

Magnamara Acácio Pimentel Barnabé

**PREDIÇÃO DE RISCO JURÍDICO DE INSTITUIÇÕES  
FINANCEIRAS**

Brasil

2022, abril



Magnamara Acácio Pimentel Barnabé

## **PREDIÇÃO DE RISCO JURÍDICO DE INSTITUIÇÕES FINANCEIRAS**

Dissertação apresentada ao Curso de Mestrado Acadêmico em Economia, Universidade de Brasília, como requisito parcial para a obtenção do título de Mestre em Economia

Universidade de Brasília - UnB  
Faculdade de Administração Contabilidade e Economia - FACE  
Departamento de Economia - ECO  
Programa de Pós-Graduação

Orientador: Daniel Oliveira Cajueiro

Brasil  
2022, abril

Magnamara Acácio Pimentel Barnabé  
PREDIÇÃO DE RISCO JURÍDICO DE INSTITUIÇÕES FINANCEIRAS/ Magnamara Acácio  
Pimentel Barnabé. – Brasil, 2022, abril-  
117p. : il. (algumas color.) ; 30 cm.

Orientador: Daniel Oliveira Cajueiro

Dissertação (Mestrado) – Universidade de Brasília - UnB  
Faculdade de Administração Contabilidade e Economia - FACE  
Departamento de Economia - ECO  
Programa de Pós-Graduação, 2022, abril.

1. Palavra-chave1. 2. Palavra-chave2. 3. Palavra-chave3. II. Universidade de Brasília. III. Faculdade  
de Administração, Contabilidade e Economia - FACE. IV. Departamento de Economia IV. PREDIÇÃO  
DE RISCO JURÍDICO DE INSTITUIÇÕES FINANCEIRAS

Magnamara Acácio Pimentel Barnabé

## **PREDIÇÃO DE RISCO JURÍDICO DE INSTITUIÇÕES FINANCEIRAS**

Dissertação apresentada ao Curso de Mestrado Acadêmico em Economia, Universidade de Brasília, como requisito parcial para a obtenção do título de Mestre em Economia

Trabalho aprovado. Brasil, 28 de abril de 2022:

---

**Daniel Oliveira Cajueiro**  
Orientador

---

**Hebert Kimura**  
Convidado 1

---

**Marina Rossi**  
Convidado 2

Brasil  
2022, abril



*Este trabalho é dedicado às mulheres, mães, esposas, profissionais, amigas,  
aquelas que lutam para que a liberdade nos permita ser e estar onde quisermos.*





# AGRADECIMENTOS

Agradeço aos professores que estiveram presentes na minha formação pessoal e profissional neste mestrado, em especial ao meu orientador Daniel Cajueiro, aos colegas de curso por compartilharem experiências e conhecimentos nas aulas e trabalhos, aos meus filhos que sempre me motivam a melhorar e aos meus pais, por me propiciarem o suporte emocional e logístico para participação nas aulas e na realização deste trabalho.



*“Quem ensina aprende ao ensinar e  
quem aprende ensina ao aprender  
(Paulo Freire)*



# RESUMO

Nesse trabalho implementamos diferentes modelos de previsão de ocorrência de desembolso financeiro em processos judiciais de instituições financeiras brasileiras a partir de técnicas estatísticas, métodos computacionais e algoritmos de *Machine Learning* (*Logistic Regression, Decision Tree, Bagging, Ada Boost, Gradient Boosting e Random Forest*) e comparamos os modelos preditivos para identificar o melhor. Para avaliar a adequação dos modelos, realizamos os testes com a obtenção dos índices de acurácia, sensibilidade, especificidade, precisão, f1-score e Area Under ROC Curve (AUROC), com o intuito de validar os modelos em 2 (dois) cenários, um com o uso exclusivo de variáveis de cadastro dos processos judiciais e outro com variáveis de andamento processual. Os melhores resultados foram obtidos com *Random Forest* para o cenário de variáveis exclusivas de cadastro (acurácia de 80.71%) e o *Gradient Boosting* para o cenário que agregou variáveis de andamento processual (acurácia de 88.94%). O uso de modelagem preditiva para aferir o risco jurídico, um tipo de risco operacional, mostrou-se adequado e com elevado grau de assertividade para atender as exigências regulamentares de provisão e registro em passivos contingentes nas demonstrações financeiras, conforme preconizado no *IAS 37 - Provisions, contingent liabilities and contingent assets*, Pronunciamento Técnico CPC 25 e Resolução nº 3.823 do Banco Central do Brasil.

**Palavras-chave:** *Machine Learning*, gerenciamento de riscos, risco jurídico, risco operacional, provisão, provisão judicial, passivo contingente, instituições financeiras, modelos preditivos, demonstrações financeiras, CPC 25, IAS 37, IASB, Resolução BACEN 3.823, Judiciário, Poder Judiciário, Jurídico, ISO 31000, COSO ERM, Python.



# ABSTRACT

In this work we implemented different models for predicting the occurrence of financial disbursement in legal proceedings of Brazilian financial institutions from statistical techniques, methods computational tools and Machine Learning algorithms (Logistic Regression, Decision Tree, Bagging, Ada Boost, Gradient Boosting and Random Forest) and compared the predictive models to identify the best. In order to evaluate the adequacy of the models, we carried out the tests to obtain the indexes of accuracy, sensitivity, specificity, precision, f1-score and Area Under ROC Curve (AUROC), with the purpose of validating the models in 2 (two) scenarios, one with the exclusive use of register variables of the lawsuits and another with procedural progress variables. The best results were obtained with Random Forest for the scenario of exclusive variables of registration (accuracy of 80.71%) and Gradient Boosting for the scenario that added procedural progress variables (accuracy of 88.94%). The use of predictive modeling to assess legal risk, a type of operational risk, proved to be adequate and with high degree of assertiveness to meet regulatory requirements for provision and recording of liabilities contingents in the financial statements, as recommended in IAS 37 - Provisions, contingent liabilities and contingent assets, Pronunciamento Técnico CPC 25 and Resolução No. 3.823 of the Central Bank from Brazil.

**Keywords:** Machine Learning, risk management, legal risk, operational risk, provision, judicial provision, contingent liability, financial institutions, predictive models, financial statements, CPC 25, IAS 37, IASB, BACEN Resolução 3.823, Judiciary, Judiciary, Legal, ISO 31000, COSO ERM, Python.





# LISTA DE ILUSTRAÇÕES

Figura 1 – Classificação de riscos financeiros e não financeiros . . . . .	28
Figura 2 – ISO 31000 - <i>Risk management principles and guidelines</i> . . . . .	31
Figura 3 – Cubo do COSO ERM 2004 . . . . .	33
Figura 4 – Nova imagem do COSO ERM 2017 . . . . .	34
Figura 5 – CPC 25 - Fatos geradores . . . . .	36
Figura 6 – CPC 25 - Provisão e passivo contingente . . . . .	37
Figura 7 – Seleção da amostra . . . . .	47
Figura 8 – Categorias de aprendizado e tarefas associadas . . . . .	55
Figura 9 – Exemplo de árvore de decisão . . . . .	57
Figura 10 – Exemplo Bagging . . . . .	58
Figura 11 – Fórmula AdaBoost 1 . . . . .	59
Figura 12 – Fórmula AdaBoost 2 . . . . .	60
Figura 13 – Exemplo Random Forest . . . . .	61
Figura 14 – Matriz de confusão . . . . .	61
Figura 15 – Interpretação dos resultados da curva ROC . . . . .	63
Figura 16 – ROC - Receiver Operating Characteristic . . . . .	63
Figura 17 – AUC - Area Under the Curve . . . . .	64
Figura 18 – Matriz de confusão do cenário 1 - Base teste . . . . .	65
Figura 19 – Curva ROC individual do cenário 1 - Base teste . . . . .	66
Figura 20 – Curva ROC consolidada do cenário 1 - Base teste . . . . .	67
Figura 21 – Matriz de confusão do cenário 2 - Base teste . . . . .	67
Figura 22 – Curva ROC individual do cenário 2 - Base teste . . . . .	68
Figura 23 – Curva ROC consolidada - cenário 2 - Base teste . . . . .	69



# LISTA DE TABELAS

Tabela 1 – Conceitos de gerenciamento de riscos . . . . .	30
Tabela 2 – Eventos da IAS 37 . . . . .	35
Tabela 3 – Balanço Patrimonial do Banco do Brasil 2021 . . . . .	39
Tabela 4 – Balanço Patrimonial da CAIXA 2021 . . . . .	39
Tabela 5 – Balanço Patrimonial do Itaú 2021 . . . . .	40
Tabela 6 – Balanço Patrimonial do Bradesco 2021 . . . . .	40
Tabela 7 – Balanço Patrimonial do Santander 2021 . . . . .	41
Tabela 8 – Provisões IF brasileiras para risco jurídico . . . . .	41
Tabela 9 – Passivo contingente IF brasileiras para risco jurídico . . . . .	41
Tabela 10 – Provisões IF brasileiras para risco jurídico em dólar . . . . .	42
Tabela 11 – Passivo contingente IF brasileiras para risco jurídico em dólar . . . . .	42
Tabela 12 – Balanço Patrimonial JP Morgan Chase 2021 . . . . .	42
Tabela 13 – Balanço Patrimonial Bank of America 2021 . . . . .	43
Tabela 14 – Balanço Patrimonial <i>MUFJ Bank</i> 2021 . . . . .	45
Tabela 15 – Balanço Patrimonial ICBC 2021 . . . . .	45
Tabela 16 – Balanço Patrimonial BNP Paribas 2021 . . . . .	46
Tabela 17 – Variáveis selecionadas . . . . .	49
Tabela 18 – Variável dependente . . . . .	49
Tabela 19 – Variáveis independentes de cadastro . . . . .	50
Tabela 20 – Variáveis independentes de andamento processual . . . . .	50
Tabela 21 – Cenários para os algoritmos . . . . .	51
Tabela 22 – Biblioteca de modelos do <i>Python</i> . . . . .	56
Tabela 23 – Bibliotecas acessórias do <i>Python</i> . . . . .	56
Tabela 24 – Resultados das validações do cenário 1 - Base teste . . . . .	66
Tabela 25 – Resultados das validações do cenário 2 - Base teste . . . . .	68



# LISTA DE ABREVIATURAS E SIGLAS

AUC	Area Under the Curve
AB	Ada Boost
BCB	Banco Central do Brasil
BG	Bagging Classifier
BP	Balço Patrimonial
CPC	Comitê de Pronunciamentos Contábeis
DF	Demonstrações Financeiras
DT	Decision Tree
GB	Gradient Boosting
IF	Instituição Financeira
ISO	International Organization for Standardization
LR	Logistic Regression
ML	Machine learning
PF	Pessoa Física
PJ	Pessoa Jurídica
RF	Random Forest
RO	Risco Operacional
ROC	Receiver Operating Characteristics



# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>25</b>
<b>2</b>	<b>PROCESSO E RISCO JURÍDICO EM INSTITUIÇÕES FINANCEIRAS</b>	<b>27</b>
<b>2.1</b>	<b>Conceitos de risco</b>	<b>27</b>
2.1.1	Classificação de riscos	27
2.1.1.1	Risco operacional	29
2.1.1.2	Risco jurídico	29
2.1.2	Gerenciamento de riscos	30
2.1.2.1	ISO 31000	30
2.1.2.2	COSO ERM	32
2.1.3	Instituições financeiras brasileiras e o risco jurídico	34
<b>2.2</b>	<b>Provisão e passivo contingente</b>	<b>35</b>
2.2.1	IAS 37 - Provisões, passivos contingentes e ativos contingentes	35
2.2.1.1	Pronunciamento técnico CPC 25	36
2.2.1.2	Resolução BCB 3.823/2009	37
<b>2.3</b>	<b>Benchmark de risco jurídico e provisão em instituições financeiras</b>	<b>37</b>
2.3.1	Instituições financeiras brasileiras	37
2.3.1.1	Balancos patrimoniais	38
2.3.1.2	Provisões e passivos contingentes das IF brasileiras	41
2.3.2	Instituições financeiras internacionais	42
2.3.2.1	JP Morgan Chase	42
2.3.2.2	Bank of America	43
2.3.2.3	MUFJ Bank – Mitsybish UFK Financial Group	44
2.3.2.4	ICBC - Industrial and Commercial Bank of China	45
2.3.2.5	BNP Paribas	46
<b>3</b>	<b>BASE DE DADOS</b>	<b>47</b>
<b>3.1</b>	<b>Seleção da amostra</b>	<b>47</b>
<b>3.2</b>	<b>Tratamento das variáveis</b>	<b>48</b>
<b>3.3</b>	<b>Seleção das variáveis</b>	<b>48</b>
3.3.1	Variável dependente	49
3.3.2	Variáveis independentes de cadastro	49
3.3.3	Variáveis independentes de andamento processual	50
<b>3.4</b>	<b>Cenários do estudo</b>	<b>50</b>
<b>4</b>	<b>METODOLOGIA DE PESQUISA</b>	<b>53</b>
<b>4.1</b>	<b>Ciência de dados e Machine Learning</b>	<b>53</b>
4.1.1	ML e instituições financeiras	53
4.1.2	ML e o judiciário	54
4.1.3	Tarefas de aprendizado	54
4.1.4	Python	55
4.1.4.1	Bibliotecas principais	55
4.1.4.2	Bibliotecas acessórias	56

<b>4.2</b>	<b>Técnicas Estatísticas</b>	<b>56</b>
4.2.1	Classificação	56
4.2.2	Regressão Logística	57
4.2.3	Métodos baseados em árvores	57
4.2.3.1	<i>Decision Tree</i>	57
4.2.3.2	<i>Bagging</i>	58
4.2.3.3	<i>Ada Boost</i>	59
4.2.3.4	<i>Gradient Boosting</i>	60
4.2.3.5	<i>Random Forest</i>	60
<b>4.3</b>	<b>Avaliação e validação de modelos</b>	<b>61</b>
4.3.1	Matriz de confusão	61
4.3.1.1	Acurácia	61
4.3.1.2	Sensibilidade / Recall	62
4.3.1.3	Especificidade	62
4.3.1.4	Precisão	62
4.3.1.5	f1 Score	62
4.3.2	<i>ROC - Receiver Operating Characteristic</i>	62
4.3.2.1	AUC - Area Under the Curve	63
<b>5</b>	<b>RESULTADOS E ANÁLISES</b>	<b>65</b>
<b>5.1</b>	<b>Cenário 1 - Resultados das validações - Base Teste</b>	<b>65</b>
5.1.1	Cenário 1 - Matriz de confusão	65
5.1.2	Cenário 1 - Acurácia, Sensibilidade, Especificidade, Precisão, f1-score e AUC	66
5.1.3	Cenário 1 - Curva ROC	66
<b>5.2</b>	<b>Cenário 2 - Resultados das validações</b>	<b>67</b>
5.2.1	Cenário 2 - Matriz de Confusão	67
5.2.2	Cenário 2 - Acurácia, Sensibilidade, Especificidade, Precisão, f1-score e AUC	68
5.2.3	Cenário 2 - Curva ROC	68
<b>6</b>	<b>CONSIDERAÇÕES FINAIS</b>	<b>71</b>
	<b>REFERÊNCIAS</b>	<b>73</b>
	<b>APÊNDICES</b>	<b>77</b>
	<b>APÊNDICE A – LEIAUTE DOS DADOS DAS VARIÁVEIS INDEPENDENTES DE CADASTRO</b>	<b>79</b>
	<b>APÊNDICE B – CÓDIGO FONTE DOS MODELOS NO PYTHON</b>	<b>81</b>
	<b>ANEXOS</b>	<b>97</b>
	<b>ANEXO A – DF 2021 - BANCO DO BRASIL</b>	<b>99</b>
	<b>ANEXO B – DF 2021 - CAIXA</b>	<b>103</b>
	<b>ANEXO C – DF 2021 - ITAÚ</b>	<b>109</b>



<b>ANEXO D – DF 2021 - BRADESCO . . . . .</b>	<b>113</b>
<b>ANEXO E – DF 2021 - SANTANDER . . . . .</b>	<b>115</b>



# 1 INTRODUÇÃO

Esta dissertação é motivada pela relevância do risco jurídico nos montantes de provisão e passivos contingentes nos balanços patrimoniais das instituições financeiras brasileiras e pelos desafios de mensuração da probabilidade e impacto para estimar esses montantes conforme regulamentações e melhores práticas nacionais e internacionais. Alguns referenciais foram importantes nessa motivação como Basiléia II (SUPERVISION, 2004), *IAS 37 - Provisions, Contingent Liabilities and Contingent Assets* (IASB, 2001), Pronunciamento Técnico CPC 25 - Provisões, Passivos Contingentes e Ativos Contingentes (CONTÁBEIS, 2009) e Resolução BCB nº 3.823/2009 (BRASIL, 2009).

Outro incentivo à pesquisa é o fato do risco jurídico ou legal estar inserido na vertente de risco operacional (SUPERVISION, 2004), uma tipologia de risco desafiadora para uso em *Machine Learning* em razão da dificuldade de identificação, seleção e composição de base de dados para estudo e com bibliografia mais escassa em comparação ao risco de crédito.

Utilizamos no trabalho uma base de dados histórica de perdas judiciais na perspectiva financeira de um grande Banco brasileiro, técnicas estatísticas e *Machine Learning* que permitiram o desenvolvimento de modelos preditivos com alto grau de acurácia e desempenho. A utilização dos dados descaracterizou autores e individualizações, com resguardo dos princípios éticos e de proteção de dados.

A abordagem do tema proposto teve o objetivo de revisar a literatura, incluindo as regulamentações, para responder as seguintes perguntas:

1. A utilização de algoritmos preditivos de *Machine Learning* para apurar risco jurídico e risco operacional de instituições financeiras é viável?
2. Em caso de resposta afirmativa da pergunta anterior, qual o melhor algoritmo preditivo de *Machine Learning* para apurar o risco jurídico?

O estudo busca responder a essas duas perguntas com revisão da literatura em ordem lógica e concatenada em 4 (quatro) capítulos:

No capítulo 2 (dois) - **Processo e risco jurídico em instituições financeiras**, abordamos conceitos de riscos e gerenciamento de riscos, regulamentações correlatas, benchmarks de provisão e passivo contingente de risco jurídico em IF brasileiras e maiores IF internacionais.

No capítulo 3 (três) **Base de dados** descrevemos a forma de seleção da amostra, variáveis e definição de 2 (dois) cenários de estudo para aplicação de modelos de *Machine Learning*, um apenas com variáveis de cadastro de processos judiciais e outro com incremento de variáveis de andamento processual.

No capítulo 4 (quatro) - **Metodologia de pesquisa** clarificamos a compreensão sobre a escolha e uso de modelos preditivos de Regressão Logística, Árvore de Decisão, *Bagging*, *Ada Boost*, *Gradient Boosting* e *Random Forest* para apuração de risco jurídico. Aprofundamos o conhecimento em técnicas estatísticas, algoritmos e ferramentas de *Machine Learning* para o desenvolvimento de modelos. Quanto avaliação e validação de modelos, checamos a bibliografia relacionada a matriz de confusão, acurácia, sensibilidade, precisão, f1-score, curva ROC e AUC.

No capítulo 5 (cinco) **Resultados e análises** apresentamos as comparações das aplicações no *Python* dos modelos preditivos selecionados para os 2 (dois) cenários de estudo, de forma a identificar o melhor estimador da probabilidade de perda financeira decorrente de risco jurídico em instituições financeiras.



