendizagem promissor para outros problemas de classificação.
Kamalloo e Abadeh (2014)	Esse trabalho propõe um novo classificador usando os princípios do sistema imunológico e regras fuzzy. Os resultados indicam que o sistema de classificação proposto é preciso na detecção de riscos de crédito.
Mues et al. (2004)	O artigo avalia empiricamente em que medida diagramas de decisão são capazes de fornecer uma descrição visual compacta. Poderosos conjunto de regras classificadoras foram obtidas utilizando técnicas de extração de rede neural.
Davis, Edelman e Gammerman (1992)	O objetivo desse trabalho é aplicar uma série de algoritmos para avaliação de cartão de crédito. Apesar do pequeno número de exemplos de treinamento e conjunto de teste, ambas as técnicas (Inferência bayesiana sem assumir independência, modelo G&T, e algoritmos de rede neural) se saíram bem.
Florez-Lopez (2007)	Esse trabalho propõe um modelo de três etapas para analisar os determinantes de rating. Em relação aos modelos multivariados, a Gini árvore de decisão oblíqua CART parece ser uma boa estratégia, proporcionando um equilíbrio bastante satisfatório entre precisão e compreensão.
Zopounidis e Doumpos (2002)	O artigo propõe um novo método para alcançar a discriminação multi-grupo com base em um processo de segmentação binária iterativo. Esse processo proporciona alta precisão e acurácia em problemas de discriminação multi-grupos. A comparação com a análise discriminante demonstrou que a abordagem proposta pode superar o método de discriminação.

continua na próxima página

continuação da Tabela 4

Twala (2010)	Esse artigo explora o comportamento previsto de cinco classificadores para diferentes tipos de ruído em termos de precisão da previsão de risco de crédito e como essa precisão pode ser melhorada usando conjuntos de classificadores. A avaliação experimental mostra que a técnica do conjunto de classificadores tem o potencial de melhorar a precisão da predição.
Gunduz e Uhrig-Homburg (2011)	Empregando o modelo Merton, o artigo analisa a capacidade de <i>credit default swaps</i> (CDS) na previsão de preços dos CDS de empresas na mesma classe de risco. Embora o Merton e os modelos de intensidade lidem com taxas de temporização e juros de mora de forma diferente, o desempenho de previsão em série transversal é, em média, similar.
Tian, Shi e Liu (2012)	O objetivo desse trabalho é compreender SVM do ponto de vista da otimização, rever vários modelos de otimização de representação em SVM, suas aplicações em economia, a fim de promover os interesses de pesquisa em ambos. Pesquisadores e engenheiros em mineração de dados, especialmente em SVM podem se beneficiar dessa pesquisa em compreender melhor a essência da relação entre SVM e otimização.
Lin, Hu e Tsai (2012)	Esse trabalho revisa 130 artigos no período entre 1995 e 2010, com foco no desenvolvimento de técnicas de aprendizado de máquina. Os autores concluem que o desenvolvimento de modelos para a previsão de crises financeiras que utilizam o aprendizado de máquina ainda tem um longo caminho a percorrer.
Huang et al. (2004)	Usando backpropagation de rede neural (BNN), esse artigo apresenta uma técnica relativamente nova de aprendizado de máquina, SVM, para fornecer um modelo com melhor poder explicativo. Os resultados mostraram que o SVM alcança uma precisão comparável a de BNN.