## 2 PROCESSO E RISCO JURÍDICO EM INSTITUIÇÕES FINANCEIRAS

### 2.1 CONCEITOS DE RISCO

Há diversas tentativas para desenvolver uma linguagem padronizada de riscos e o Guia ISO 73 - *Risk Management Vocabulary* foi desenvolvido como a terminologia comum que deve ser utilizada em todos os padrões ISO, conceitos que também serão utilizados neste trabalho (STANDARDIZATION, 2009b).

A ISO 31000 - *Risk Management Principles and Guidelines* assim define risco: “efeito da incerteza nos objetivos. Um efeito é um desvio do esperado. Pode ser positivo, negativo ou ambos, e pode abordar, criar ou resultar em oportunidades e ameaças. Os objetivos podem ter diferentes aspectos e categorias e podem ser aplicados em diferentes níveis. O risco é geralmente expresso em termos de fontes de risco, eventos potenciais, suas consequências e sua probabilidade” (STANDARDIZATION, 2009a). Para bom entendimento do conceito de riscos, destacamos outros conceitos relevantes:

**Evento:** pode ter uma ou mais ocorrências e pode ter várias causas e várias consequências. Pode ser algo esperado que não acontece ou algo que não é esperado que acontece.

**Consequência:** resultado de um evento que afeta os objetivos. Pode ser certa ou incerta e pode ter efeitos diretos ou indiretos, positivos ou negativos sobre os objetivos. As consequências podem ser expressas qualitativa ou quantitativamente.

**Probabilidade:** chance de algo acontecer. Na avaliação de riscos é utilizada para se referir a chance de algo acontecer somada a uma medida objetiva ou subjetiva, qualitativa ou quantitativa, e descrita em termos gerais ou matemáticos.

**Controle:** medida que mantém ou modifica o risco. Os controles incluem, mas não se limitam a qualquer processo, política, dispositivo, prática ou outras condições e ações que mantenham ou modifiquem o risco. Os controles nem sempre podem exercer o efeito modificador pretendido ou presumido.

Outro *framework* conhecido no tema de riscos corporativos é o COSO - Gerenciamento de riscos corporativos - Estrutura integrada.

Em sua introdução, o COSO de 2004 registra que os eventos podem gerar impacto tanto negativo quanto positivo ou ambos. Os que geram impacto negativo representam riscos que podem impedir a criação de valor ou mesmo destruir o valor existente. Os de impacto positivo podem contrabalançar os de impacto negativo ou podem representar oportunidades, que por sua vez representam a possibilidade de um evento ocorrer e influenciar favoravelmente a realização dos objetivos, apoiando a criação ou a preservação de valor (COSO, 2004).

#### 2.1.1 CLASSIFICAÇÃO DE RISCOS

Conforme Moosa registra (MOOSA, 2007), os riscos podem ser classificados de várias formas, como por exemplo, é aceitável organizá-los dependendo de quão quantificáveis são. Em um extremo estão os tipos de riscos decorrentes de mudanças nos valores de ativos líquidos. Nesse caso, os dados sobre o histórico anterior estão disponíveis, o que torna o risco, independentemente da definição, totalmente quantificável. No outro extremo estão os riscos decorrentes de eventos pouco frequentes, como uma crise financeira contagiosa, com consequências potencialmente massivas, em que o risco é bastante difícil de

quantificar.

O esquema de classificação mais simples e comum é baseado na distinção entre risco financeiro e não financeiro. O risco financeiro é decorrente de perdas inerentes às operações financeiras, por isso é classificado em risco de crédito, possibilidade de inadimplência por parte dos tomadores, e risco de mercado, decorrente de flutuações de preços financeiros. A Figura 1 demonstra essa separação (MOOSA, 2007):

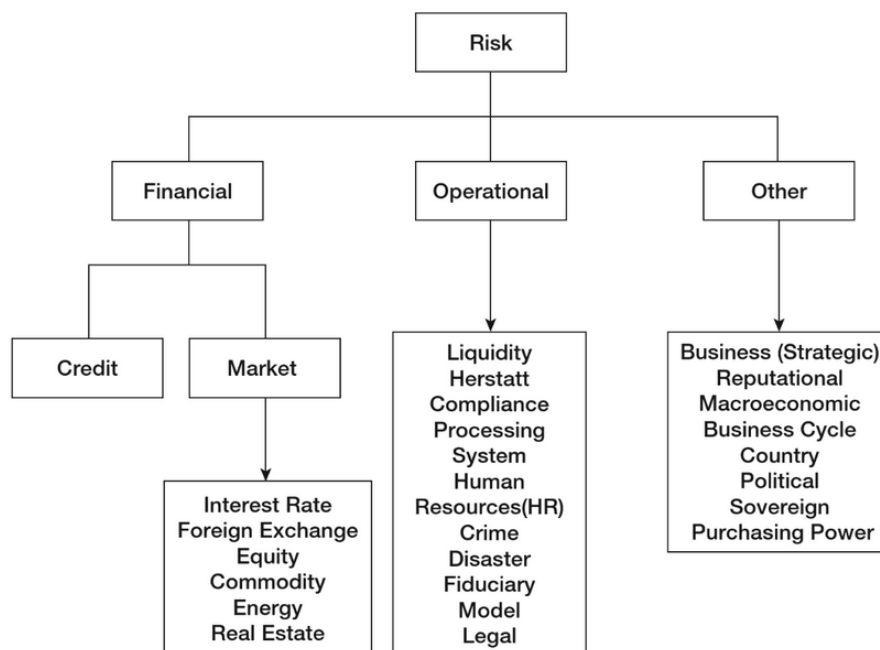


Figura 1 – Classificação de riscos financeiros e não financeiros

A Resolução nº 4.557/2017 do Banco Central do Brasil conceitua os tipos de riscos, os quais sintetizamos a seguir (BRASIL, 2017b):

**Risco de crédito:** possibilidade de ocorrência de perdas associadas ao descumprimento pela contraparte de suas obrigações nos termos pactuados; desvalorização, redução de remunerações e ganhos esperados em instrumento financeiro decorrentes da deterioração da qualidade creditícia da contraparte, do interveniente ou do instrumento mitigador; reestruturação de instrumentos financeiros; ou custos de recuperação de exposições caracterizadas como ativos problemático.

**Risco operacional:** possibilidade da ocorrência de perdas resultantes de eventos externos ou de falha, deficiência ou inadequação de processos internos, pessoas ou sistemas.

**Risco de mercado:** possibilidade de ocorrência de perdas resultantes da flutuação nos valores de mercado de instrumentos detidos pela instituição.

**Risco de taxa de juros:** risco, atual ou prospectivo, do impacto de movimentos adversos das taxas de juros no capital e nos resultados da instituição financeira, para os instrumentos classificados na carteira bancária.

**Risco de liquidez:** possibilidade da instituição não ser capaz de honrar eficientemente suas obrigações esperadas e inesperadas, correntes e futuras, incluindo as decorrentes de vinculação de garantias, sem afetar suas operações diárias e sem incorrer em perdas significativas.

### 2.1.1.1 RISCO OPERACIONAL

Considerando o foco deste estudo, detalharemos mais o risco operacional. Conforme citado anteriormente, trata-se da possibilidade de ocorrência de perdas resultantes de eventos externos ou de falha, deficiência ou inadequação de processos internos, pessoas ou sistemas.

Essa definição inclui o risco jurídico ou legal, relacionado à inadequação ou deficiência em contratos firmados pela instituição, a sanções em razão de descumprimento de dispositivos legais e às indenizações por danos a terceiros decorrentes das atividades desenvolvidas pela instituição financeira.

A Resolução nº 4.557/2017 do Banco Central do Brasil traz entre os riscos operacionais ([BRASIL, 2017b](#)):

- I - fraudes internas.
- II - fraudes externas.
- III - demandas trabalhistas e segurança deficiente do local de trabalho.
- IV - práticas inadequadas relativas a clientes, produtos e serviços.
- V - danos a ativos físicos próprios ou em uso pela instituição.
- VI - situações que acarretem a interrupção das atividades da instituição.
- VII - falhas em sistemas, processos ou infraestrutura de tecnologia da informação (TI).
- VIII - falhas na execução, no cumprimento de prazos ou no gerenciamento das atividades da instituição.

### 2.1.1.2 RISCO JURÍDICO

Basiléia II classificou o risco jurídico ou legal como um subconjunto do risco operacional ([SUPERVISION, 2004](#)). É possível detalhar mais o risco jurídico como decorrente do descumprimento real ou potencial de leis, normas e regulamentos incidentes sobre determinada atividade ou conjunto de atividades, com impacto direto no resultado das organizações.

A materialização do risco jurídico para as empresas enseja em perdas financeiras, de imagem, reputação e a depender da extensão, inviabilizam a execução da estratégia organizacional.

O Judiciário tem sido utilizado no Brasil como instrumento para solucionar questões que, em grande parte, poderiam ser tratadas no âmbito administrativo. Nesse sentido, o gerenciamento do risco jurídico permite a atuação proativa para que os impasses sejam evitados e quando existentes sejam prontamente mitigados para que tenham menor prejuízo a todas as partes.

É plenamente cabível que a gestão do risco jurídico seja inserida no escopo dos demais riscos, utilizando-se de metodologias e *frameworks* de identificação, mensuração e mitigação. Numa avaliação apressada pode parecer que ao corpo jurídico de uma instituição não compete a atividade de gerenciamento de riscos, contudo em análise mais minuciosa verifica-se oportunidade de ganhos com a incorporação dessa gestão.

Conforme visto, a gestão de riscos é parte de todas as atividades da empresa, sendo mais eficaz e eficiente quando adotada pelos detentores de conhecimento especializado, personalizado e que dispõe das melhores e mais atualizadas informações.

Recomenda-se, portanto, integração entre o corpo jurídico e a área de riscos das instituições na gestão do risco jurídico, além de outras áreas como a de vendas, financeira, tecnologia, logística e pessoas.

## 2.1.2 GERENCIAMENTO DE RISCOS

Gerenciar riscos faz parte do escopo das organizações sendo por vezes algo facultativo, ora obrigatório por leis e regulamentações, a depender do ramo de atuação. De toda sorte, gerenciar bem riscos é um diferencial competitivo para as instituições porque permite a adequada precificação de produtos e serviços, atuação preventiva de perdas, instituição de controles adequados e assertivos e melhora o ambiente empresarial dos países.

Segundo Hopkin (HOPKIN, 2018) os eventos que podem impactar uma organização podem inibir o que ela está buscando alcançar (risco de perigo), aumentar esse objetivo (risco de oportunidade) ou criar incerteza sobre os resultados (riscos de controle). Há uma quarta categoria de riscos relacionada a obrigações mandatórias impostas às organizações e esses riscos são chamados de riscos de conformidade.

Em sua obra atualizada, Hopkin e Thompson apresentam conceitos de gerenciamento de riscos de 4 (quatro) organizações, conforme Tabela 1 (THOMPSON; HOPKIN, 2021):

Organização	Definição
RIMS	O gerenciamento de riscos corporativos é uma disciplina estratégica de negócios que apoia a realização de objetivos de uma organização, abordando todo o espectro de seus riscos e gerenciando o impacto combinado desses riscos como um portfólio de riscos inter-relacionados (RIMS, 2022).
COSO 2017	A cultura, capacidade e prática que as organizações integram com a definição da estratégia e aplicam quando executam essa estratégia, com o objetivo de gerenciar riscos na criação, preservação e realização de valor (COSO, 2017).
IIA	Uma abordagem rigorosa e coordenada para avaliar e responder a todos os riscos que afetam a realização dos objetivos estratégicos e financeiros de uma organização (IIA, 2022).
<i>The Orange Book</i>	Atividades coordenadas, projetadas e operadas para gerenciar riscos e exercer o controle interno de uma organização (UK, 2020).

Tabela 1 – Conceitos de gerenciamento de riscos

Conforme visto, existem vários padrões de gerenciamento de riscos, porém neste estudo abordaremos em detalhes a ISO 31000 (STANDARDIZATION, 2009a) e o COSO (COSO, 2004) (COSO, 2017) diante da relevância e aplicabilidade mundiais.

### 2.1.2.1 ISO 31000

A ISO 31000 estabelece princípios, estrutura e um processo sistemático de gestão de riscos, o que fornece certo nível de padronização e traz credibilidade aos resultados. A Figura 2 ilustra essa tríade (STANDARDIZATION, 2009a).

Os princípios da gestão de riscos conforme ISO 31000 são:

**Integrado:** a gestão de riscos é parte integrante de todas as atividades organizacionais.

**Estruturado e abrangente:** uma abordagem estruturada e abrangente da gestão de riscos contribui para resultados consistentes e comparáveis.

**Personalizado:** a estrutura e o processo de gestão de riscos são customizados e proporcionais ao contexto externo e interno da organização relacionado aos seus objetivos.

**Inclusivo:** o envolvimento adequado e oportuno das partes interessadas permite que seus conhecimentos, visões e percepções sejam considerados. Isso resulta em mais conscientização e gerenciamento de risco informado.



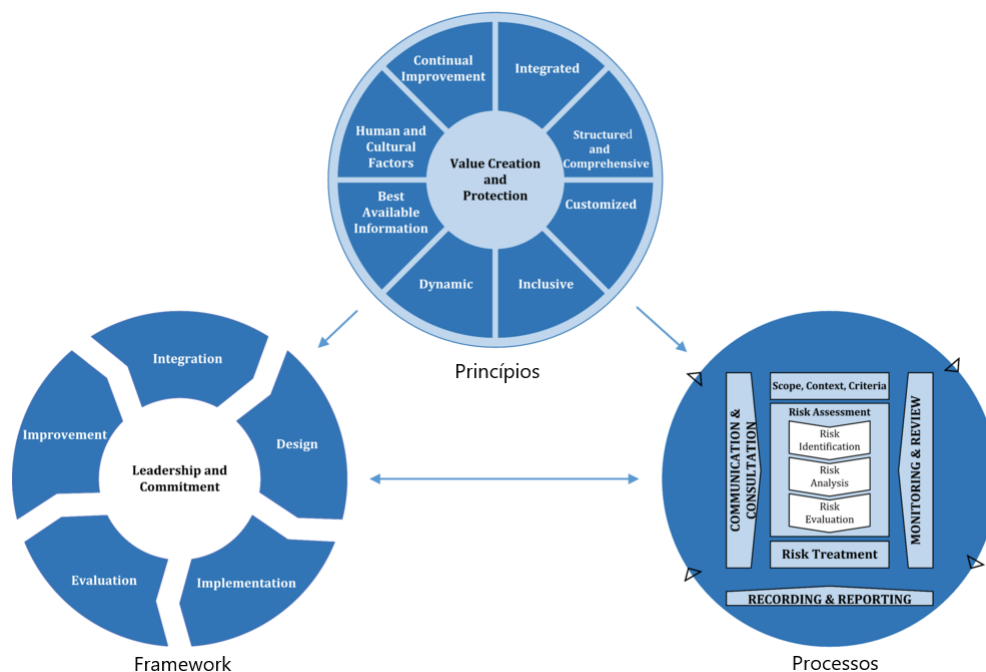


Figura 2 – ISO 31000 - *Risk management principles and guidelines*

**Dinâmico:** os riscos podem surgir, mudar ou desaparecer à medida que o contexto externo e interno de uma organização muda. A gestão de risco antecipa, detecta, reconhece e responde a essas mudanças e eventos de forma adequada e oportuna.

**Melhor informação disponível:** as entradas para a gestão de riscos são baseadas em informações históricas, atuais e futuras. As informações devem ser oportunas, claras e disponíveis para as partes interessadas relevantes.

**Fatores humanos e culturais:** o comportamento e a cultura humanos influenciam todos os aspectos da gestão de risco em cada nível e estágio.

**Melhoria contínua:** a gestão de riscos é continuamente aprimorada pelo aprendizado e experiência.

Em relação ao *framework*, a ISO 31000 estabelece como ponto central a **liderança e compromisso** da alta administração e dos órgãos de supervisão, quando aplicável, para garantirem que a gestão de riscos seja integrada em todas as atividades organizacionais. Os outros pontos do *framework* são (STANDARDIZATION, 2009a):

**Integração:** gerenciar riscos depende de uma compreensão das estruturas e do contexto organizacional. As estruturas diferem dependendo do propósito, objetivos e complexidade da organização. O risco é gerenciado em todas as partes da estrutura da organização. Todos em uma organização têm a responsabilidade de gerenciar riscos.

**Design:** entendimento da organização e do contexto em que está inserida; articulação do compromisso da alta administração e órgãos de supervisão; atribuição de papéis e responsabilidades em todos os níveis; alocação de recursos adequados (pessoas, processos, métodos, ferramentas, sistemas de informação, conhecimento); e fluxo adequado de comunicação.

**Implementação:** plano apropriado, com tempo e recursos estimados; identificar onde e como as decisões são tomadas; ajustar os processos de tomada de decisão, se necessário; garantir que o plano de gerenciamento de riscos seja claramente compreendido e praticado.

**Avaliação:** medir periodicamente o desempenho da estrutura de gerenciamento de riscos frente a seu propósito.

**Melhoria:** à medida que lacunas relevantes ou oportunidades de melhoria são identificadas, a organização deve desenvolver planos e tarefas e atribuí-los aos responsáveis pela implementação.

Quanto aos processos, a ISO 31000 estabelece a interatividade dos seguintes elementos (STANDARDIZATION, 2009a):

**Comunicação e consulta:** ajudar as partes interessadas relevantes a entenderem os riscos, a base sobre a qual as decisões são tomadas e as razões pelas quais ações específicas são necessárias. A comunicação busca promover a conscientização e a compreensão do risco, enquanto a consulta envolve a obtenção de *feedback* e informações para apoiar a tomada de decisão.

**Escopo, contexto e critério:** definir o escopo do processo, entender o contexto externo e interno, quais decisões precisam ser tomadas, quais são os resultados esperados das etapas a serem tomadas no processo, escolha de ferramentas e técnicas adequadas de avaliação de risco, estabelecer os recursos necessários, as responsabilidades e os registros a serem mantidos, verificar os relacionamentos com outros projetos, processos e atividades.

**Identificação do risco:** encontrar, reconhecer e descrever os riscos que podem ajudar ou impedir uma organização de atingir seus objetivos. Devem ser considerados: fontes de risco tangíveis e intangíveis, causas, eventos, ameaças, oportunidades, vulnerabilidades, capacidades, mudanças no contexto externo e interno, indicadores de riscos emergentes, a natureza e o valor dos ativos e recursos, consequências e seu impacto nos objetivos, limitações de conhecimento e confiabilidade das informações, fatores relacionados ao tempo, e preconceitos, suposições e crenças dos envolvidos.

**Mensuração do risco:** compreender a natureza do risco e suas características, incluindo, quando apropriado, o nível de risco. Deve considerar fatores como a probabilidade e impacto da materialização do risco.

**Avaliação do risco:** comparar os resultados com os critérios de risco estabelecidos para determinar onde uma ação adicional é necessária.

**Tratamento do risco:** selecionar e implementar opções para lidar com o risco e envolve um processo iterativo de formular e selecionar opções de tratamento de risco, planejar e implementar o tratamento de riscos, avaliar a eficácia desse tratamento, decidir se o risco remanescente é aceitável e se não, continuar o tratamento.

**Monitoramento:** garantir e melhorar a qualidade e a eficácia do projeto.

**Relatórios e reportes:** o processo de gerenciamento de riscos e seus resultados devem ser documentados e relatados por meio de mecanismos apropriados. A gravação e o relatório têm como objetivos comunicar as atividades de gestão de risco e os resultados em toda a organização, fornecer informações para a tomada de decisões, melhorar as atividades de gestão de risco e auxiliar a interação com as partes interessadas, incluindo aquelas com responsabilidade e prestação de contas pelas atividades de gestão de risco.

#### 2.1.2.2 COSO ERM

O COSO, *The Committee of Sponsoring Organizations*, é uma entidade sem fins lucrativos, dedicada à melhoria dos relatórios financeiros com foco em controles internos e governança corporativa. Suas recomendações são referência para controles internos.

O COSO ERM (COSO, 2004) trouxe 8 (oito) componentes para o gerenciamento de riscos corporativos:

**Ambiente interno:** o ambiente interno compreende o tom de uma organização e fornece a base em que os riscos são identificados e abordados, inclusive a filosofia de gerenciamento de riscos, o apetite por riscos, a integridade e os valores éticos.

**Fixação de objetivos:** os objetivos devem existir antes que a administração possa identificar os eventos em potencial que poderão afetar a sua realização. O gerenciamento de riscos corporativos assegura que a administração disponha de um processo implementado para estabelecer os objetivos que propiciem suporte e estejam alinhados com a missão da organização e sejam compatíveis com o seu apetite a riscos.

**Identificação de eventos:** os eventos internos e externos que influenciam o cumprimento dos objetivos de uma organização devem ser identificados e classificados entre riscos e oportunidades.

**Avaliação de riscos:** os riscos são analisados considerando a probabilidade e o impacto como base para determinar a forma de enfrentamento. Esses riscos são avaliados quanto a condição de inerentes e residuais.

**Resposta a risco:** a administração escolhe as respostas aos riscos - evitar, aceitar, reduzir ou compartilhar, com medidas para alinhar com o apetite por riscos.

**Atividades de controle:** políticas e procedimentos são estabelecidos e implementados para assegurar que as respostas aos riscos sejam executadas com eficácia.

**Informações e comunicações:** as informações relevantes são identificadas, colhidas e comunicadas na forma e no prazo que permitam que cumpram suas responsabilidades. A comunicação eficaz ocorre em sentido amplo em todos níveis da organização.

**Monitoramento:** a integridade da gestão de riscos corporativos é monitorada por atividades gerenciais contínuas ou avaliações independentes ou de ambas as formas.

Conforme detalha Moeller (MOELLER, 2007), no COSO 2004 (COSO, 2004) existe um relacionamento direto entre os objetivos que a organização empenha-se em alcançar e os componentes do gerenciamento de riscos corporativos, que representam aquilo que é necessário para o seu alcance.

Esse relacionamento é apresentado em uma matriz tridimensional em forma de cubo, com 3 (três) dimensões: 4 (quatro) categorias de objetivo em colunas verticais, estratégicos, operacionais, de comunicação e conformidade; os 8 (oito) componentes em linhas horizontais; e as unidades de uma organização na terceira dimensão, conforme Figura 3:

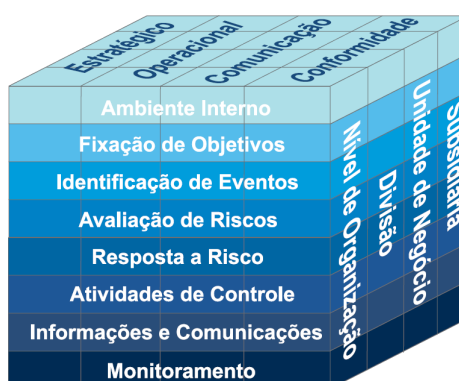


Figura 3 – Cubo do COSO ERM 2004

Em 2017 houve atualização do COSO, que passou a contar com 5 (cinco) componentes inter-relacionados, em resposta ao reconhecimento da necessidade de haver vínculos mais fortes entre estratégia, risco e desempenho (COSO, 2017):

**Governança e cultura:** a governança demonstra seu protagonismo com o reforço da importância e responsabilidades de supervisão sobre o gerenciamento de riscos corporativos. A cultura diz respeito a valores éticos, a comportamentos esperados e ao atendimento do risco em toda a entidade.

**Estratégia e definição de objetivos:** gerenciamento de riscos corporativos, estratégia e definição de objetivos atuam juntos no processo de planejamento estratégico. O apetite por riscos é estabelecido e alinhado com a estratégia; os objetivos de negócios colocam a estratégia em prática e, ao mesmo tempo, servem como base para identificar, avaliar e responder aos riscos.

**Performance:** os principais riscos precisam ser identificados e avaliados. Os resultados desse processo são comunicados aos principais *stakeholders* envolvidos na supervisão dos riscos.

**Análise e revisão:** a organização deve avaliar sua performance em gerenciamento de riscos, com a oportunidade de verificar pontos de melhoria.

**Informação, comunicação e divulgação:** gerenciar riscos demanda um processo contínuo de obtenção e compartilhamento de informações precisas, de fontes internas e externas, originadas em várias camadas e processos da organização.

A Figura 4 apresenta a atualização visual (COSO, 2017):



Figura 4 – Nova imagem do COSO ERM 2017

Segundo (THOMPSON; HOPKIN, 2021), essa estrutura atualizada do COSO ERM 2017 (COSO, 2017) adotou uma estrutura de componentes e princípios. Diferencia claramente ERM de controle interno, bem como aprimora as referências de apetite por riscos e tolerância ao risco. A intenção da estrutura revisada é elevar a discussão da estratégia, aprimorar o desempenho e vincular mais explicitamente o ERM à tomada de decisões. Há maior ênfase na relação entre risco e valor. Além disso os benefícios da integração do ERM são enfatizados. Por fim, a estrutura revisada destaca o papel da cultura na obtenção de uma gestão empresarial de riscos mais bem sucedida.

### 2.1.3 INSTITUIÇÕES FINANCEIRAS BRASILEIRAS E O RISCO JURÍDICO

As instituições financeiras mundiais sujeitam-se a regulamentações como forma de resguardar riscos de colapsos financeiros, crises e risco de contágio. Inúmeras iniciativas foram adotadas ao longo do tempo para melhorar a gestão e o controle dos riscos financeiros. Os acordos de Basileia são um exemplo

do compromisso de inúmeros bancos centrais de todo o mundo para prevenir riscos, criando exigências mínimas de reserva de capital.

A gestão e controle são realizados especialmente sobre as Demonstrações Financeiras (DF), que seguem regras e critérios rígidos, padronizados e preestabelecidos.

Os riscos relevantes devem ser reconhecidos e provisionados nas DF e é nesse contexto que destacamos o risco jurídico. A extensão das discussões judiciais é tão ampla, demorada e impactante financeiramente que a elas aplicam-se as leis, normas e regulamentações nacionais e internacionais para tratar das providências contábeis de provisão e contingenciamento nos balanços das empresas. Aos acionistas é importante prever e considerar tais estimativas nos fluxos futuros de resultados, sob pena de colocarem a sustentabilidade das empresas em risco.

Podemos considerar que a medida de mensuração quantitativa do risco jurídico é o cálculo do impacto e da probabilidade de ocorrência de perda financeira, desembolso, decorrente de condenação em processos judiciais. A esse contexto de mensuração, as instituições financeiras brasileiras, além observar as melhores práticas de gestão de riscos, sujeitam-se a leis, normas e regulamentos, como por exemplo a IAS 37 (IASB, 2001), CPC 25 (CONTÁBEIS, 2009), Resolução BCB 3823/2009 (BRASIL, 2009).

## 2.2 PROVISÃO E PASSIVO CONTINGENTE

### 2.2.1 IAS 37 - PROVISÕES, PASSIVOS CONTINGENTES E ATIVOS CONTINGENTES

O IAS 37 - *Provisions, Contingent liabilities and contingent assets* (IASB, 2001) é uma norma internacional que tem por objetivo definir critérios de reconhecimento e bases de mensuração aplicáveis a provisões, contingências passivas e contingências ativa a serem divulgadas de forma suficiente nas Demonstrações Contábeis das instituições em geral. Entre os conceitos trazidos, a IAS 37 define:

**Provisões:** passivos de prazo ou valor incertos.

**Passivos:** são obrigações presentes, provenientes de eventos passados, com expectativa de liquidação financeira.

**Obrigações:** são eventos que gerados por imposições legais ou não formalizadas.

**Contingência passiva:** uma possível obrigação presente cuja existência será confirmada somente pela ocorrência ou não de um ou mais eventos futuros, que não estejam totalmente sob o controle da entidade; ou uma obrigação presente que surge de eventos passados, mas que não é reconhecida porque:

- 1) É improvável que a entidade tenha de liquidá-la; ou
- 2) O valor da obrigação não pode ser mensurado com suficiente segurança.

Ainda sobre conceitos, a IAS 37 (IASB, 2001) caracteriza eventos prováveis, possíveis e remotos com a probabilidade de ocorrência de eventos, conforme Tabela 2:

Evento	Probabilidade
Provável	Probabilidade de ocorrência de um ou mais eventos futuros é <b>maior</b> que a de não ocorrência.
Possível	Probabilidade de ocorrência de um ou mais eventos futuros é <b>menor</b> que a de não ocorrência.
Remoto	Probabilidade de ocorrência de um ou mais eventos futuros é <b>substancialmente pequena</b> .

Tabela 2 – Eventos da IAS 37

Para cada tipo de provisão relevante, a entidade também deve divulgar:

- I - o valor contábil no início e no fim do período.

II - provisões adicionais feitas no período, incluindo aumentos nas provisões existentes.

III - montantes utilizados (ou seja, incorridos e baixados contra a provisão) durante o período.

IV - montantes não utilizados, estornados durante o período.

V - despesas financeiras apropriadas no período para as provisões ajustadas ao valor presente e qualquer mudança na taxa de desconto.

VI - uma breve descrição de sua natureza e o cronograma esperado de quaisquer desembolsos.

VI - uma indicação das incertezas sobre o valor ou o cronograma dos desembolsos. Quando for necessário fornecer informações adequadas, uma entidade deve divulgar as principais premissas adotadas em relação a eventos futuros.

VII - o montante de qualquer reembolso esperado, declarando o valor de qualquer ativo que tenha sido reconhecido por conta desse reembolso.

Exceto quando considerada remota a possibilidade de ocorrência de qualquer desembolso, a entidade deve divulgar, para cada tipo de contingência passiva relevante, uma breve descrição da natureza da contingência passiva e, quando praticável:

I - uma estimativa do efeito financeiro.

II - uma indicação das incertezas relacionadas ao montante ou ao tempo de qualquer desembolso.

III - a possibilidade de qualquer reembolso.

#### 2.2.1.1 PRONUNCIAMENTO TÉCNICO CPC 25

O Pronunciamento Técnico CPC 25 é uma correlação brasileira às Normas Internacionais de Contabilidade previstas na *IAS 37 - Provisions, Contingent liabilities and contingent assets* (IASB, 2001).

O objetivo do CPC 25 é assegurar a aplicação de critérios de reconhecimento e bases de mensuração para provisões, passivos contingentes e ativos contingentes e que seja divulgada informação suficiente nas notas explicativas, para permitir que os interessados verifiquem sua natureza, oportunidade e valor (CONTÁBEIS, 2009).

Contabilmente a **provisão** “é um passivo de prazo ou valor incertos”, conceituada no CPC 25 como situações advindas de resultados de eventos passados que podem ensejar em saída financeira para liquidação de uma obrigação presente ou obrigação possível, cuja existência será confirmada apenas pela ocorrência de um ou mais eventos incertos. A Figura 5 sintetiza os momentos dos fatos geradores:

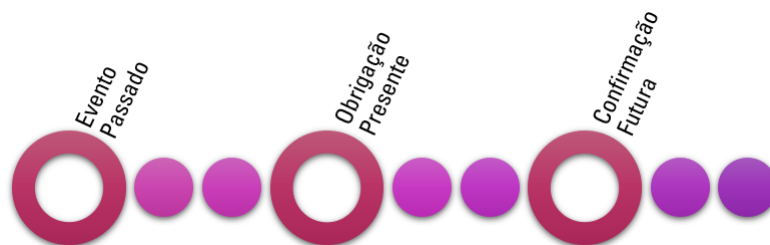


Figura 5 – CPC 25 - Fatos geradores

Outro conceito importante do CPC 25 é o de passivos contingentes, caracterizado por:

(a) uma obrigação possível que resulta de eventos passados e cuja existência será confirmada apenas pela ocorrência, ou não, de um ou mais eventos futuros incertos não totalmente sob o controle da

entidade; ou

(b) uma obrigação presente que resulta de eventos passados, mas que não é reconhecida porque não é provável que uma saída de recursos que incorporam benefícios econômicos seja exigida para liquidar a obrigação ou o valor da obrigação não pode ser mensurado com suficiente confiabilidade.

Contabilmente não se deve reconhecer um passivo contingente, contudo deve-se divulgá-lo em notas explicativas das Demonstrações Financeiras. Uma entidade deve divulgar um passivo contingente, salvo se a possibilidade de saída de recursos que incorporam benefícios econômicos seja remota.

Para clarificar os conceitos de provisão e passivo contingente, elaboramos na Figura 6 uma árvore de decisão baseada na probabilidade de ocorrência de desembolso financeiro:

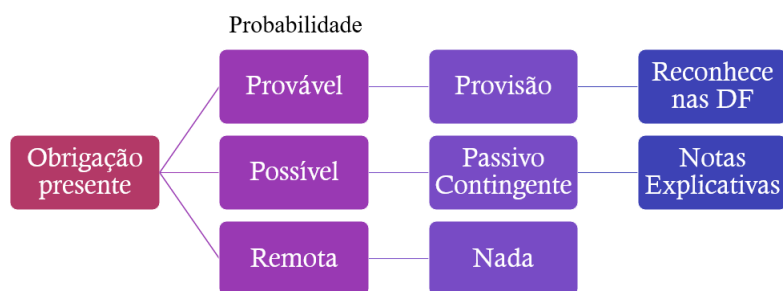


Figura 6 – CPC 25 - Provisão e passivo contingente

### 2.2.1.2 RESOLUÇÃO BCB 3.823/2009

A Resolução nº 3.823 do Banco Central do Brasil dispõe sobre procedimentos aplicáveis no reconhecimento, mensuração e divulgação de provisões, contingências passivas e contingências ativas (BRASIL, 2009):

*Art. 1º As instituições financeiras e as demais instituições autorizadas a funcionar pelo Banco Central do Brasil devem observar o Pronunciamento Técnico CPC 25, emitido pelo Comitê de Pronunciamentos Contábeis (CPC), no reconhecimento, mensuração e divulgação de provisões, contingências passivas e contingências ativas.*

*Art. 2º As instituições de que trata o art. 1º devem manter à disposição do Banco Central do Brasil, pelo prazo de cinco anos, toda a documentação e detalhamento utilizados no reconhecimento, mensuração e divulgação de provisões, contingências passivas e contingências ativas.*

*Art. 3º Verificada impropriedade ou inconsistência nos processos de classificação, divulgação e registro contábil das provisões, contingências passivas e contingências ativas, o Banco Central do Brasil poderá determinar os ajustes necessários, com o conseqüente reconhecimento contábil dos efeitos nas demonstrações contábeis.*

## 2.3 BENCHMARK DE RISCO JURÍDICO E PROVISÃO EM INSTITUIÇÕES FINANCEIRAS

### 2.3.1 INSTITUIÇÕES FINANCEIRAS BRASILEIRAS

As Instituições financeiras apresentam provisões e passivos contingentes decorrentes principalmente de processos judiciais e administrativos, intrínsecos ao curso normal dos seus negócios, movidos por terceiros, empregados e órgãos públicos em ações cíveis, trabalhistas e fiscais.

Esses processos são encerrados após um longo período e envolvem discussões quanto ao mérito e aspectos processuais. Além dos aspectos subjetivos na determinação da possibilidade de perda atribuída a cada processo, a evolução da jurisprudência sobre determinadas causas nem sempre é uniforme.

É necessário que as IF atuem para a identificação, avaliação, monitoramento, mensuração, registro das provisões e divulgação dos passivos contingentes, considerando a totalidade e integridade da base de dados.

Modelos estatísticos podem ser utilizados para quantificar processos judiciais massificados. A avaliação individualizada deve ser realizada para processos de valores relevantes, com o apoio de especialistas. Neste trabalho conceituamos processos massificados como aqueles relativos a ações semelhantes e cujo valor individual não seja relevante.

Para fins de *benchmark*, utilizaremos como referencial a Resolução nº 4.553/2017 do Banco Central do Brasil (BRASIL, 2017a), que estabelece a segmentação das instituições financeiras brasileiras. Por essa Resolução é possível identificar os maiores bancos do país a partir do segmento "S1".

O S1 é composto pelos bancos múltiplos, bancos comerciais, bancos de investimento, bancos de câmbio e caixas econômicas que tenham porte igual ou superior a 10% (dez por cento) do Produto Interno Bruto (PIB) ou que exerçam atividade internacional relevante, independentemente do porte da instituição.

Em consulta ao site do BCB (BRASIL, 2022), verificamos a categorização de 6 (seis) IF para o Segmento 1, em ordem alfabética:

BANCO DO BRASIL

BRADESCO

BTG PACTUAL

CAIXA ECONÔMICA FEDERAL

ITAU

SANTANDER

Neste trabalho excluiremos o banco BTG Pactual em razão de sua característica predominante ser a de Banco de Investimentos e focaremos nos bancos comerciais Banco do Brasil, Bradesco, Caixa, Itaú e Santander.

### 2.3.1.1 BALANÇOS PATRIMONIAIS

#### **BANCO DO BRASIL**

O Banco do Brasil é um banco múltiplo com atuação em todo o território nacional, que desenvolve algumas atividades em centros financeiros mundiais. Tem por objeto a prestação de serviços bancários, de intermediação e suprimento financeiro sob suas múltiplas formas, inclusive nas operações de câmbio e nas atividades complementares, destacando-se seguros, previdência privada, capitalização, corretagem de títulos e valores mobiliários, administração de cartões de crédito/débito, consórcios, fundos de investimentos e carteiras administradas e o exercício de quaisquer atividades facultadas às instituições integrantes do Sistema Financeiro Nacional (BRASIL, 2021).

O valor total dos ativos do Banco do Brasil é aproximadamente de **US\$ 427 bilhões de dólares**, conforme Tabela 3:



ATIVO	PASSIVO	ATIVO	PASSIVO
R\$ 1.988.645.574	R\$ 1.854.419.676	US\$ 427.822.123	US\$ 398.945.782
	PL R\$ 134.225.898		PL US\$ 28.876.341
R\$ 1.988.645.574	R\$ 1.988.645.574	US\$ 427.822.123	US\$ 427.822.123

Valores em mil.

Conversão: 13/04/22, [www.bcb.gov.br/conversao](http://www.bcb.gov.br/conversao).

Tabela 3 – Balanço Patrimonial do Banco do Brasil 2021

### CAIXA ECONÔMICA FEDERAL

A Caixa Econômica Federal é uma instituição financeira que desenvolve atividades bancárias de captação, carteiras comerciais, operações de câmbio, crédito ao consumidor, imobiliário e rural, prestação de serviços bancários, negócios com cartões de débito e crédito, administração de fundos e carteiras de investimento, intermediação de títulos e valores mobiliários, seguros, previdência privada, capitalização, administração de consórcios, administração em caráter de exclusividade dos serviços das loterias federais e de operações de penhor de jóias (CAIXA, 2021).