continua na próxima página

continuação da Tabela 4

Kim e Ahn (2012)	Os autores propõem uma abordagem computacional para MSVMs que leva em conta as características ordinais para lidar com várias classes ordinais de forma eficiente e eficaz. Os resultados mostraram que o modelo proposto melhora o desempenho de classificação, em comparação com outras técnicas de classificação típica multi-classe e utiliza menos recursos computacionais.
Piramuthu (2004)	Nesse estudo são avaliados métodos de seleção de características baseadas na distância probabilística no pré processamento de dados de entrada para a indução de DT. Os resultados desse estudo mostram que medidas de distância inter classe resultam em um melhor desempenho em relação as medidas probabilísticas em geral.
Chen e Cheng (2013)	Esse trabalho aplica dois modelos híbridos que resolvem problemas práticos de classificação de rating de crédito. Resultados experimentais demonstram que os modelos híbridos propostos para a classificação de rating de crédito superam os modelos anunciados no trabalho.
Hajek e Michalak (2013)	Os autores selecionam recursos utilizando a abordagem envoltória e comparam os métodos de seleção de recursos usados para a previsão do rating de crédito. Os resultados sugerem que a metodologia de rating dos EUA prefere o tamanho das empresas e das relações de valor de mercado, enquanto a metodologia europeia se baseia mais em índices de rentabilidade e alavancagem.
Piramuthu (2006)	O artigo utiliza ferramentas de apoio a decisão, para avaliar o risco de crédito a partir de uma perspectiva de aprendizado de máquina. O artigo conclui que é melhor escolher uma ferramenta que leva as características dos dados de avaliação de crédito financeiros em consideração.

continua na próxima página

continuação da Tabela 4

Lee (2007)	O autor sugere um novo modelo com melhor poder explicativo e estabilidade aplicando SVM para o problema de rating de crédito corporativo. Os resultados do experimento mostram que SVM supera os outros métodos.
Gonen, Gonen e Gargen (2012)	Os autores propõem a utilização do classificador proibit com uma estrutura definida previamente e aprendizado múltiplo de kernel com a construção de um kernel apropriado para realizar a seleção de recurso de grupo. Os resultados experimentais nos dois conjuntos de dados mostram a validade e eficácia do algoritmo de classificação binária proposto.
Tsai e Chen (2010)	Nesse trabalho quatro tipos diferentes de modelos híbridos são comparados por a "Classificação + Classificação", "Classificação + Clustering", "Clustering + Classificação" e as técnicas "Clustering + Clustering". Os resultados experimentais mostram que a "Classificação + Classificação" de modelo híbrido com base na combinação de regressão logística e redes neurais podem fornecer precisão, maior previsão e maximizar o lucro.
Zhou, Jiang e Shi (2010)	Esse artigo apresenta o uso do método de classificação do "subespaço mais perto" para avaliação de crédito. Experimentos mostram que o método de avaliação do risco de crédito subespaço mais próximo é um método competitivo.
Danenas, Garsva e Gudas (2011)	O artigo apresenta um estudo sobre o desenvolvimento do modelo de avaliação de risco de crédito usando classificadores baseados em SVM. A pesquisa mostrou que diferentes classificadores SVM produziram resultados semelhantes. Contudo, a seleção adequada do classificador e seus parâmetros continua a ser um problema importante.

continua na próxima página

continuação da Tabela 4

Ben-David e Frank (2009)	O artigo compara vários aspectos de modelos de aprendizado de máquina e sistemas especialistas. Os resultados mostram que, apesar de alguns modelos de aprendizado de máquina poderem superar a precisão do sistema especialista com significância estatística, a maioria dos modelos não supera.
Tsai, Hsu e Yen (2014)	Esse trabalho realiza um estudo de comparar conjuntos de classificadores por três técnicas de classificação amplamente utilizados, redes neurais perceptron multicamadas (MLP), SVM e DT. Os resultados experimentais mostram que conjuntos DT compostas por 80-100 classificadores utilizando o método de impulsão apresentam um melhor desempenho.
Ahn (2011)	Esse estudo propõe novos tipos de classificadores ANN e MSVM, que são nomeados OMANN e OMSVM. Eles são projetados para estender classificadores binários ANN ou SVM através da aplicação de estratégia ordinal particionamento pairwise (OPP). Os resultados experimentais mostraram que os modelos propostos melhoram a precisão da classificação em comparação com técnicas de classificação multiclasse típicas com a redução do esforço computacional.
Yang (2007)	Este trabalho apresenta um sistema de scoragem prático com base em um método de kernel. Os experimentos mostraram que os procedimentos adaptativos são comparáveis aos que usam aprendizagem de lote e que o método proposto é em robusto para os dados com um grande número de variáveis de entrada e distribuições de classe desequilibradas.
Saberi et al. (2013)	O estudo analisa pioneiramente o conceito de granularidade para selecionar o tamanho ideal de grupo de teste e de treinamento na área de aprendizado de máquina. Propõe um algoritmo integrado capaz de prever o status de crédito de clientes. O algoritmo apresentado prediz bem o status.