O valor total dos ativos CAIXA é aproximadamente de **US\$ 312 bilhões de dólares**, conforme Tabela 4:

ATIVO	PASSIVO	ATIVO	PASSIVO
R\$ 1.450.370.927	R\$ 1.374.215.958	US\$ 312.021.799	US\$ 295.638.397
	PL R\$ 76.154.969		PL US\$ 16.383.402
R\$ 1.450.370.927	R\$ 1.450.370.927	US\$ 312.021.799	US\$ 312.021.799

Valores em mil.

Conversão: 13/04/22, [www.bcb.gov.br/conversao](http://www.bcb.gov.br/conversao).

Tabela 4 – Balanço Patrimonial da CAIXA 2021

### ITAÚ

O Itaú está presente em 18 (dezoito) países e fornece produtos e serviços financeiros em todas as modalidades nas carteiras: comercial, de investimento, imobiliário, investimento, de arrendamento mercantil e de operações de câmbio. Suas operações são divididas em três segmentos: Banco de Varejo, Banco de Atacado e Atividades com Mercado e Corporação (ITAÚ, 2021).

O valor total dos ativos do Itaú é aproximadamente de **US\$ 445 bilhões de dólares**, conforme Tabela 5:

ATIVO	PASSIVO	ATIVO	PASSIVO
R\$ 2.069.206.000	R\$ 1.904.730.000	US\$ 445.153.282	US\$ 09.769.163
	PL R\$ 164.476.000		PL US\$ 35.384.119
R\$ 2.069.206.000	R\$ 2.069.206.000	US\$ 445.153.282	US\$ 445.153.282

Valores em mil.

Conversão: 13/04/22, [www.bcb.gov.br/conversao](http://www.bcb.gov.br/conversao).

Tabela 5 – Balanço Patrimonial do Itaú 2021

## BRADESCO

O Bradesco opera na forma de Banco Múltiplo, desenvolve atividades bancárias em todas as modalidades autorizadas, com carteiras comerciais, de operações de câmbio, de crédito ao consumidor e de crédito imobiliário. Por suas controladas atua direta e indiretamente em diversas outras atividades com destaque para arrendamento mercantil, banco de investimentos, corretora de títulos e valores mobiliários, administração de consórcios, cartões de crédito, empreendimentos imobiliários, seguros, previdência e capitalização (BRADESCO, 2021).

O valor total dos ativos do Bradesco é aproximadamente de **US\$ 355 bilhões de dólares**, conforme Tabela 6:

ATIVO	PASSIVO	ATIVO	PASSIVO
R\$ 1.653.665.568	R\$ 1.505.830.602	US\$ 355.757.066	US\$ 323.952.973
	PL R\$ 147.834.966		PL US\$ 31.804.093
R\$ 1.653.665.568	R\$ 1.653.665.568	US\$ 355.757.066	US\$ 355.757.066

Valores em mil.

Conversão: 13/04/22, [www.bcb.gov.br/conversao](http://www.bcb.gov.br/conversao).

Tabela 6 – Balanço Patrimonial do Bradesco 2021

## SANTANDER

O Santander opera como banco múltiplo e desenvolve as operações em carteiras comerciais, de investimento, de crédito imobiliário, arrendamento mercantil e câmbio. Com empresas controladas atua nos mercados de instituição de pagamento, administração de consórcios, corretagem de valores mobiliários, corretagem de seguros, financiamento ao consumo, plataformas digitais, gestão de benefícios, recuperação de créditos não performados, capitalização, previdência privada, e administração de vales alimentação, refeição e outros (SANTANDER, 2021)

O valor total dos ativos do Santander é aproximadamente de **US\$ 211 bilhões de dólares**, conforme Tabela 7:

ATIVO	PASSIVO	ATIVO	PASSIVO
R\$ 980.816.744	R\$ 901.972.673	US\$ 211.005.474	US\$ 194.043.559
	PL R\$ 78.844.071		PL US\$ 16.961.915
R\$ 980.816.744	R\$ 980.816.744	US\$ 211.005.474	US\$ 211.005.474

Valores em mil.

Conversão: 13/04/22, www.bcb.gov.br/conversao.

Tabela 7 – Balanço Patrimonial do Santander 2021

### 2.3.1.2 PROVISÕES E PASSIVOS CONTINGENTES DAS IF BRASILEIRAS

Em pesquisa nas Demonstrações Financeiras de 2021 dos 5 (cinco) maiores bancos brasileiros, informações públicas, foi possível verificar os montantes para as provisões e passivos contingentes relacionados ao risco jurídico.

Em 2021 o montante para provisões foi de **R\$ 80 bilhões de reais**, enquanto que para os passivos contingentes o valor foi de **R\$ 79 bilhões**, conforme Tabelas 8 e 9 :

Provisões	BB	CAIXA	ITAÚ	BRADESCO	SANTANDER
Cíveis	11.303.452	3.279.719	3.317.000	9.178.471	3.092.507
Trabalhistas	4.684.907	6.991.311	8.219.000	6.729.107	1.941.169
Fiscais	537.780	378.813	6.498.000	8.072.037	4.312.234
Outros riscos	não informado	não informado	1.558.000	não informado	não informado
<b>Total</b>	16.526.139	10.649.843	19.592.000	23.979.615	9.345.910

Valores em R\$ mil

Tabela 8 – Provisões IF brasileiras para risco jurídico

Passivo Contingente	BB	CAIXA	ITAÚ	BRADESCO	SANTANDER
Cíveis	2.248.542	1.985.291	4.903.000	7.979.276	2.380.000
Trabalhistas	87.689	não informado	448.000	não informado	267.000
Fiscais	12.192.801	8.966.993	35.855	37.556.235	29.726
<b>Total</b>	14.529.032	10.952.284	5.386.855	45.535.511	2.676.726

Valores em R\$ mil

Tabela 9 – Passivo contingente IF brasileiras para risco jurídico

As ações cíveis decorrem de pleitos relacionados a revisão de contratos e de ações de indenização por danos materiais e morais. As trabalhistas da discussão de pretensos direitos trabalhistas específicos à categoria profissional, tais como: horas extras, equiparação salarial, reintegração, adicional de transferência, complemento de aposentadoria, entre outros. Já as ações fiscais correspondem ao valor principal dos tributos envolvidos em discussões fiscais administrativas ou judiciais, objeto de lançamento de ofício, acrescido de juros, multa e, encargos, quando aplicável.

A título de comparação, preparamos as Tabelas 10 e 11 com os montantes de provisões e passivos contingentes das instituições financeiras brasileiras convertidos em dólar, que totalizaram **US\$ 17 bilhões** e **US\$ 16,8 bilhões de dólares** respectivamente:

Provisões	BB	CAIXA	ITAÚ	BRADESCO	SANTANDER
Cíveis	2.431.739	700.630	708.594	1.960.751	660.637
Trabalhistas	1.007.875	1.493.519	1.755.784	1.437.506	414.682
Fiscais	115.694	80.924	1.388.135	1.724.389	921.201
Outros riscos	não informado	não informado	332.828	não informado	não informado
<b>Total</b>	<b>3.555.308</b>	<b>2.275.073</b>	<b>4.185.341</b>	<b>5.122.646</b>	<b>1.996.520</b>

Valores em US\$ mil.

Conversão: 13/04/22, www.bcb.gov.br/conversao.

Tabela 10 – Provisões IF brasileiras para risco jurídico em dólar

Passivo Contingente	BB	CAIXA	ITAÚ	BRADESCO	SANTANDER
Cíveis	480.345	424.108	1.047.403	1.704.573	508.428
Trabalhistas	18.733	não informado	95.704	não informado	57.038
Fiscais	2.604.687	1.915.574	7.660	8.022.951	6.350
<b>Total</b>	<b>3.103.765</b>	<b>2.339.681</b>	<b>1.150.767</b>	<b>9.727.524</b>	<b>571.816</b>

Valores em US\$ mil.

Conversão: 13/04/22, www.bcb.gov.br/conversao.

Tabela 11 – Passivo contingente IF brasileiras para risco jurídico em dólar

## 2.3.2 INSTITUIÇÕES FINANCEIRAS INTERNACIONAIS

### 2.3.2.1 JP MORGAN CHASE

JP Morgan é considerado o maior banco dos Estados Unidos. A seguir apresentamos resumo de seu Balanço Patrimonial de 2021 (CHASE, 2021), com ativos na casa de **US\$ 3,7 trilhões de dólares**, Tabela 12:

ATIVO	PASSIVO
US\$ 3.743.567	US\$ 3.449.440
	PL
	US\$ 294.127
US\$ 3.743.567	US\$ 3.743.567

Valores milhão.

Tabela 12 – Balanço Patrimonial JP Morgan Chase 2021

Em análise das Demonstrações Financeiras de 2021, identificamos a nota explicativa nº 30 - Contencioso Contingências:

*"Em 31 de dezembro de 2021, a Firma, suas subsidiárias e afiliadas são réus em vários processos legais, incluindo litígios privados, civis, investigações ou questões de aplicação regulatória.*

*Os litígios vão desde ações individuais envolvendo um único demandante a ações coletivas com potencialmente milhões de membros da classe.*

*Esses processos judiciais estão em estágios variados de adjudicação, arbitragem ou investigação, e envolvem cada uma das linhas de negócios da Firma, diversas regiões geográficas e uma ampla variedade de reivindicações (incluindo delito de direito comum, reivindicações contratuais, antitruste estatutária, valores mobiliários e reivindicações de proteção ao consumidor), algumas das quais apresentam teorias jurídicas.*

### PERDA PROVÁVEL

O JP Morgan registra que estabeleceu reservas para os processos atualmente pendentes de acordo com as disposições do US GAAP para contingências. Nesses casos provisionam um passivo para litígios quando a perda for provável e o valor possa ser razoavelmente estimado.

A cada trimestre avaliam os processos judiciais pendentes, verificam a suficiência de suas reservas para litígios e realizam ajustes a maior ou a menor, com base no julgamento da administração e após consulta ao corpo jurídico.

A reserva para a despesa com litígios judiciais para o ano de 2021 foi de **US\$ 426 milhões** e de **US\$ 1,1 bilhão de dólares** para os exercícios de 2021, 2020 e 2019.

### PERDA POSSÍVEL

O JP Morgan acredita que a estimativa da faixa agregada de perdas razoavelmente possíveis, além das reservas estabelecidas para o seu processo legal é de **US\$ 0** a aproximadamente **US\$ 1,5 bilhão de dólares** em 31 de dezembro de 2021.

### PERDA REMOTA

O banco JP Morgan descreve ainda, na nota explicativa nº 30, as justificativas para a impossibilidade de realização de estimativas, em linha com o que preconiza a IAS 37 (IASB, 2001):

- Número, variedade e estágios distintos dos processos, incluindo o fato de que muitos estão em etapas preliminares;
- Existência de múltiplos réus cuja parcela de responsabilidade ainda não foi determinada;
- Inúmeras questões ainda não resolvidas em muitos dos procedimentos, incluindo questões relativas ao escopo de muitas reivindicações; e
- Incerteza dos vários potenciais resultados de tais processos, inclusive quando a IF fez suposições sobre decisões futuras pelo tribunal ou outro juiz, ou sobre o comportamento ou incentivos de partes adversas ou autoridades reguladoras e essas suposições foram incorretas.

#### 2.3.2.2 BANK OF AMERICA

O *Bank of America Corporation* fornece uma gama diversificada de serviços e produtos financeiros nos EUA e em alguns mercados internacionais. É um o 2o maior banco americano, com ativos na ordem de **US\$ 3,1 trilhões de dólares** (AMERICA, 2021):

ATIVO	PASSIVO
US\$ 3.169.495	US\$ 2.899.429
	PL
	US\$ 270.066
US\$ 3.169.495	US\$ 3.169.495

Valores milhão.

Tabela 13 – Balanço Patrimonial Bank of America 2021

Em análise das Demonstrações Financeiras de 2021, identificamos a nota explicativa nº 12 - *Litigation and Regulatory Matters* (Contencioso e assuntos regulatórios), informações de risco jurídico, perda possível, provável e remota:

No curso normal dos negócios, a Corporação e suas subsidiárias são rotineiramente réus ou partes em muitas ações e processos legais, regulatórios e governamentais pendentes e ameaçados. Em vista da dificuldade inerente de prever o resultado de tais questões, particularmente quando os reclamantes buscam danos muito grandes ou indeterminados ou quando as questões apresentam novas teorias jurídicas ou envolvem um grande número de partes, a Corporação geralmente não pode prever o resultado final do processo, pendências, prazo para a resolução final dessas questões, ou eventual perda, multas ou penalidades relacionadas a cada pendência.

"À medida que o assunto se desenvolve, a Corporação, em conjunto com qualquer advogado externo que esteja lidando com o assunto, avalia se tal assunto apresenta uma contingência de perda provável e estimável e, para os assuntos descritos abaixo, se uma perda além de qualquer passivo acumulado é razoavelmente possível em períodos futuros. Uma vez que a contingência de perda seja considerada provável e estimável, a Corporação estabelecerá um passivo provisionado e registrará um valor correspondente de despesas relacionadas a litígios. A Corporação continua monitorando o assunto em busca de desenvolvimentos que possam afetar o valor do passivo provisionado que foi estabelecido anteriormente" (AMERICA, 2021).

#### **PERDA PROVÁVEL**

A despesas (provisões) relacionadas a litígios em 2021 é de **US\$ 164 milhões** e de **US\$ 823 milhões de dólares** em 2020.

#### **PERDA POSSÍVEL**

Para qualquer assunto para o qual uma perda em períodos futuros seja razoavelmente possível e estimável e para exposições de declarações e garantias, a faixa estimada de perda possível é de **US\$ 0 a US\$ 1,0 bilhão de dólares** em excesso do passivo acumulado, se houver, em 31 de dezembro de 2021.

"O passivo acumulado e o intervalo estimado de perda possível são baseados em informações atualmente disponíveis e sujeitos a julgamento significativo, uma variedade de premissas e incertezas conhecidas e desconhecidas. As questões subjacentes ao passivo provisionado e a faixa estimada de perda possível são imprevisíveis e podem mudar de tempos em tempos e as perdas reais podem variar significativamente da estimativa e provisão atuais".(AMERICA, 2021)

#### **BALANÇO PATRIMONIAL - PASSIVO**

No balanço patrimonial de 2021, o *Bank of America* apresenta no passivo o valor de **US\$ 200,42 milhões de dólares** para despesas acumuladas e outros passivos, item em que acreditamos estarem incorporadas as provisões de anos anteriores (*Accrued expenses and other liabilities*).

#### **2.3.2.3 MUFJ BANK – MITSUBISHI UFJ FINANCIAL GROUP**

O *MUFJ Bank* é o maior banco japonês e o quinto maior banco do mundo, com ativos totais no valor **US\$ 2,5 trilhões de dólares**:

ATIVO	PASSIVO	ATIVO	PASSIVO
¥ 321.292.847	¥ 305.216.168	US\$ 2.558.878	US\$ 2.430.839
	PL ¥ 16.076.679		PL US\$ 128.039
¥ 321.292.847	¥ 321.292.847	US\$ 2.558.878	US\$ 2.558.878

Valores ¥ em bilhão, US\$ milhão.

Conversão: 11/04/22, www.bcb.gov.br/conversao.

Tabela 14 – Balanço Patrimonial *MUFJ Bank* 2021

Ao avaliar as demonstrações financeiras do *MUFJ Bank* verificamos um ponto que pode ser traduzido como uma diferença cultural e/ou do modelo judicial japonês: nota-se a pouca judicialização das demandas decorrentes de litígios e questões regulatórias, o que ocasiona ausência de provisões e estimativas para passivos contingentes nas demonstrações financeiras.

Na nota explicativa nº 26 - Passivos contingentes (Contingent liabilities) consta ([GROUP, 2021](#)):

*"No curso normal dos negócios, o Grupo MUFJ está sujeito a diversos litígios e questões regulatórias. De acordo com a orientação contábil aplicável, o Grupo MUFJ estabelece um passivo acumulado para perdas e contingências decorrentes de litígios e questões regulatórias quando forem determinadas como prováveis e o valor provável da perda possa ser razoavelmente estimado.*

*Com base no conhecimento atual e consulta com o advogado, a administração acredita no resultado final de tais litígios e questões regulatórias, onde as perdas são prováveis e os montantes das perdas prováveis podem ser razoavelmente estimados, não têm impacto na posição financeira, resultados operacionais ou fluxos de caixa do Grupo MUFJ.*

*Adicionalmente, a administração acredita que o valor da perda é razoavelmente possível, mas não provável, de vários litígios e questões regulatórias não são relevantes para a posição financeira, resultados operacionais ou fluxos de caixa do Grupo MUFJ."*

#### 2.3.2.4 ICBC - INDUSTRIAL AND COMMERCIAL BANK OF CHINA

O *ICBC - Industrial and Commercial Bank of China* é o maior banco do mundo, com ativos totais no valor de **US\$ 5,5 trilhões de dólares**:

ATIVO	PASSIVO	ATIVO	PASSIVO
RMB 35.171.383	RMB 31.896.125	US\$ 5.509.043	US\$ 4.996.026
	PL RMB 3.275.258		PL US\$ 513.017
RMB 35.171.383	RMB 35.171.383	US\$ 5.509.043	US\$ 5.509.043

Valores em milhão.

Conversão: 11/04/22, www.bcb.gov.br/conversao.

Tabela 15 – Balanço Patrimonial *ICBC* 2021

Ao avaliarmos as demonstrações financeiras do *ICBC*, nota explicativa nº 7.4.14 Compromissos e passivos contingentes (*Commitments and contingent liabilities*) ([INDUSTRIAL; CHINA, 2021](#)):

*"Nos processos judiciais e de arbitragens o Grupo está envolvido em processos judiciais e arbitragens durante o curso normal de suas operações. Em 31 de dezembro de 2021, havia vários processos*

judiciais e arbitragens em aberto contra o Banco e/ou suas subsidiárias com um valor total reivindicado de RMB 6.165 milhões (31 de dezembro de 2020: RMB 4.928 milhões).

*Na opinião da administração, o Grupo constituiu provisão adequada para quaisquer prováveis perdas com base nos fatos e circunstâncias atuais, e o resultado final desses processos judiciais e arbitragens não terão impacto significativo sobre a situação financeira ou operações do Grupo."*

O valor reivindicado nas ações judiciais contra o ICBC de **RMB 6.165 milhões de remimbis** equivale a **US\$ 1.311 milhões de dólares**. Não há mais detalhamento a respeito de provisão e contingências nesse sentido, visto que a Instituição entende que os riscos legais relacionados não são relevantes.

### 2.3.2.5 BNP PARIBAS

O banco BNP Paribas é o maior banco europeu, com ativos totais no valor de **US\$ 2,8 trilhões de dólares**:

ATIVO	PASSIVO	ATIVO	PASSIVO
€ 2.634.444	€ 2.511.937	US\$ 2.858.635	US\$ 2.725.703
	PL € 122.507		PL US\$ 132.932
€ 2.634.444	€ 2.634.444	US\$ 2.858.635	US\$ 2.858.635

Valores em milhão.

Conversão: 11/04/22, [www.bcb.gov.br/conversao](http://www.bcb.gov.br/conversao).

Tabela 16 – Balanço Patrimonial BNP Paribas 2021

Quanto aos riscos jurídicos, o BNP - Paribas traz na nota explicativa nº 7.b - Procedimentos Legais e Arbitragem (Legal Proceedings and Arbitration) ([PARIBAS, 2021](#)):

*"O Banco é parte como réu em várias ações, disputas e processos judiciais (incluindo investigações judiciais ou autoridades supervisoras) em várias jurisdições decorrentes do curso de seus negócios, incluindo, entre outros, em conexão com suas atividades como contraparte de mercado, credor, empregador, investidor e contribuinte."*

#### PERDA PROVÁVEL

O valor total provisionado no balanço patrimonial da BNP - Paribas em 2021 é de **€ 992 milhões de euros**, sendo **€ 519 milhões** vindos do balanço de 2020, **€ 558 milhões** de novas provisões realizadas em 2021 e deduzidos do valor utilizado de (112) milhões.

O Banco ainda traz a seguinte informação: *Os riscos relacionados foram avaliados pelo Banco e estão sujeitos às provisões para passivos e encargos das demonstrações financeiras. Uma provisão é reconhecida quando for provável que uma saída de recursos que incorporam benefícios econômicos serão necessários para liquidar um obrigação decorrente de um evento passado e uma estimativa confiável pode ser feita do valor da obrigação.*

#### PERDA POSSÍVEL

Na nota explicativa nº 7.b, o BNP Paribas registra que algumas de suas subsidiárias são rés em diversas ações pendentes perante o Tribunal de Falências dos Estados Unidos no valor possível de **US\$ 1,3 bilhão de dólares**.



### 3 BASE DE DADOS

Utilizamos neste estudo uma base de dados com informações de processos judiciais de uma instituição financeira brasileira caracterizada como ré, ou seja, acusada de culpa, e que tiveram ou não algum tipo de desembolso financeiro no decorrer de vigência do processo por condenação total, parcial ou acordo.

Utilizamos várias bases de dados e realizamos as consultas com a chave "número do processo judicial". A seguir algumas informações quantitativas das principais bases utilizadas:

- Base de processos judiciais : 8.498.553 processos.
- Base de pagamentos realizados : 13.234.506 pagamentos (total ou parcial).
- Base de fases percorridas no processo : 151.691.446 registros de fases.

Dada a volumetria das bases, entendemos necessário realizar alguns filtros, que resultaram numa base de dados de **403.921 processos** processos judiciais:

- Apenas processos extintos nos últimos 36 (trinta e seis) meses, período de 01/12/2018 a 01/12/2021.
- Assuntos da própria IF, excluídos os assuntos externos em que há prestação de serviços a terceiros como empresas, governos, fundos.
- Processos em que a IF é caracterizada como ré, excluídos os processos em que a IF ingressou contra alguma parte.
- Apenas as áreas judicial fiscal, cível (feitos diversos) e trabalhista.

#### 3.1 SELEÇÃO DA AMOSTRA

As informações foram extraídas da base de dados de 403.921 registros e para o adequado processamento dos modelos foi necessário excluir os registros em branco, nulos, com valores diferentes da expectativa (valor da causa = 0), bem como aqueles com inconsistência no registro das informações, como por exemplo processos com duração negativa e parte contrária sem informação se pessoa física ou jurídica.

A base de dados da amostra, após as exclusões, totalizou 398.207 registros, sendo que 48.74% tiveram desembolso financeiro, enquanto 51.26% não tiveram, 67.6% foram propostos exclusivamente por pessoas físicas, 4.6% exclusivamente por pessoas jurídicas e 27.8% por ambos, conforme Figura 7:

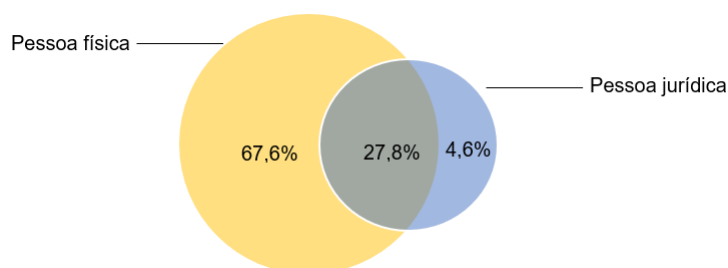


Figura 7 – Seleção da amostra

A amostra soma, em valores atualizados, o pagamento de R\$ 7,8 bilhões de reais nos 36 (trinta e seis) meses do estudo, o que confirma a importância da estruturação de modelos preditivos para que movimentos de provisão sejam adequadamente realizados, conforme regulamentações, leis e melhores práticas, bem como sejam adotados movimentos de mitigação de riscos jurídicos.

Para a modelagem, a amostra foi subdividida em base treino e base teste na proporção de 80% e 20% e totalizaram 318.566 e 79.641 registros respectivamente.

## 3.2 TRATAMENTO DAS VARIÁVEIS

Dentro do universo da base de dados, centenas de variáveis foram analisadas, testadas e excluídas. Essa seleção exigiu avaliação pormenorizada e consumiu parte relevante do tempo deste estudo. Verificamos que um dos maiores desafios da modelagem de risco operacional, no qual está inserido o risco jurídico, foi entender, reunir e concatenar as informações numa base de dados.

Vale destacar que todas as variáveis são internas, capturadas dos processos judiciais em que a instituição financeira foco da análise é parte, ou seja, não foram objeto deste estudo informações do cenário jurídico do país, de estratégia, orientações e jurisprudências do Poder Judiciário.

O tratamento das variáveis permitiu separá-las em dois conceitos relacionados à disponibilidade, facilidade de extração e tratamento: 1) Cadastro e 2) Andamento do processo.

As variáveis de cadastro caracterizam-se por serem objetivas, indiscutíveis, sem juízo de valor e disponíveis desde o primeiro momento do processo judicial. Extraímos as variáveis de cadastro da **Base de processos judiciais**.

As variáveis de andamento do processo são subjetivas, pois dependem da avaliação e do comportamento das partes e juízes, disponíveis conforme andamento e evolução do processo judicial. Destaca-se que para processos com as mesmas características podem existir fases processuais e julgamentos distintos. Extraímos as variáveis da **Base de fases percorridas no processo**.

## 3.3 SELEÇÃO DAS VARIÁVEIS

Após o tratamento e entendimento das variáveis e bases, das 36 (trinta e seis) opções de variáveis da **Base de processos judiciais** e 481 (quatrocentas e oitenta e uma) opções da **Base de fases percorridas**, selecionamos 25 (vinte e cinco) variáveis: 13 (treze) de cadastro e 12 (doze) relacionadas ao andamento dos processos judiciais.

Excluimos na seleção das variáveis de cadastro aquelas com a maior parte dos campos nulos e com informações equivalentes, como "valor da causa" e "valor da causa atualizado". Quanto às variáveis de andamento processual, foi a parte mais trabalhosa da seleção, que envolveu a análise de 481 (quatrocentas e oitenta e uma) tipologias de fases e exclusão daquelas com forte correlação com a variável dependente.

Quanto à variável dependente deste estudo, essa se relaciona com a ocorrência ou não de pagamento financeiro em algum momento durante a vigência do processo judicial, nominada de "Houve pagamento", com resposta binária de 0 - Não e 1 - Sim. As demais são do tipo independentes.

Na Tabela 17 apresentamos todas as variáveis selecionadas com as respectivas características e tipos:

Variável	Característica	Tipo
1. Duração do processo	Cadastro	Independente
2. Unidade Jurídica	Cadastro	Independente
3. Área Judicial	Cadastro	Independente
4. Assunto	Cadastro	Independente
5. Quantidade de PF	Cadastro	Independente
6. Quantidade de PJ	Cadastro	Independente
7. Base / Região do Judiciário	Cadastro	Independente
8. Comarca	Cadastro	Independente
9. Vara	Cadastro	Independente
10. Foro	Cadastro	Independente
11. Valor da causa atualizado	Cadastro	Independente
12. Tipo de ação	Cadastro	Independente
13. Macro assunto	Cadastro	Independente
14. Houve pagamento	Andamento	Dependente
15. Acórdão desfavorável (fase 3)	Andamento	Independente
16. Acórdão favorável (fase 4)	Andamento	Independente
17. Acórdão parcialmente favorável (fase 5)	Andamento	Independente
18. Acórdão retorno para julgamento do mérito (fase 6)	Andamento	Independente
19. Sentença desfavorável (fase 98)	Andamento	Independente
20. Sentença favorável (fase 99)	Andamento	Independente
21. Sentença parcialmente favorável (fase 100)	Andamento	Independente
22. Sentença incompetência de foro (fase 112)	Andamento	Independente
23. Sentença homologatória na execução (fase 164)	Andamento	Independente
24. Conversão em execução (fase 182)	Andamento	Independente
25. Acórdão para correção de erro material (fase 346)	Andamento	Independente

Fonte: Base de dados da IF estudada

Tabela 17 – Variáveis selecionadas

### 3.3.1 VARIÁVEL DEPENDENTE

A variável dependente, “Houve pagamento”, representa o evento de desembolso financeiro durante a vigência do processo judicial, seja por acordo ou decisão do juiz de direito.

Em análise da base amostral, 48.74% dos processos judiciais tiveram desembolso financeiro, enquanto 51.26% não. Nota-se, portanto, que a base está balanceada, visto que há equilíbrio entre os percentuais de ocorrência ou não de desembolso.

Sintetizamos na Tabela 18 as informações da variável dependente do estudo:

Variável Dependente	Descrição	Leiaute	Base Amostral
Houve pagamento	Informa se houve desembolso financeiro durante a vigência do processo judicial	0 - Não	51.26%
		1 - Sim	48.74%

Tabela 18 – Variável dependente

### 3.3.2 VARIÁVEIS INDEPENDENTES DE CADASTRO

Detalhamos melhor as 13 (treze) variáveis independentes de cadastro, visto que as denominações podem variar de uma IF para outra, contudo o conteúdo e objetivos são os mesmos.

Na Tabela 19 apresentamos as descrições e exemplos, importantes para correto entendimento:

Variável	Descrição	Exemplo
1. Duração do processo	Duração do processo em meses	24 meses
2. Unidade Jurídica	Região administrativa do advogado	Brasília
3. Área Judicial	Área do Direito	Trabalhista, Fiscal, Cível
4. Assunto	Assuntos tratados na ação	Hora extra
5. Quantidade de PF	Número de PF autoras da ação	1, 20, 400
6. Quantidade de PJ	Número de PJ autoras da ação	1, 10, 15
7. Base	Região do Judiciário	DF, PE, RJ, RS ou SP
8. Comarca	Cidade onde corre o processo	Salvador, São Paulo
9. Vara	Indica a vara do foro	1a Vara
10. Foro	Tribunal onde corre o processo	Federal, Estadual
11. Valor da causa atualizado	Potencial ganho financeiro do autor	R\$ 10.000,00
12. Tipo de ação	Tipo de ação na justiça	Cobrança, Civil pública
13. Macro assunto	Detalhamento do assunto da ação	Empréstimo, Seguro

Fonte: Base de dados da Instituição Financeira estudada

Tabela 19 – Variáveis independentes de cadastro

No [Apêndice A](#) detalhamos o leiaute da base de dados das variáveis independentes de cadastro, cujos dados qualitativos foram transformados em número para melhor aplicação dos modelos.

### 3.3.3 VARIÁVEIS INDEPENDENTES DE ANDAMENTO PROCESSUAL

As 11 (onze) variáveis independentes de andamento processual são binárias, conforme detalhado na Tabela 20:

Variável	Leiaute
15. Acórdão desfavorável	0 - Não 1 - Sim
16. Acórdão favorável	0 - Não 1 - Sim
17. Acórdão parcialmente favorável	0 - Não 1 - Sim
18. Acórdão retorno para julgamento do mérito	0 - Não 1 - Sim
19. Sentença desfavorável	0 - Não 1 - Sim
20. Sentença favorável	0 - Não 1 - Sim
21. Sentença parcialmente favorável	0 - Não 1 - Sim
22. Sentença incompetência de foro	0 - Não 1 - Sim
23. Sentença homologatória na execução	0 - Não 1 - Sim
24. Conversão em execução	0 - Não 1 - Sim
25. Acórdão para correção de erro material	0 - Não 1 - Sim

Fonte: Base de dados da Instituição Financeira estudada

Tabela 20 – Variáveis independentes de andamento processual

## 3.4 CENÁRIOS DO ESTUDO

Para este estudo consideramos 2 (dois) cenários diferentes com o objetivo de:

1) Avaliar a suficiência para a modelagem utilizando-se exclusivamente as variáveis independentes de cadastro; e

2) Avaliar o ganho de performance para a modelagem em agregar às variáveis independentes de cadastro as variáveis independentes de andamento processual.

A Tabela 21 detalha os 2 (dois) cenários de estudo:

---

<b>Cenário</b>	<b>Objetivo</b>	<b>Var.independentes</b>	<b>Quant. Var.independentes</b>
1	Avaliar a suficiência para a modelagem utilizando-se exclusivamente as variáveis independentes de cadastro.	Cadastro	13
2	Avaliar o ganho de performance para a modelagem em agregar às variáveis independentes de cadastro as variáveis independentes de andamento processual.	Cadastro andamento	24

---

Tabela 21 – Cenários para os algoritmos



## 4 METODOLOGIA DE PESQUISA

A proposta deste trabalho é identificar modelos preditivos de ML capazes de estimar a probabilidade de ocorrência de perda financeiras em processos judiciais em que a IF analisada figura como parte ré. O objetivo é que essa modelagem sirva de parâmetro para a adequada estimativa da probabilidade em provável, possível e remota, conforme regramento da IAS 37 (IASB, 2001) e do CPC 25 (CONTÁBEIS, 2009) para provisão e passivos contingentes.

As técnicas estatísticas preditivas utilizadas envolvem 6 (seis) modelos classificação: Regressão Logística, *Random Forest*, *Bagging*, *Decision Tree*, *Ada Boost* e *Gradient Boosting*.

A avaliação e validação dos modelos tiveram a intenção de comparar a adequação, robustez e eficácia da previsão de pagamento em processos judiciais da amostra de dados do estudo. Utilizamos para isso as técnicas de matriz de confusão, acurácia, *recall*, especificidade, precisão, *f1-score*, ROC e AUC.

### 4.1 CIÊNCIA DE DADOS E MACHINE LEARNING

A ciência de dados (*data science*) serve para transformar problemas empresariais em problemas de dados: coletar, entender e formatar dados (GRUS, 2021). A captação e tratamento de dados é de extrema relevância para o *Machine Learning*.

Segundo Pinheiro (PINHEIRO et al., 2018), *Machine Learning* ou Aprendizagem de Máquina é a área da ciência da computação composta por uma combinação de tecnologias que permitem que computadores tomem decisões com a ajuda de algoritmos que reconhecem padrões e se tornam capazes de fazer previsões. Além de deixar as escolhas mais inteligentes, o *Machine Learning* faz previsões sobre eventos futuros desconhecidos com a identificação de probabilidades.

Recentemente houve um crescimento expressivo em ciência de dados. Os avanços podem ser explicados por vários fatores, dentre eles destaca (CARVALHO, 2021): 1) **Extração de dados**: desenvolvimento de novos sensores, incluindo câmeras, cada vez mais rápidos, baratos e sofisticados; 2) **Armazenamento de dados**: aumento da capacidade de armazenamento, confiabilidade e redução do custo; e 3) **Transmissão de dados**: queda do custo, aumento da capacidade dos meios físicos utilizados e revolução trazida pelas redes de computadores e pela internet, que conecta mais dispositivos, ou coisas, do que pessoas.

Para abordar aprendizado de máquina é necessário falar sobre *modelos*. Um modelo é uma especificação matemática ou probabilística entre diferentes variáveis. Segundo Grus (GRUS, 2021), *Machine Learning* refere-se à criação e ao uso de modelos que aprendem com dados, o que também pode ser denominado de *modelagem preditiva ou mineração de dados*.

#### 4.1.1 ML E INSTITUIÇÕES FINANCEIRAS

A Conferência Internacional sobre Serviços de Computação (SANZ; ZHU, 2021) destacou que a inovação em Inteligência Artificial (IA) produz cada vez mais uma riqueza de técnicas, especialmente da forma indutiva, como *machine learning (ML)*. A maioria dos algoritmos de ML é neutra em relação ao setor e independente dos processos de negócio. A inovação é impulsionada por pesquisas disponíveis publicamente, colhidas em código aberto, com ampla distribuição por fornecedor de software e nuvem. O trabalho contínuo de IA cria uma imensa fonte de ativos para modelagem orientada a dados, entregues

como bibliotecas de software.

Nessa Conferência (SANZ; ZHU, 2021) debateram que a aplicação desses ativos de dados em finanças não acontece com sucesso ou velocidade comparáveis. Os dados são limitados em termos de acesso e uso na maior parte dos países por leis ou regras de governança, risco e *compliance* (GRC). As implicações legais somam-se ao fato de que as técnicas de IA dever ser explicadas em linguagem leiga para tomadores de decisão e reguladores antes da implantação em campo, que pode não ser satisfatória. Uma grande porcentagem de projetos falha, parte devido à modelagem inadequada de ML para problemas de análise e previsão em finanças.

Por fim, neste ponto concluímos que a variedade e complexidade do comportamento humano presente na maioria dos processos financeiros exige a compreensão de IA em um nível de profundidade cognitiva sem comparação com outros setores e a automação de processos em instituições financeiras usando "caixas pretas" de ML está sob o escrutínio da ética, leis, consumidores, reguladores (locais, estaduais, nacionais e internacionais).

#### 4.1.2 ML E O JUDICIÁRIO

A previsão do julgamento legal é um tópico de longa duração na teoria e na prática do direito para melhorar a consistência judicial, o acesso à justiça e a eficiência administrativa. Vários métodos e técnicas surgiram ao longo do tempo, desde modelos simples de cálculo até algoritmos analíticos altamente avançados para prever julgamentos jurídicos. Trabalhos recentes abordam a utilização de ML no Judiciário, conforme abordado a seguir.

Marzieh Karimi aborda em artigo recente (KARIMI-HAGHIGHI; CASTILLO, 2021) a eficácia da utilização algorítmica na justiça criminal. Em seu estudo avalia os riscos de reincidência violenta em presos sentenciados a partir de uma base de dados. Desde a década de 1920, as ferramentas de avaliação de risco de violência têm sido progressivamente usadas na justiça criminal por agentes de liberdade condicional, policiais e psicólogos para avaliar o risco de ofensas sexuais, criminosas e violentas em mais de 44 (quarenta e quatro) países. A autora destaca em seu estudo que a introdução de algoritmos para avaliação de risco na justiça criminal é um tema controverso e motiva pesquisas sobre justiça algorítmica.

Minjung Park (PARK; CHAI, 2021) apresenta um estudo sobre desenvolvimento de um modelo preditivo automatizado para decisões jurídicas relacionados a casos de violações de privacidade causadas por empresas, prevendo o aumento desse tipo de matéria no judiciário e a necessidade de redução dos esforços dos profissionais do direito e das empresas. Em seu artigo verificou um sistema de automação legal chamado *AI lawyer*, inventado para prever veredictos nos Estados Unidos em maio de 2016 que lê um grande número de documentos de julgamento e analisa o conteúdo com base em um algoritmo especial para desenhar decisões para julgar um caso automaticamente.