continua na próxima página

continuação da Tabela 4

Gestel et al. (2006)	O artigo apresenta um modelo genérico para desenvolver um sistema interno de rating avançado de análise de risco no contexto de países desenvolvidos e em desenvolvimento. O modelo sugerido é muito genérico no sentido de que ele pode ser facilmente aplicado em outros contextos de avaliação de risco.
Doumpos e Zopounidis (2006)	Esse estudo explora a combinação de diferentes métodos de classificação no desenvolvimento de modelos eficientes para a avaliação do risco de crédito. Os resultados mostram que os modelos combinados podem superar modelos individuais para análise de risco de crédito.
ONG, HUANG e TZENG (2005)	os autores usam programação genética para construir modelos de scoring de crédito. Comparam a taxa de erro para outros modelos de scoring de crédito, incluindo a ANN, DT, conjuntos irregulares e de regressão logística. A programação genética pode proporcionar uma melhor desempenho do que os outros modelos.
Tsai (2014)	O artigo desenvolve um novo modelo de socorro financeiro híbrido baseado na combinação da técnica de cluster e classificadores conjuntos. Combinando Self-Organizing Maps com MLP, conjuntos de classificadores tem um desempenho melhor, o que proporciona precisão predicação superior e erros tipo I e II inferiores.
Shi, Zhang e Qiu (2013)	Nesse trabalho um recurso de ponderação SVM é apresentado para avaliação de risco de crédito em que um F-score é adotado para o ranking importância de recurso. As duas versões ponderadas de SVM são testados contra o SVM tradicional e os resultados revelam a validade do método proposto.

continua na próxima página

continuação da Tabela 4

Tsai (2008a)	Este trabalho tem como objetivo a comparação de cinco métodos de seleção de característica utilizados na previsão de falências, teste t, matriz de correlação, regressão passo a passo, análise de componentes principais e análise fatorial. O método de seleção de recursos teste t supera os outros por duas medições de desempenho.
Tsai (2008b)	O objetivo desse artigo é examinar e compreender plenamente o desempenho de SVM. Usando MLP como o valor de referência, o classificador SVM apenas executa melhor em um dos quatro conjuntos de dados.
WU (2011)	O autor propõe uma solução de integração de estratégias de pré processamento de dados e o classificador de redes Bayesiano com o algoritmo de busca de árvore Naive Bayes, a fim de melhorar a precisão da classificação. Este artigo contribui para estender as aplicações práticas da incorporação de vários classificadores e estratégias de pré-processamento de dados.
Wu, Hu e Huang (2014)	O trabalho tem como objetivo discutir a influência do pré-processamento de dados antes da construção do classificador. Os resultados mostram que a <i>DT-Bagging</i> com o método de reamostragem de dados permite obter uma excelente precisão, o que indica que o modelo de previsão de duas fases proposto é melhor do que os modelos convencionais de uma só fase.
Ben-David (1995)	O artigo propoe uma métrica de seleção de atributo que leva em conta o erro e a monotonicidade na construção de DT. A métrica empiricamente mostrada é capaz de reduzir significativamente o grau de não-monotonicidade de DT sem sacrificar a sua exatidão indutiva.