Esse artigo relata que torna-se mais difícil e demorado para os advogados humanos sentenciarem veredictos corretos em julgamentos legais, porque houve um aumento significativo no número de ações judiciais no passado recente. Assim, há uma necessidade de desenvolver sistemas de previsão de julgamentos judiciais precisamente baseados em um grande número de precedentes legais, com o surgimento desses advogados de IA e sistemas jurídicos automáticos baseados em *big data*. No entanto, ainda há pouca pesquisa sobre *big data analytics* no campo do direito.

#### 4.1.3 TAREFAS DE APRENDIZADO

Algoritmos de ML são utilizados em tarefas que podem ser divididas em **preditivas** e **descritivas**. Em tarefas preditivas, algoritmos são aplicados a um conjunto de dados de treinamento rotulados para



induzir um modelo capaz de prever, para um novo objeto apresentado, o valor de seu atributo alvo. Nas tarefas preditivas os algoritmos seguem o aprendizado supervisionado. Em tarefas de descrição, ao invés de prever um valor, os algoritmos extraem padrões de um conjunto de dados e seguem o aprendizado não supervisionado (FACELII et al., 2021).

Neste trabalho é utilizada a tarefa supervisionada e, diante dos dados e valor do rótulo a ser predito, a classificação. Na figura 8 ilustramos a hierarquia das categorias de aprendizado e as tarefas associadas (FACELII et al., 2021):

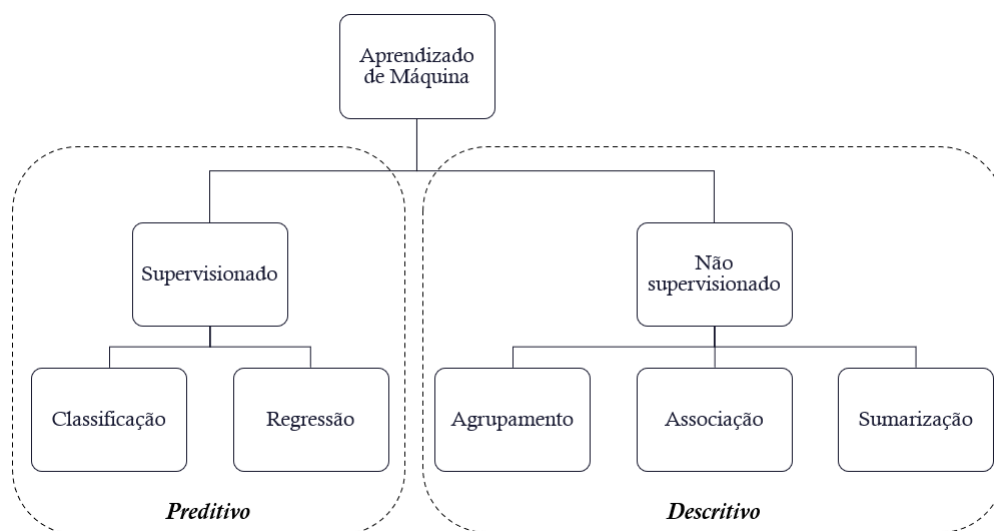


Figura 8 – Categorias de aprendizado e tarefas associadas

#### 4.1.4 PYTHON

Para aplicação dos modelos estatísticos, a linguagem de programação escolhida foi *Python*, código aberto, versão 3.9.5.

*Python* é uma linguagem *open-source* utilizada em *data science*, *machine learning*, desenvolvimento de *web*, desenvolvimento de aplicativos, automação de *scripts*, dentre outros. É orientada a objetos, uma forma específica de organizar softwares em que os procedimentos estão submetidos a classes, o que possibilita mais controle e estabilidade de códigos para projetos de grandes proporções (DEITEL; DEITEL LIPERI, 2002).

Foi desenvolvida pelo matemático Guido Van Rossum na década de 90 com o objetivo de otimizar a leitura de códigos e estimular a produtividade de quem o utiliza.

Segundo Chun (CHUN, 2000), *Python* é uma descomplicada e robusta linguagem de programação com os seguintes atributos: escalável, fácil de aprender, manter e ler, gerencia bem a memória e efetivo.

##### 4.1.4.1 BIBLIOTECAS PRINCIPAIS

Neste estudo utilizamos 6 (seis) modelos preditivos e suas respectivas bibliotecas do Python, conforme descrito na Tabela 22:

Modelos	Bibliotecas
Regressão logística	sklearn.linear_model (LogisticRegression)
<i>Decision Tree</i>	sklearn.tree (DecisionTreeClassifier)
<i>Random Forest</i>	sklearn.ensemble (RandomForestClassifier)
<i>Bagging</i>	sklearn.ensemble (BaggingClassifier)
<i>Ada Boost</i>	sklearn.ensemble (AdaBoostClassifier)
<i>Gradient</i>	sklearn.ensemble (GradientBoostingClassifier)

Tabela 22 – Biblioteca de modelos do *Python*

#### 4.1.4.2 BIBLIOTECAS ACESSÓRIAS

Outras bibliotecas foram utilizadas complementarmente, para avaliação, validação e plotagem de resultados dos modelos, Tabela 23:

Objetivos	Bibliotecas
Análise de dados	pandas
Operações matemáticas	numpy
Visualização de dados	matplotlib.pyplot, seaborn
Selecionar amostras	sklearn.model_selection (train_test_split)
Refinamento de modelos ( <i>tuning</i> )	from sklearn.model_selection import GridSearchCV
Validação e desempenho dos modelos	sklearn.metrics auc roc_curve roc_auc_score classification_report confusion_matrix accuracy_score balanced_accuracy_score fbeta_score f1_score precision_score recall_score precision_recall_curve precision_recall fscore_support mean_squared_error make_scorer

Tabela 23 – Bibliotecas acessórias do *Python*

## 4.2 TÉCNICAS ESTATÍSTICAS

### 4.2.1 CLASSIFICAÇÃO

As variáveis deste estudo são eminentemente categóricas, ou seja, as respostas tem características qualitativas, como por exemplo a “Base Região do Judiciário” dos processos judiciais, que podem ser do DF – Distrito Federal, de PE – Pernambuco, do RJ – Rio de Janeiro, do RS – Rio Grande do Sul ou de SP – São Paulo.

Dentre as abordagens estatísticas para prever respostas qualitativas a uma observação temos a classificação, que atribui uma observação a uma categoria ou classe (JAMES et al., 2013).

### 4.2.2 REGRESSÃO LOGÍSTICA

A Regressão Logística caracteriza-se por avaliar a relação entre a variável dependente e as variáveis independentes e mostra-se adequada quando a variável resposta é categórica binária, 0 e 1. Descreve matematicamente o comportamento de  $Y$  em função dos valores de  $X_i$  a partir do método da máxima verossimilhança (JR; LEMESHOW; STURDIVANT, 2013).

A probabilidade de ocorrência de um evento pode ser estimada diretamente. No caso da variável dependente  $Y$  assumir apenas dois possíveis estados, 0 e 1, e houver um conjunto de variáveis independentes  $X_1, X_2, \dots, X_i$ , o modelo pode ser escrito conforme equação 4.1 (KLEINBAUM et al., 2002):

$$P(Y) = \frac{1}{1 + e^{-z}} \quad (4.1)$$

Onde  $z = \alpha + \sum_i^n b_i X_i$

### 4.2.3 MÉTODOS BASEADOS EM ÁRVORES

Os métodos baseados em árvores são simples e úteis para interpretação. Neste trabalho utilizamos abordagens que envolvem a produção de várias árvores que são combinadas para culminarem num consenso, como *Decision Tree*, *Random Forest*, *Bagging* e *Boosting*. A combinação de muitas árvores resulta em melhorias substanciais na precisão da previsão (JAMES et al., 2013).

#### 4.2.3.1 DECISION TREE

As *Decision Tree* são utilizadas frente às árvores de regressão quando a variável resposta é categórica, binária, 0 e 1. O modelo de árvore de decisão adota uma sequência de decisões lógicas, como num fluxo ou passo a passo, com “nós” de decisão que sinalizam a necessidade de tomada de decisão em dado momento ou atributo.

As decisões são tomadas como se fossem regras de “se/então”, que tomam como ponto de partida uma determinada situação inserida num pacote de opções e retornam com uma decisão (MAIMON; ROKACH, 2014), conforme Figura 9:

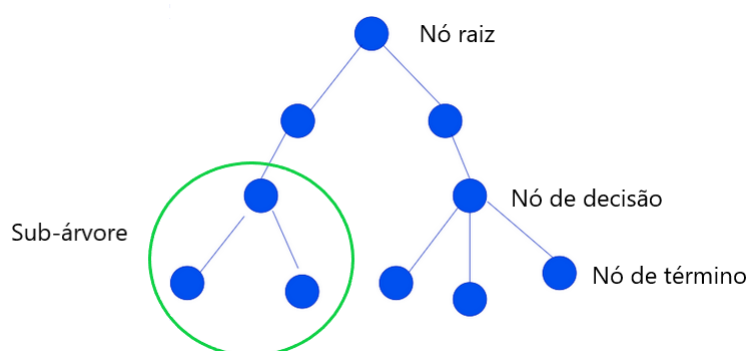


Figura 9 – Exemplo de árvore de decisão

Na DT, a probabilidade de uma amostra pertencer a uma classe  $C_j$  é estimada pela Equação 4.2, em que  $|Z|$  é o número de amostras do conjunto  $Z$  (BREIMAN et al., 2017):

$$p_i = \frac{freq(C_j, Z)}{|Z|} \quad (4.2)$$

Para se fazer a divisão dos "nós" da árvore de classificação, utiliza-se o algoritmo *Gini*, que diz que ao selecionarmos aleatoriamente duas amostras da mesma população, ambas devem ter a mesma classe e a probabilidade é igual a 1 se a população for pura. Funciona apenas com variáveis categóricas, "sucesso" e "fracasso", e, quanto maior o Gini, maior a homogeneidade. A equação 4.3 detalha a fórmula:

$$gini = 1 - \sum_{i=1}^C (p_i)^2 \quad (4.3)$$

A entropia é um conceito utilizado para a classificação do DT. A entropia é o grau de pureza do conjunto e define o número de *bits* necessários, em média, para representar a informação em falta, usando codificação ótima (TSAI; HSU; YEN, 2014). Quanto menor for a entropia, melhor.

A fórmula da entropia segue descrita na equação 4.4, dado um conjunto de dados Z, com instâncias pertencentes à classe i, com probabilidade  $p_i$  :

$$Entropia(Z) = \sum_{i=1}^C p_i * \log_2(p_i) \quad (4.4)$$

#### 4.2.3.2 BAGGING

A técnica *Bagging* pode ser usada com muitos métodos de classificação e métodos de regressão para reduzir a variação associada a previsão e, assim, melhorar o processo de previsão (BREIMAN, 1996).

É uma ideia relativamente simples: muitas amostras são extraídas dos dados disponíveis, algum método de previsão é aplicado a cada amostra e, em seguida, os resultados são combinados por média para regressão e votação simples para classificação para obter a previsão geral, com a variação sendo reduzida devido à média (SUTTON, 2005).

O método *Bagging* é utilizado para diminuir a variância das previsões. Combina os resultados de múltiplos classificadores, modelados em diferentes sub-amostras do mesmo conjunto de dados. A Figura 10 clarifica a ideia do *Bagging* (TEAM, 2016):

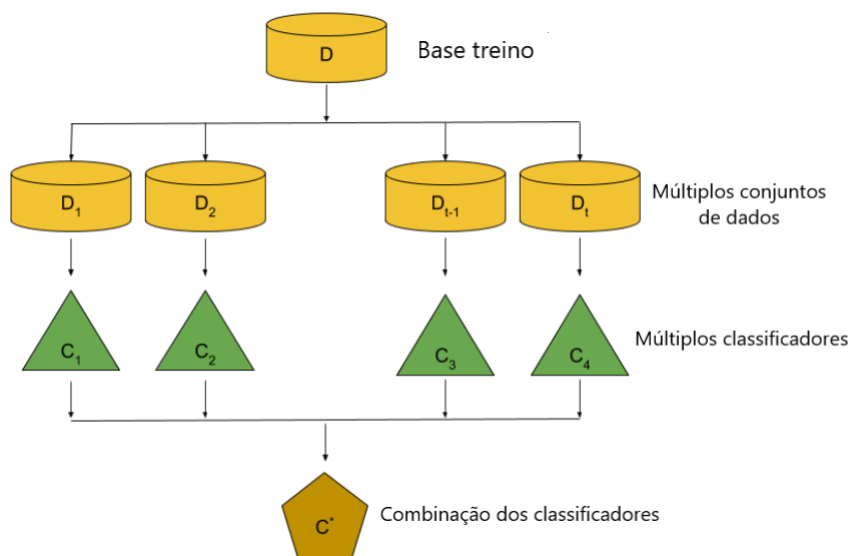


Figura 10 – Exemplo Bagging

Breiman (BREIMAN, 1996) descreve o modelo *Bagging* com as seguintes etapas:

1. Construção de uma amostra aleatória,  $t$ , selecionada do conjunto de dados;
2. Cálculo do estimador  $C_t$  no conjunto de dados do passo 1;
3. Repetição dos dois primeiros passos por  $t = 1, \dots, T$ , em que  $T$  é o total de iterações definidas pelo executor; e
4. A partir daí, cada classificador determina um voto, em que  $x$  comporta os dados de cada elemento do conjunto de treinamento, conforme Equação 4.5:

$$C(x) = T^{-1} \sum_{t=1}^T C_t(x) \quad (4.5)$$

#### 4.2.3.3 ADA BOOST

Yoav Freund e Robert Schapire desenvolveram em 1996 o algoritmo de reforço *AdaBoost*. Nas Figuras 11 e 12 apresentamos a extração de um dos artigos que traz a fórmula desse algoritmo em 2 (duas versões: AdaBoost.M1 e AdaBoost.M2). As duas versões são equivalentes para problemas de classificação binária e diferem apenas em sua manipulação (FREUND; SCHAPIRE et al., 1996):

##### Algorithm AdaBoost.M1

**Input:** sequence of  $m$  examples  $\langle (x_1, y_1), \dots, (x_m, y_m) \rangle$  with labels  $y_i \in Y = \{1, \dots, k\}$   
 weak learning algorithm **WeakLearn**  
 integer  $T$  specifying number of iterations

**Initialize**  $D_1(i) = 1/m$  for all  $i$ .

**Do for**  $t = 1, 2, \dots, T$

1. Call **WeakLearn**, providing it with the distribution  $D_t$ .
2. Get back a hypothesis  $h_t : X \rightarrow Y$ .
3. Calculate the error of  $h_t$ :  $\epsilon_t = \sum_{i: h_t(x_i) \neq y_i} D_t(i)$ . If  $\epsilon_t > 1/2$ , then set  $T = t - 1$  and abort loop.
4. Set  $\beta_t = \epsilon_t / (1 - \epsilon_t)$ .
5. Update distribution  $D_t$ :  $D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} \beta_t & \text{if } h_t(x_i) = y_i \\ 1 & \text{otherwise} \end{cases}$   
 where  $Z_t$  is a normalization constant (chosen so that  $D_{t+1}$  will be a distribution).

**Output** the final hypothesis:  $h_{\hat{n}}(x) = \arg \max_{y \in Y} \sum_{t: h_t(x)=y} \log \frac{1}{\beta_t}$ .

Figura 11 – Fórmula AdaBoost 1

**Algorithm AdaBoost.M2**

**Input:** sequence of  $m$  examples  $\langle (x_1, y_1), \dots, (x_m, y_m) \rangle$  with labels  $y_i \in Y = \{1, \dots, k\}$   
 weak learning algorithm **WeakLearn**  
 integer  $T$  specifying number of iterations

Let  $B = \{(i, y) : i \in \{1, \dots, m\}, y \neq y_i\}$

**Initialize**  $D_1(i, y) = 1/|B|$  for  $(i, y) \in B$ .

**Do for**  $t = 1, 2, \dots, T$

1. Call **WeakLearn**, providing it with mislabel distribution  $D_t$ .
2. Get back a hypothesis  $h_t : X \times Y \rightarrow [0, 1]$ .
3. Calculate the pseudo-loss of  $h_t$ :  $\epsilon_t = \frac{1}{2} \sum_{(i,y) \in B} D_t(i, y) (1 - h_t(x_i, y_i) + h_t(x_i, y))$ .
4. Set  $\beta_t = \epsilon_t / (1 - \epsilon_t)$ .
5. Update  $D_t$ :  $D_{t+1}(i, y) = \frac{D_t(i, y)}{Z_t} \cdot \beta_t^{(1/2)(1+h_t(x_i, y_i)-h_t(x_i, y))}$   
 where  $Z_t$  is a normalization constant (chosen so that  $D_{t+1}$  will be a distribution).

**Output** the hypothesis:  $h_{fm}(x) = \arg \max_{y \in Y} \sum_{t=1}^T \left( \log \frac{1}{\beta_t} \right) h_t(x, y)$ .

Figura 12 – Fórmula AdaBoost 2

#### 4.2.3.4 GRADIENT BOOSTING

O termo *Boosting* refere-se a uma família de algoritmos que converte uma aprendizagem fraca em uma aprendizagem forte. O *Boosting*, assim como o *Bagging*, é uma abordagem baseada em comitês que pode ser usada para melhorar a precisão dos métodos de classificação ou regressão. Ao contrário do *Bagging*, que usa uma média simples dos resultados para obter uma previsão geral, o reforço usa uma média ponderada dos resultados obtidos da aplicação de um método de previsão a várias amostras (SUTTON, 2005).

Além disso, com o reforço, as amostras usadas em cada etapa não são todas extraídas da mesma maneira da mesma população, mas os casos previstos incorretamente de uma determinada etapa recebem um peso aumentado durante a próxima etapa. Assim, o *boosting* é um procedimento iterativo, incorporando pesos, em vez de se basear em uma simples média de previsões, como é o caso do *Bagging*.

#### 4.2.3.5 RANDOM FOREST

*Random Forest* é uma combinação de preditores de árvores de modo que cada árvore depende dos valores de um vetor aleatório de uma amostra com a mesma distribuição para todas as árvores da floresta. É um classificador que consiste em uma coleção de classificadores conforme equação 4.6 (BREIMAN, 2001):

$$RF = h(x, \Theta_k), k = 1 \quad (4.6)$$

Onde os  $\Theta_k$  :

São distribuídos identicamente, em vetores independentes e aleatórios e cada árvore lança um voto unitário para a classe mais popular na entrada  $X$ .

O *Random Forest* faz parte dos métodos *Bagging*, que tem a característica de combinar diferentes modelos para se obter um único resultado, de forma mais potente. Por isso o RF é mais robusto e complexo, conforme demonstramos na Figura 13:

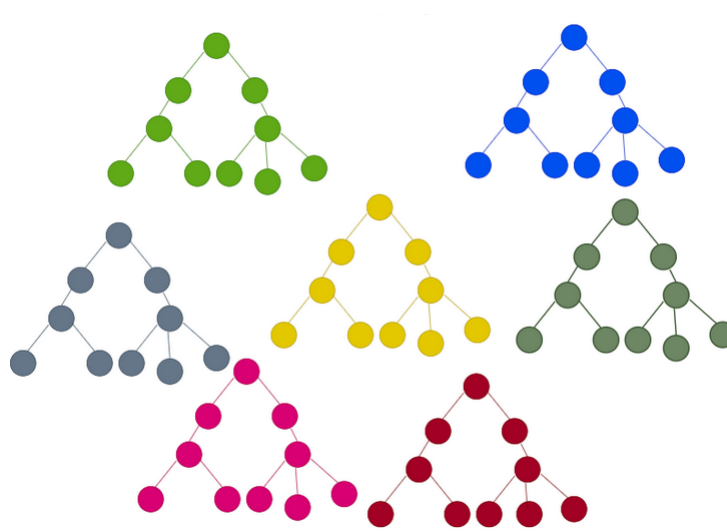


Figura 13 – Exemplo Random Forest

### 4.3 AVALIAÇÃO E VALIDAÇÃO DE MODELOS

Para a avaliação dos modelos utilizamos como referencial as medidas-padrão para validação de modelos de classificação de risco de crédito, como por exemplo a acurácia, sensibilidade, especificidade, precisão e f1-score, que podem ser explicadas a partir de uma matriz de confusão (WANG et al., 2011).