continua na próxima página

continuação da Tabela 4

Luo, Cheng e Hsieh (2009)	Os autores comparam a taxa de precisão de dois modelos de classificação, incluindo o SVM, e o <i>Clustering-Launched Classification</i> . Os resultados mostram que, na maioria dos casos, a precisão da CLC supera os outros. Além disso, a CLC é fácil de aplicar e não precisa de muito conhecimento estatístico.
Wang et al. (2012)	O estudo propõe dois conjuntos de DT dupla: <i>RS-bagging DT</i> e <i>bagging-RS DT</i> , que são baseados em duas estratégias de conjunto: bagging e subespaço aleatório, para reduzir a influência dos dados sobre o ruído e os atributos redundantes de dados e para obter a precisão da classificação relativamente maior. Os resultados mostram que as duas técnicas podem ser usados como alternativas de classificação de crédito.
Huang, Tzeng e Ong (2006)	Esse trabalho propõe-se a lidar com o problema de classificação de crédito utilizando a programação genética de dois estágios (2SGP), incorporando as vantagens das regras <i>if-then</i> e a função discriminante. Com base dos resultados numéricos, pode-se concluir que 2SGP pode proporcionar a melhor precisão do que outros modelos.
Won, Kim e Bae (2012)	O artigo propõe um modelo de refinamento do conhecimento que pode redefinir as regras múltiplas de algoritmos baseados em regras de conjuntos de dados de dividendos usando GA. A nova técnica é chamada de <i>GAKR (genetic algorithm knowledge refinement)</i> . Os experimentos mostram que o modelo GAKR sempre supera outros modelos.
TSAI e WU (2008)	Esse artigo investiga o desempenho de um único classificador como o classificador de linha de base para comparação com múltiplos classificadores e múltiplos classificadores diversificados utilizando ANN. Os múltiplos classificadores diversificados se desempenham pior em todos os conjuntos de dados. No entanto, para o tipo de erros do tipo II e I, não há vencedor exato.

continua na próxima página

continuação da Tabela 4

Zhang et al. (2010)	Os autores propõe um novo modelo de <i>credit scoring</i> , chamado vertical bagging decision tree model (VBDTM). Os resultados da análise mostram que o desempenho do método proposto é excelente na precisão da previsão.
Liu, Fu e Lin (2010)	Esse artigo apresenta um novo modelo de <i>credit scoring</i> quantitativo baseado em SVM com algoritmo genético adaptativo, gr-GA-SVM. Os resultados numéricos indicam que o modelo apresentado é mais preciso e eficiente que SVM e SVM-GA.
Siarni, Gholamian e Basiri (2013)	Nesse trabalho, uma aplicação do algoritmo modelo de árvore localmente linear (LOLIMOT) foi testado para avaliar a superioridade de seu desempenho para prever o estado de crédito do cliente. Os resultados analíticos indicam que o LOLIMOT aumenta significativamente a precisão da previsão.
Angelini, Tollo e Roli (2008)	Nesse trabalho foram desenvolvidos dois sistemas de ANN, uma rede com um padrão de alimentação de entrada e outro com uma arquitetura para fins especiais. As redes neurais podem ser muito bem sucedidas na aprendizagem e na estimativa de default, desde que a análise, os dados de pré-processamento e de formação sejam realizadas de forma cuidadosa.
Antonakis e Sfakianakis (2009)	O estudo utiliza bases reais de crédito para alusão NBR contra análise linear discriminante, análise de regressão logística, k-nearest neighbours, árvores de classificação e redes neurais. NBR tem uma capacidade de previsão menor do que alguns dos outros cinco métodos sob todas as medidas utilizadas.
Liu e Schumann (2005)	O estudo ilustra como quatro métodos de seleção de atributos- 'ReliefF', 'baseado correlação', 'baseado na consistência' e 'algoritmos Wrapper' ajudam a melhorar a simplicidade, rapidez e precisão dos modelos de <i>scoring</i> . Entre os quatro métodos, o baseado em consistência e os métodos de seleção de recurso Wrapper possuem um desempenho relativamente melhor.

continua na próxima página

continuação da Tabela 4

Tsakonas e Dounias (2007)	O objetivo do trabalho é demonstrar como gerar ANN lógicas eficientes, com o auxílio de métodos de programação genética de forma adaptativa por meio de um esquema inovador. A metodologia proposta supera uma série de metodologias estatísticas competitivas, ao mesmo tempo que produz regras de decisão acessíveis, curtas e transparentes em significado e uso.
Twala (2013)	Esse artigo investiga a robustez de cinco algoritmos de aprendizado de máquina para ambiente de risco de crédito com ruído. Na presença de ruído, é o ruído na variável de classe que é responsável pela má precisão da previsão do conceito de aprendizagem.
Oreski e Oreski (2014)	Nesse trabalho, um algoritmo de heurística avançada é apresentado, o algoritmo genético com redes neurais (HGA-NN), é usado para identificar um subconjunto de recurso ideal e para aumentar a precisão da classificação e escalabilidade na avaliação de risco de crédito. Os resultados experimentais são promissores para a seleção de recursos e classificação.
Hashemi et al. (1998)	Nesse artigo um modelo composto de componentes conjuntos ANN foi construído. O modelo final foi capaz de classificar corretamente 96 % de um conjunto de testes de quatro tipos de estruturas do banco.
Garcia, Marques e Sanchez (2012)	O objetivo do trabalho é investigar sistematicamente se a aplicação de algoritmos de filtragem leva a um aumento na precisão de classificadores. Os conjuntos processados possuem desempenho significativamente melhor do que os conjuntos de treinamento não-pré-processados ao usar a regra de decisão de <i>K nearest neighbour</i> .
Marques, Garcia J.a e Sanchez (2012)	O artigo combina algoritmos de reamostragem de dados (<i>bagging e AdaBost</i>) e atribui métodos de seleção de subconjuntos. O novo classificador conjunto de dois níveis constitui uma solução adequada para os problemas de classificação de crédito, possui desempenho melhor do que os conjuntos individuais tradicionais do que os classificadores individuais.

continua na próxima página

continuação da Tabela 4

Yu, Wang e Lai (2009)	Nesse estudo, um novo modelo fuzzy de grupo tomador de decisão é proposto como uma eficaz ferramenta de análise de decisão multicritério. Os resultados revelam que o modelo pode fornecer uma solução promissora para tarefas de classificação de crédito, a técnica proposta tem um grande potencial para ser aplicado a outros problemas financeiros.
Marques, Garcia e Sanchez (2012)	O artigo avalia o desempenho de sete técnicas de previsão individuais quando usado como parte de cinco métodos. Os resultados experimentais e testes estatísticos mostram que DT constitui a melhor solução para métodos de conjuntos, seguido de perto pelo ANN perceptron multicamadas e da regressão logística.
Sanchez, Monelos e Lopez (2013)	Esse artigo analisa a capacidade preditiva de dois modelos de previsão de falha financeira corporativa (regressão logística e ANN), derivados de processos de auditoria. Os modelos indicam que o fluxo de informações de auditoria contém elementos suficientes para antecipar o processo de falência.
Varetto (1998)	O estudo compara análise discriminante linear (LDA) com Algoritmo Genético (GA). GA tem-se revelado um instrumento muito eficaz de análise de risco de insolvência. Em comparação com a análise discriminante, GAs tornam possível obter funções lineares otimizadas que não dependem de hipóteses estatísticas restritiva.
Wang et al. (2011)	Esse estudo realiza uma avaliação comparativa do desempenho de três métodos de conjunto, Bagging, Boosting, e Stacking, com base em quatro bases de aprendizado, regressão logística, DT, ANN e SVM. Os três métodos de conjunto podem melhorar substancialmente o aprendizado das base. Bagging tem um melhor desempenho em todos os conjuntos de dados.

continua na próxima página

continuação da Tabela 4

Bellotti e Crook (2009)	Os autores comparam SVM com os métodos tradicionais regressão logística e análise de discriminantes. SVMs são competitivos e podem ser usados como base de um método de seleção de recurso para descobrir as características que são mais importante na determinação do risco de inadimplência.
Chen e Li (2010)	Esse estudo propõe quatro abordagens com o classificador SVM para seleção de características que retém informações suficientes para classificação. O resultado deste estudo sugere que a abordagem de classificação de crédito híbrido é mais robusto e eficaz na busca subconjuntos ótimos e é um método promissor para as áreas de mineração de dados.
Chen, Ma e Ma (2009)	Nesse trabalho é proposto SVM híbrido baseado em três estratégias: (1) usando CART, (2) usando MARS e (3) utilizando pesquisa de rede para otimizar os parâmetros. Resultados analíticos demonstram que a técnica SVM híbrido não só tem a melhor taxa de classificação, mas também tem o menor erro de tipo II, em comparação com CART, MARS e SVM.
Finlay (2009)	O artigo utiliza GA para determinar o impacto sobre o desempenho resultante de ter objetivos diferentes para o modelo de construção e avaliação. Os resultados empíricos mostram que todos os modelos funcionam de forma semelhante bem, o que sugere que os objetivos de modelagem e de negócios estão bem alinhados.
Hoffmann et al. (2007)	Os autores investigam o poder e utilidade das regras de classificação fuzzy para fins de mineração de dados. Propõe duas regras fuzzy evolutivas: uma estratégia de evolução e um algoritmo genético. Os classificadores fuzzy genéticos se equiparam favoravelmente com os outros classificadores em termos de exatidão de classificação.