#### 4.3.1 MATRIZ DE CONFUSÃO

A matriz de confusão é uma tabela 2 x 2 utilizada em ML para identificar os quatro tipos de classificação de modelo binário (0 ou 1), conforme Figura 14:

		Valor previsto	
		Positivo	Negativo
Valor real	Positivo	VP (Verdadeiro Positivo)	FN (Falso Negativo)
	Negativo	FP (Falso Positivo)	VN (Verdadeiro Negativo)

Figura 14 – Matriz de confusão

A matriz de erro, como também é chamada a matriz de confusão, permite calcular valores como acurácia, sensibilidade, precisão, especificidade e f1 score (SOKOLOVA; LAPALME, 2009).

##### 4.3.1.1 ACURÁCIA

A acurácia reflete em percentual a quantidade de acertos em relação à classificação total, conforme equação 4.7:

$$Acurácia = \frac{VP + VN}{VP + FP + FN + VN} \quad (4.7)$$

#### 4.3.1.2 SENSIBILIDADE / RECALL

A sensibilidade ou recall é a relação de VP - Verdadeiros Positivos frente à quantidade total de positivos. Pode ser entendida como "Taxa de Positivos Verdadeiros" e mede o quão assertivo o classificador é para classificar a classe de interesse, conforme equação 4.8

$$\text{Sensibilidade} = \frac{VP}{VP + FN} \quad (4.8)$$

#### 4.3.1.3 ESPECIFICIDADE

A especificidade é a parte de VN - Verdadeiros Negativos em relação ao total de negativos. Pode ser nominada também de "Taxa de Positivos Falsos" e mede o quão assertivo o classificador é para classificar corretamente a outras classe, conforme equação 4.9:

$$\text{Especificidade} = \frac{VN}{VN + FP} \quad (4.9)$$

#### 4.3.1.4 PRECISÃO

A precisão é o valor predito positivo de forma correta, VP - Verdadeiro Positivo, em relação ao total de classificados como positivos, conforme equação 4.10:

$$\text{Precisão} = \frac{VP}{VP + FP} \quad (4.10)$$

#### 4.3.1.5 F1 SCORE

O escore f1 é definido como a média harmônica da sensibilidade (recall) e precisão, conforme equação 4.11:

$$f1score = \frac{2 * sensibilidade * precisão}{sensibilidade + precisão} \quad (4.11)$$

Como a pontuação f1 é uma média de sensibilidade e precisão, isso significa que a pontuação f1 dá peso igual à ambos e:

- Um modelo obterá uma pontuação alta na f1 se a sensibilidade e precisão forem altos;
- Um modelo obterá uma pontuação f1 baixa se sensibilidade e precisão forem baixos;
- Um modelo obterá uma pontuação f1 média se um for baixo e o outro for alto.

### 4.3.2 ROC - RECEIVER OPERATING CHARACTERISTIC

A *Receiver Operating Characteristic Curve*, ROC, é um gráfico que expõe de forma visual o desempenho de um modelo de classificação (FAWCETT, 2006). Utiliza-se de dois eixos, a sensibilidade e especificidade, com o objetivo de ilustrar a relação entre a "Taxa de Positivos Verdadeiros" e a "Taxa de Positivos Falsos".

A seguir, na Figura 15, uma representação gráfica da curva ROC. Vários pontos no espaço ROC são importantes a serem observados. O ponto inferior esquerdo (0, 0) representa a estratégia de nunca emitir uma classificação positiva, tal classificador não comete erros falsos positivos, mas também não obtém nenhum positivo verdadeiro. A estratégia oposta, de emitir classificações positivas



incondicionalmente, é representada pelo ponto superior direito (1, 1) e o ponto "D"(0, 1) representa a classificação perfeita(FAWCETT, 2006):

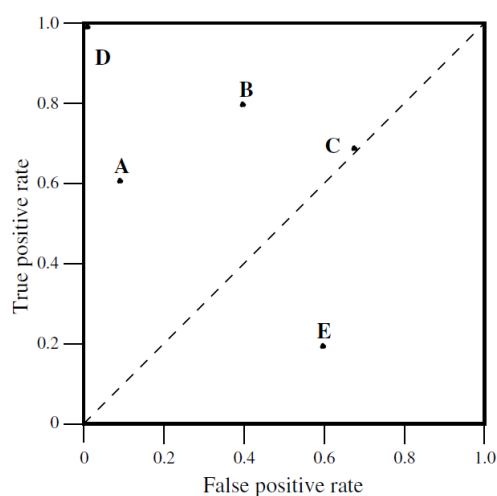


Figura 15 – Interpretação dos resultados da curva ROC

Os gráficos ROC têm sido usados há muito tempo na teoria de detecção de sinal para representar a compensação entre as taxas de acerto e as taxas de alarme falso dos classificadores (SWETS, 2014).

Nos últimos anos houve um aumento no uso de gráficos ROC na comunidade de aprendizado de máquina devido, em parte, à percepção de que a precisão de classificação simples geralmente é uma métrica ruim para medir o desempenho (PROVOST; FAWCETT, 1998).

Concluimos este tópico com a Figura 16, que sinaliza em cores e formas as curvas ROC conforme serão demonstradas neste estudo:

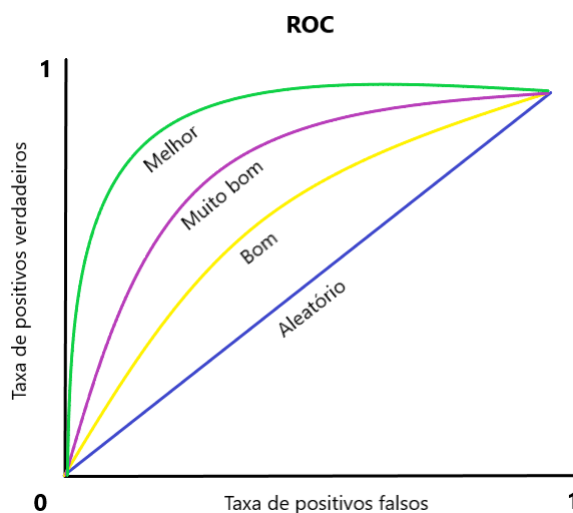


Figura 16 – ROC - Receiver Operating Characteristic

#### 4.3.2.1 AUC - AREA UNDER THE CURVE

Para comparar classificadores, podemos querer reduzir o desempenho do ROC a um único valor escalar que representa o desempenho esperado. Um método comum é calcular a área sob a curva ROC,

abreviada AUC (BRADLEY, 1997). Pode-se entender que a AUC é uma forma sintetizar a curva ROC num único valor, juntando todos os limites e calculando a "área sob a curva".

O valor da AUC varia entre 0 e 1, com limiar entre as classes de 0,5, e quanto maior o valor, melhor e mais preciso o modelo, conforme Figura 17:

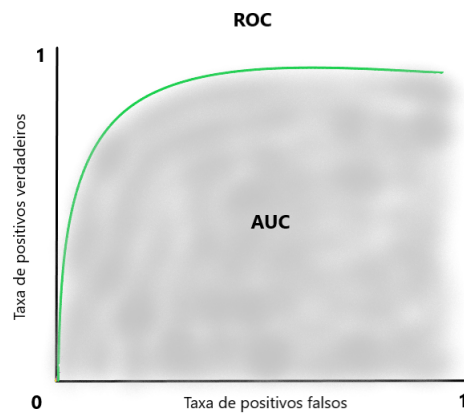


Figura 17 – AUC - Area Under the Curve

Conforme citado anteriormente, como a AUC é uma porção da área do quadrado unitário, seu valor estará sempre entre 0 e 1. No entanto, como a adivinhação aleatória produz a linha diagonal entre (0, 0) e (1, 1), que tem uma área de 0,5, nenhum classificador realista deve ter uma AUC menor que 0,5. A AUC também está intimamente relacionada com o coeficiente de Gini (BREIMAN et al., 2017), que é o dobro da área entre a diagonal e a curva ROC:  $Gini + 1 = 2 \cdot AUC$  (FAWCETT, 2006).

## 5 RESULTADOS E ANÁLISES

Para avaliar a performance dos modelos diante das características das variáveis independentes, utilizamos os 6 (seis) algoritmos selecionados nos cenários de estudo 1 e 2, conforme códigos descritos no Apêndice B.

A disponibilidade de informações do cenário 1, apenas com variáveis independentes de cadastro, é maior e de mais fácil acesso ao cientista de dados em qualquer instituição financeira, motivo pelo qual seu estudo é relevante para se analisar a suficiência quanto à matriz de confusão, acurácia, sensibilidade, especificidade, f1-score, AUC e ROC. Para o cenário 2, é necessário conhecimento técnico especializado do campo judicial e rigoroso controle dos registros das fases processuais na base de dados. Contudo, é necessário verificar se há ganhos significativos na modelagem ao agregarmos outras variáveis.

### 5.1 CENÁRIO 1 - RESULTADOS DAS VALIDAÇÕES - BASE TESTE

Os modelos foram avaliados considerando-se a separação da base em 80% para treino e 20% para teste, no total de 318.566 e 79.641 registros respectivamente. Os instrumentos de validação foram aplicados nas bases de teste dos cenários 1 e 2 e os resultados da performance dos modelos serão apresentados a seguir.

#### 5.1.1 CENÁRIO 1 - MATRIZ DE CONFUSÃO

Apresentamos as 6 (seis) matrizes de confusão do cenário 1, aplicadas na base de teste, com destaque às performances dos modelos *Random Forest*, que melhor estimou os Verdadeiros Positivos e os Falsos Negativos, e *Gradient Boosting*, que melhor estimou os Verdadeiros Negativos e Falsos Positivos.

##### *Logistic Regression*

		Valor previsto	
		Houve pagamento	Não houve pagamento
Valor real	Houve pagamento	27950	12978
	Não houve pagamento	10362	28352

##### *Random Forest*

		Valor previsto	
		Houve pagamento	Não houve pagamento
Valor real	Houve pagamento	33107	7821
	Não houve pagamento	7505	31209

##### *Bagging*

		Valor previsto	
		Houve pagamento	Não houve pagamento
Valor real	Houve pagamento	32331	8597
	Não houve pagamento	10563	28151

##### *Decision Tree*

		Valor previsto	
		Houve pagamento	Não houve pagamento
Valor real	Houve pagamento	30698	10230
	Não houve pagamento	7460	31254

##### *Ada Boost*

		Valor previsto	
		Houve pagamento	Não houve pagamento
Valor real	Houve pagamento	32430	8498
	Não houve pagamento	7693	31021

##### *Gradient Boosting*

		Valor previsto	
		Houve pagamento	Não houve pagamento
Valor real	Houve pagamento	32655	8273
	Não houve pagamento	7094	31620

Figura 18 – Matriz de confusão do cenário 1 - Base teste

Concluimos que apenas com a avaliação da matriz de confusão ainda não é possível aferir qual o melhor modelo, sendo necessário avançar nas análises.

### 5.1.2 CENÁRIO 1 - ACURÁCIA, SENSIBILIDADE, ESPECIFICIDADE, PRECISÃO, F1-SCORE E AUC

A partir dos dados apresentados na matriz de confusão do cenário 1, da base de teste, é possível traçar um comparativo entre a acurácia, sensibilidade, especificidade, precisão, f1-score e AUC. A tabela 24 demonstra esse comparativo:

Modelo	Acurácia	Sensibilidade	Especificidade	Precisão	f1 score	ROC_AUC
LR	0.70694%	0.73234%	0.68291%	0.68599%	0.70841%	0.70763%
RF	0.80756%	0.80614%	0.80891%	0.79962%	0.80287%	0.80753%
BG	0.75942%	0.72715%	0.78995%	0.76606%	0.74610%	0.75855%
DT	0.77788%	0.80730%	0.75005%	0.75340%	0.77942%	0.77868%
AB	0.79670%	0.80129%	0.79237%	0.78496%	0.79304%	0.79683%
GB	0.80705%	0.81676%	0.79786%	0.79262%	0.80451%	0.80731%

Tabela 24 – Resultados das validações do cenário 1 - Base teste

Observamos que o modelo *Random Forest* apresentou o melhor resultado como classificador, com melhores acurácia (0.80756%), especificidade (0.80891%), precisão (0.79962%) e AUC (0.80753%), seguido do *Gradient Boosting*, que apresentou os melhores resultados para sensibilidade (0.81676%) e f1-score (0.80451%).

Em contrapartida, o modelo de Regressão Logística teve o pior desempenho e deixou a desejar frente aos demais, com taxas chegando a 10 p.p a menor frente ao melhor resultado, como a especificidade (0.68291%) e precisão (0.68599%).

### 5.1.3 CENÁRIO 1 - CURVA ROC

Analisamos as curvas ROC do cenário 1 e notamos que todos os modelos ultrapassaram o limiar de 0.5, faixa pontilhada em azul, com tendência ao canto superior esquerdo, o que significa boa relação entre a Taxa de Positivos Verdadeiros e Taxa de Positivos Falsos, conforme Figura 19:

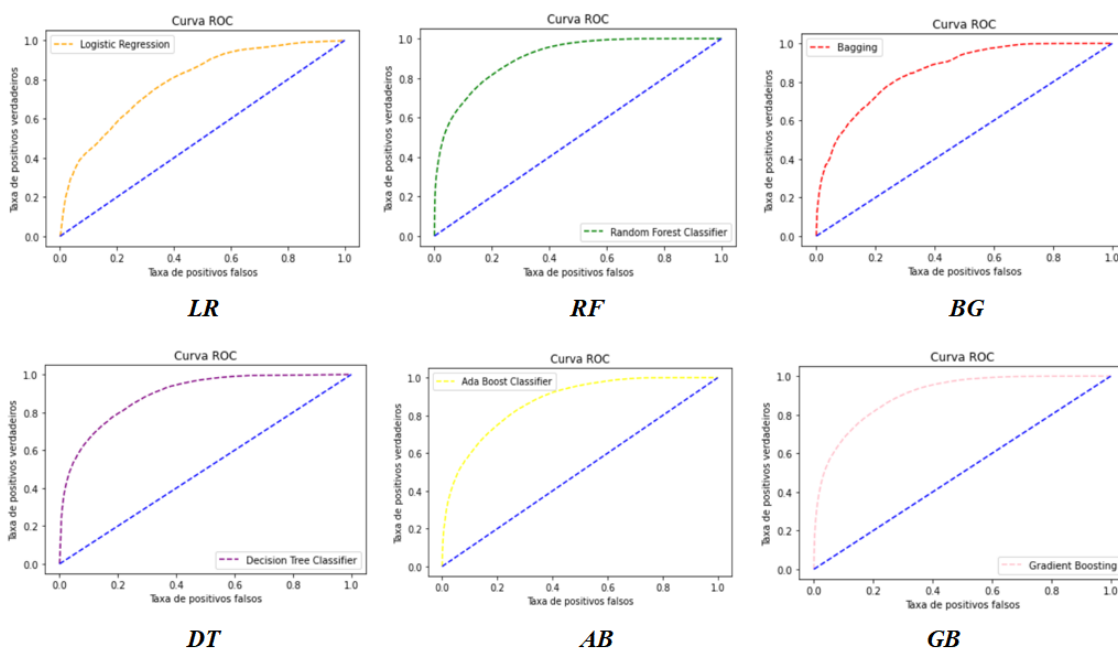


Figura 19 – Curva ROC individual do cenário 1 - Base teste

A consolidação das Curvas ROC, Figura 20, permite identificar que os modelos *Random Forest* e *Gradient Boosting* tiveram desempenho quase idêntico, com as cores das linhas se sobrepondo uma a outra, sendo possível o desempate apenas ao checarmos as notas de AUC de 0.80753% e 0.80731%, respectivamente, em que o RF apresenta leve superioridade. O pior desempenho da curva ROC é da Regressão Logística, linha laranja:

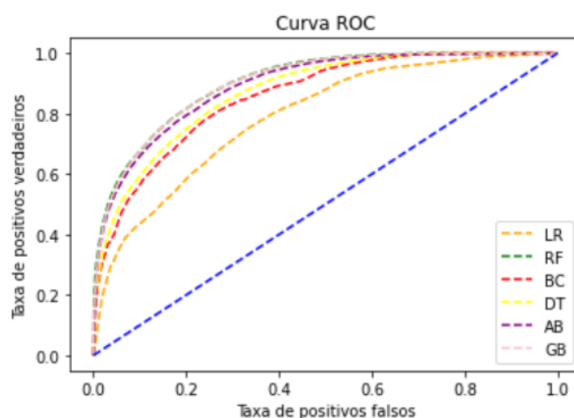


Figura 20 – Curva ROC consolidada do cenário 1 - Base teste

## 5.2 CENÁRIO 2 - RESULTADOS DAS VALIDAÇÕES

### 5.2.1 CENÁRIO 2 - MATRIZ DE CONFUSÃO

No cenário 2, base de teste, as 6 (seis) matrizes de confusão foram extraídas e 3 (três) modelos tiveram os melhores desempenhos: *Random Forest*, *Gradient Boosting* e *Ada Boost*. O RF melhor estimou os Verdadeiros Negativos e Falsos Positivos e o GB melhor estimou os Verdadeiros Positivos e os Falsos Negativos, conforme Figura 21:

#### *Logistic Regression*

		Valor previsto	
		Houve pagamento	Não houve pagamento
Valor real	Houve pagamento	32555	8373
	Não houve pagamento	5653	33061

#### *Random Forest*

		Valor previsto	
		Houve pagamento	Não houve pagamento
Valor real	Houve pagamento	35093	5835
	Não houve pagamento	3074	35640

#### *Bagging*

		Valor previsto	
		Houve pagamento	Não houve pagamento
Valor real	Houve pagamento	33435	7493
	Não houve pagamento	4984	33730

#### *Decision Tree*

		Valor previsto	
		Houve pagamento	Não houve pagamento
Valor real	Houve pagamento	34519	6409
	Não houve pagamento	4539	34175

#### *Ada Boost*

		Valor previsto	
		Houve pagamento	Não houve pagamento
Valor real	Houve pagamento	35495	5433
	Não houve pagamento	3956	34758

#### *Gradient Boosting*

		Valor previsto	
		Sim	Não
Valor real	Houve pagamento	35570	5358
	Não houve pagamento	3454	35260

Figura 21 – Matriz de confusão do cenário 2 - Base teste

Assim como no cenário 1, concluímos que apenas com a avaliação da matriz de confusão ainda não é possível aferir qual o melhor modelo, sendo necessário avançar nas análises.