continua na próxima página

continuação da Tabela 4

Hsieh (2005)	O autor apresenta uma abordagem híbrida de mineração na concepção de um modelo de <i>credit scoring</i> eficaz com base em técnicas de agrupamento e de ANN. No método proposto a abordagem de mineração híbrida pode ser usada para construir modelos eficazes de avaliação de crédito.
Huang, Chen e Wang (2007)	O artigo demonstra a precisão do classificador SVM utilizando três estratégias para construir modelos de <i>credit scoring</i> baseados em SVM híbrido. Em comparação com ANN, GA e DT, o classificador SVM conseguiu uma precisão classificatória idêntica com relativamente poucos recursos de entrada. O SVM é um promissor complemento aos métodos de mineração de dados existentes.
Lee et al. (2002)	O objetivo do estudo é explorar o desempenho de classificação de crédito, integrando BNN com a abordagem tradicional de análise discriminante. A abordagem híbrida proposta converge muito mais rápido do que o modelo ANN convencional. O aumento de precisão de classificação do BNN supera a análise discriminante tradicional e métodos de regressão logística.
Lin (2009)	O artigo propõe uma nova abordagem de regressão logística-ANN, para explorar se o modelo híbrido de dois estágios supera os tradicionais e para a construção de um sistema de alerta de socorro financeiro para a indústria bancária em Taiwan. Os resultados mostram que fatores como taxa de recuperação de ativos sensíveis e juros são significativamente relacionados com a dificuldade financeira de bancos em Taiwan.
Martens et al. (2007)	Os autores fornecem uma visão geral das técnicas de extração de regras recentemente propostas para SVMs e introduzir outras duas retiradas de redes neurais artificiais, Trepan e G-REX. As técnicas de extração regra SVM perde somente uma pequena percentagem de desempenho em comparação com o SVM.

continua na próxima página

continuação da Tabela 4

Seow e Thomas (2006)	O objetivo deste trabalho é desenvolver um modelo de programação dinâmica adaptativa onde os métodos de atualização de Bayes são empregados para melhor estimar a take-up distribuição de probabilidade. No futuro, as empresas terão de ser mais claras, não só na maneira de explorar os dados em suas bases de dados, mas também na maneira de recolher os seus dados.
Setiono, Baesens e Mues (2009)	O artigo abordar uma questão importante utilizando ANN que foi deixado de fora em um artigo recente “ <i>Knowledge discovery using a neural network simultaneous optimization algorithm on a real world classification problem</i> ” por Sexton et al. O algoritmo produz conjuntos de regras precisas que são concisas e compreensíveis.
Tang e Chi (2005)	O artigo estende trabalhos anteriores de ANN aplicando análise <i>receiver operating characteristic</i> para comparar o desempenho do modelo de Logit a de Lógica Fuzzy (FL). FL é superior em exatidão global e na classificação de empresas padrão, Logit é preferível em situações em que uma maior precisão na classificação de empresas não-padrão é o preferido.
West (2000)	Esse artigo investiga a precisão de classificação de crédito de cinco modelos ANN: <i>multilayer perceptron</i> , <i>mixture-of-experts</i> , <i>radial basis function</i> , <i>learning vector quantization</i> e <i>fuzzy adaptive resonance</i> . Os resultados demonstram que o perceptron multilayer pode não ser o modelo ANN mais preciso e que tanto os especialistas <i>mixture-of-experts</i> e <i>radial basis function</i> devem ser considerados.
Xu, Zhou e Wang (2009)	O artigo propõe três algoritmos de análise de ligação com base no pré-processamento de SVM. Mostra-se que os métodos de análise de classificação de conexão genética têm maior desempenho em termos de precisão da classificação.