## 5.2.2 CENÁRIO 2 - ACURÁCIA, SENSIBILIDADE, ESPECIFICIDADE, PRECISÃO, F1-SCORE E AUC

Observamos que o modelo *Gradient Boosting* apresentou o melhor resultado como classificador sobre a base de teste, com os maiores resultados para acurácia (0.88935%), especificidade (0.86909%), precisão (0.86809%), f1-score (0.88892%) e AUC (0.88993%). Em segundo lugar o modelo *Random Forest* se sobressaiu com a maior taxa de sensibilidade (0.92060%).

A Tabela 25 sintetiza os resultados do cenário 2:

Modelo	Acurácia	Sensibilidade	Especificidade	Precisão	f1 score	ROC_AUC
LR	0.82389%	0.85398%	0.79542%	0.79792%	0.82500%	0.82470%
RF	0.88814%	0.92060%	0.85743%	0.85931%	0.88890%	0.88901%
BG	0.84334%	0.87126%	0.81692%	0.81823%	0.84391%	0.84409%
DT	0.86253%	0.88276%	0.84341%	0.84208%	0.86194%	0.86308%
AB	0.88211%	0.89781%	0.86725%	0.86482%	0.88101%	0.88253%
GB	0.88935%	0.91078%	0.86909%	0.86809%	0.88892%	0.88993%

Tabela 25 – Resultados das validações do cenário 2 - Base teste

O modelo de Regressão Logística teve o menor desempenho, porém as taxas alcançadas mostram a boa eficácia desse classificador para o objeto deste estudo.

## 5.2.3 CENÁRIO 2 - CURVA ROC

As curvas ROC do cenário 2 apresentaram resultados superiores ao limiar de 0.5, com menor AUC para o modelo de Regressão Logística (0.82470%) e maior para o *Gradient Boosting* (0.88993%), conforme Figura 22:

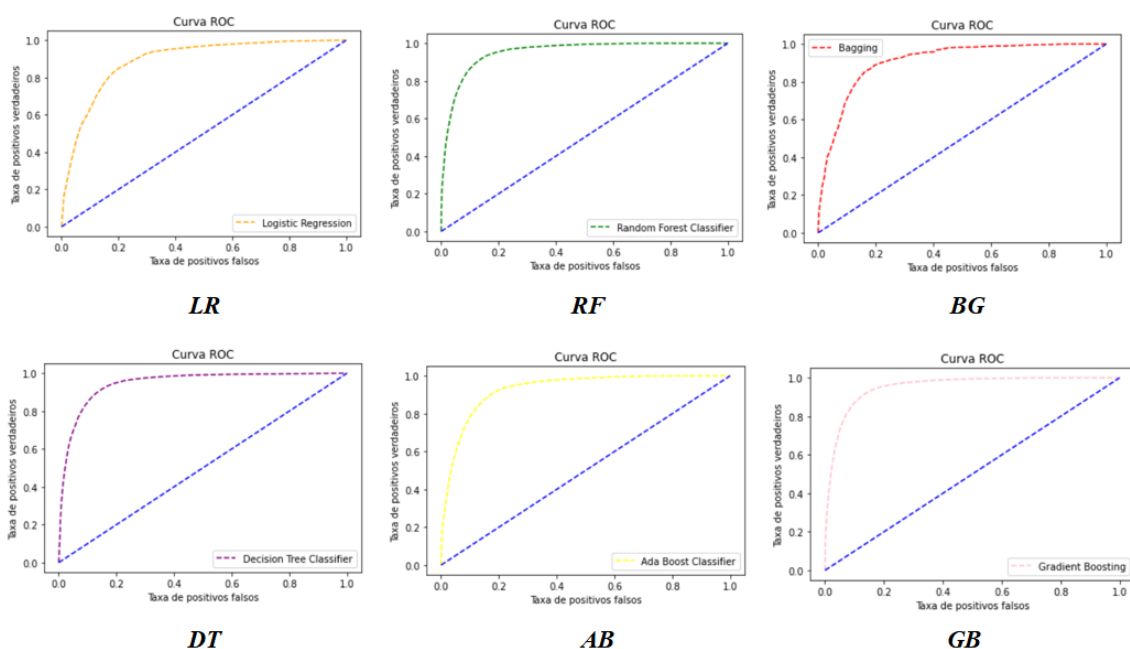


Figura 22 – Curva ROC individual do cenário 2 - Base teste

A consolidação das curvas ROC, figura 23, permite identificar que os modelos *Random Forest* e *Gradient Boosting* tem desempenhos extremamente semelhantes, com as cores das linhas praticamente unificadas, com o modelo GB levemente superior ao RF, AUC 0.88993% e 0.88901%.

Como citado anteriormente, o menor desempenho apresentado foi para o modelo de Regressão Logística, com AUC de 0.82470%, contudo é uma performance bem satisfatória.

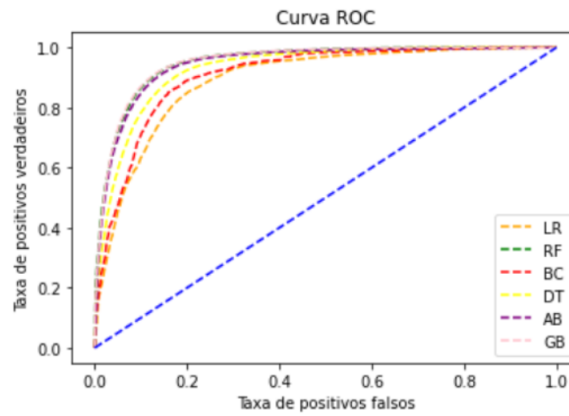


Figura 23 – Curva ROC consolidada - cenário 2 - Base teste





## 6 CONSIDERAÇÕES FINAIS

A sustentabilidade das organizações de qualquer segmento perpassa o adequado gerenciamento de riscos que engloba, dentre outros pontos, a completude e correção das demonstrações financeiras. A responsabilidade por esse gerenciamento é de todos e, conforme preconiza a ISO 31000, deve ser inclusivo e levar em consideração fatores humanos e culturais (STANDARDIZATION, 2009a).

No aspecto humano o estudo permitiu identificar, de forma objetiva, que o risco jurídico é responsabilidade de forma direta da área jurídica e da área de riscos e de forma indireta de todas as áreas das instituições, considerando que a mitigação é responsabilidade dessas unidades em suas respectivas atuações.

No aspecto cultural os *benchmarks* de risco jurídico e legal possibilitaram analisar diferenças significativas entre as instituições financeiras brasileiras e as internacionais. Os montantes provisionados e registrados em notas explicativas como passivos contingentes no Brasil são mais relevantes nas DF comparados aos 5 (cinco) maiores bancos mundiais.

Aqui é válido um destaque, as IF asiáticas *MUFJ Bank – Mitsybish UFK Financial Group* e *ICBC - Industrial and Commercial Bank of China* não tem montantes provisionados ou registros em notas explicativas para o risco jurídico (INDUSTRIAL; CHINA, 2021) (GROUP, 2021). Isso demonstra a baixa probabilidade e impacto desse tipo de risco operacional na China e Japão por motivos culturais, relacionados à baixíssima judicialização de litígios.

No Brasil o acionamento do Poder Judiciário é prática recorrente e, por vezes, questões que poderiam ter sido resolvidas no âmbito administrativo sobrecarregam a estrutura jurídica, tornando seu acesso e resolubilidade difíceis e demorados. Para os 5 (cinco) maiores bancos brasileiros a soma em 2021 verificada nas DF para provisão de riscos judiciais foi de **R\$ 80 bilhões de reais** (US\$ 17 bilhões de dólares), Tabela 8, e de **R\$ 79 bilhões de reais** (US\$ 16,8 bilhões de dólares) para passivos contingentes, Tabela 9.

Esses montantes e a quantidade de processos judiciais torna o objetivo deste estudo interessante às instituições financeiras brasileiras, em especial as respostas às duas perguntas iniciais, sustentadas em técnicas estatísticas e de *Machine Learning*.

Os algoritmos selecionados foram: *Logistic Regression*, *Decision Tree*, *Random Forest*, *Bagging*, *Ada Boost* e *Gradient Boosting*, processados na ferramenta *Python* com bibliotecas específicas para a finalidade deste estudo. A partir dos resultados e análises das validações apurados no capítulo 5, respondemos as perguntas:

1. A utilização de algoritmos preditivos de *Machine Learning* para apurar risco jurídico e risco operacional de instituições financeiras é viável? **Resposta: sim. Os resultados demonstraram elevada acurácia dos algoritmos.**

2. Em caso de resposta afirmativa da pergunta anterior, qual o melhor algoritmo preditivo de *Machine Learning* para apurar o risco jurídico? **Resposta: *Random Forest* para o cenário 1, Tabela 24, e *Gradient Boosting* para o cenário 2, Tabela 25.**

Este estudo permite afirmar que a utilização de métodos estatísticos e de *Machine Learning* é altamente eficiente para prever a ocorrência ou não de desembolso financeiro em processos judiciais de instituições financeiras brasileiras, conceituado neste trabalho como risco jurídico.

Destacamos, contudo, uma limitação relacionada ao fato de que as técnicas de IA devem ser explicadas em linguagem leiga para tomadores de decisão e reguladores antes da implantação em campo, que pode não ser satisfatória. A variedade e complexidade do comportamento humano presente na maioria dos processos financeiros exige a compreensão de IA em um nível de profundidade cognitiva sem comparação com outros setores e a automação de processos em instituições financeiras usando "caixas pretas" de ML está sob o escrutínio de leis, consumidores, reguladores (locais, estaduais, nacionais internacionais) (SANZ; ZHU, 2021).

Por fim, sugerimos para estudos futuros, avaliar a utilização de redes neurais para a predição do risco jurídico e a apuração dos valores a serem provisionados e registrados em passivos contingentes.

## REFERÊNCIAS

- AMERICA, B. of. Bank of america financial statements and notes 2021. *Bank of America*, New York, EUA, 2021. Disponível em: <<https://investor.bankofamerica.com/annual-reports-and-proxy-statements>>. Citado 2 vezes nas páginas 43 e 44.
- BRADESCO. Demonstrações contábeis do bradesco de 2021, 4º trim. *BRADESCO*, São Paulo, DF, 2021. Disponível em: <<https://www.bradesco.com.br/informacoes-ao-mercado/central-de-resultados/>>. Citado na página 40.
- BRADLEY, A. P. The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recognition*, v. 30, n. 7, p. 1145–1159, 1997. Cited By :3825. Disponível em: <[www.scopus.com](http://www.scopus.com)>. Citado na página 64.
- BRASIL. Resolução bacen nº 3.823, de 16 de dezembro de 2009, procedimentos aplicáveis no reconhecimento, mensuração e divulgação de provisões, contingências passivas e contingências ativas. *Conselho Monetário Nacional*, Brasília, DF, 2009. Disponível em: <[https://www.bcb.gov.br/pre/normativos/res/2009/pdf/res\\_3823\\_v1\\_o.pdf](https://www.bcb.gov.br/pre/normativos/res/2009/pdf/res_3823_v1_o.pdf)>. Citado 3 vezes nas páginas 25, 35 e 37.
- BRASIL. Resolução bacen nº 4.553, de 30 de janeiro de 2017, estabelece a segmentação do conjunto das instituições financeiras e demais instituições autorizadas a funcionar pelo banco central do brasil para fins de aplicação proporcional da regulação prudencial. *Conselho Monetário Nacional*, Brasília, DF, 2017. Disponível em: <[https://normativos.bcb.gov.br/Lists/Normativos/Attachments/50335/Res\\_4553\\_v2\\_P.pdf](https://normativos.bcb.gov.br/Lists/Normativos/Attachments/50335/Res_4553_v2_P.pdf)>. Citado na página 38.
- BRASIL. Resolução bacen nº 4.557, de 23 de fevereiro de 2017, estrutura de gerenciamento de riscos e a estrutura de gerenciamento de capital. *Conselho Monetário Nacional*, Brasília, DF, 2017. Disponível em: <[https://www.bcb.gov.br/pre/normativos/busca/downloadNormativo.asp?arquivo=/Lists/Normativos/Attachments/50344/Res\\_4557\\_v1\\_O.pdf](https://www.bcb.gov.br/pre/normativos/busca/downloadNormativo.asp?arquivo=/Lists/Normativos/Attachments/50344/Res_4557_v1_O.pdf)>. Citado 2 vezes nas páginas 28 e 29.
- BRASIL. Site banco central do brasil - s1. Brasília, DF, 2022. Disponível em: <<https://www.bcb.gov.br/estabilidade financeira/regprudencialsegmentacao>>. Citado na página 38.
- BRASIL, B. do. Demonstrações contábeis do banco do brasil de 2021, 4º trim. *BB*, Brasília, DF, 2021. Disponível em: <<https://ri.bb.com.br/informacoes-financeiras/central-de-resultados/>>. Citado na página 38.
- BREIMAN, L. Bagging predictors. *Machine learning*, Springer, v. 24, n. 2, p. 123–140, 1996. Citado 2 vezes nas páginas 58 e 59.
- BREIMAN, L. Random forests. *Machine learning*, Springer, v. 45, n. 1, p. 5–32, 2001. Citado na página 60.
- BREIMAN, L. et al. *Classification and regression trees*. [S.l.]: Routledge, 2017. Citado 2 vezes nas páginas 57 e 64.
- CAIXA. Demonstrações contábeis da caixa econômica federal de 2021, 4º trim. *CAIXA*, Brasília, DF, 2021. Disponível em: <[https://www.caixa.gov.br/Downloads/caixa-governanca/BrGaap\\_Demonstracoes\\_Contabeis\\_4T21.pdf](https://www.caixa.gov.br/Downloads/caixa-governanca/BrGaap_Demonstracoes_Contabeis_4T21.pdf)>. Citado na página 39.
- CARVALHO, A. C. P. D. L. F. D. Inteligência artificial: riscos, benefícios e uso responsável. *Estudos Avancados*, v. 35, n. 101, p. 21–35, 2021. Disponível em: <[www.scopus.com](http://www.scopus.com)>. Citado na página 53.
- CHASE, J. M. Jp morgan consolidated financial statements and notes and supplementary information 2021. *JP Morgan Chase*, New York, EUA, 2021. Disponível em: <<https://www.jpmorganchase.com/content/dam/jpmc/jpmorgan-chase-and-co/investor-relations/documents/audited-financial-statements-2021.pdf>>. Citado na página 42.
- CHUN, W. J. *Core Python programming*. [S.l.]: Prentice Hall PTR, 2000. v. 1. Citado na página 55.

- CONTÁBEIS, C. de P. Pronunciamento técnico cpc 25 - provisões, passivos contingentes e ativos contingentes. *IASB*, Brasília, DF, 2009. Disponível em: <<http://www.cpc.org.br/CPC/Documents-Emitidos/Pronunciamentos/Pronunciamento?Id=56>>. Citado 4 vezes nas páginas 25, 35, 36 e 53.
- COSO, I. Coso 2004 - enterprise risk management-integrated framework. *Committee of Sponsoring Organizations of the Treadway Commission*, v. 1, 2004. Citado 3 vezes nas páginas 27, 30 e 33.
- COSO, I. Coso 2017 - enterprise risk management-integrated framework. *Committee of Sponsoring Organizations of the Treadway Commission*, v. 1, 2017. Citado 2 vezes nas páginas 30 e 34.
- DEITEL, P.; DEITEL LIPERI, W. *Python, how to program*. [S.l.]: Prentice Hall, 2002. Citado na página 55.
- FACELII, K. et al. *Inteligência Artificial*. [S.l.]: Grupo Editorial Nacional, 2021. v. 2. Citado na página 55.
- FAWCETT, T. An introduction to roc analysis. *Pattern Recognition Letters*, v. 27, n. 8, p. 861–874, 2006. Cited By :11828. Disponível em: <[www.scopus.com](http://www.scopus.com)>. Citado 3 vezes nas páginas 62, 63 e 64.
- FREUND, Y.; SCHAPIRE, R. E. et al. Experiments with a new boosting algorithm. In: *CITeseer.icml*. [S.l.], 1996. v. 96, p. 148–156. Citado na página 59.
- GROUP, M. U. F. Mufj bank financial statements and notes 2021. *MUFJ Bank*, Japão, 2021. Disponível em: <[https://www.mufj.jp/dam/ir/report/annual\\_report/pdf/ar2021\\_en.pdf#page178](https://www.mufj.jp/dam/ir/report/annual_report/pdf/ar2021_en.pdf#page178)>. Citado 2 vezes nas páginas 45 e 71.
- GRUS, J. *Data Science do Zero*. [S.l.]: Alta Books Editora, 2021. v. 2. Citado na página 53.
- HOPKIN, P. *Fundamentals of risk management: understanding, evaluating and implementing effective risk management*. [S.l.]: Kogan Page Publishers, 2018. Citado na página 30.
- IASB, I. A. S. B. Ias 37 provisions, contingent liabilities and contingent assets. *International Financial Reporting Standards Foundation*, United States of America, 2001. Citado 5 vezes nas páginas 25, 35, 36, 43 e 53.
- IIA. The institute of internal auditors. 2022. Disponível em: <<https://www.theiia.org/>>. Citado na página 30.
- INDUSTRIAL, I.; CHINA, C. B. of. Icbc financial statements and notes 2021. *ICBC*, China, 2021. Disponível em: <<http://www.icbc-ltd.com/ICBCLtd/InvestorRelations/DownloadCenter/>>. Citado 2 vezes nas páginas 45 e 71.
- ITAÚ. Demonstrações contábeis do itaú de 2021, 4º trim. *ITAÚ*, São Paulo, DF, 2021. Disponível em: <<https://www.itaubr.com.br/relacoes-com-investidores/Download.aspx?Arquivo=DQiqC7Br+09tg0tmYVqmvvg==>>>. Citado na página 39.
- JAMES, G. et al. *An introduction to statistical learning*. [S.l.]: Springer, 2013. v. 112. Citado 2 vezes nas páginas 56 e 57.
- JR, D. W. H.; LEMESHOW, S.; STURDIVANT, R. X. *Applied logistic regression*. [S.l.]: John Wiley & Sons, 2013. v. 398. Citado na página 57.
- KARIMI-HAGHIGHI, M.; CASTILLO, C. Enhancing a recidivism prediction tool with machine learning: Effectiveness and algorithmic fairness. In: *Proceedings of the 18th International Conference on Artificial Intelligence and Law, ICAIL 2021*. [s.n.], 2021. p. 210–214. Disponível em: <[www.scopus.com](http://www.scopus.com)>. Citado na página 54.
- KLEINBAUM, D. G. et al. *Logistic regression*. [S.l.]: Springer, 2002. Citado na página 57.
- MAIMON, O. Z.; ROKACH, L. *Data mining with decision trees: theory and applications*. [S.l.]: World scientific, 2014. v. 81. Citado na página 57.
- MOELLER, R. R. *COSO enterprise risk management: understanding the new integrated ERM framework*. [S.l.]: John Wiley & Sons, 2007. Citado na página 33.

- MOOSA, I. A. Operational risk: a survey. *Financial Markets, Institutions & Instruments*, Wiley Online Library, v. 16, n. 4, p. 167–200, 2007. Citado 2 vezes nas páginas 27 e 28.
- PARIBAS, B. Bnp - paribas financial statements and notes 2021. *BNP - Paribas*, França, 2021. Disponível em: <<https://invest.bnpparibas/en/search/reports/documents/financial-reports>>. Citado na página 46.
- PARK, M.; CHAI, S. Ai model for predicting legal judgments to improve accuracy and explainability of online privacy invasion cases. *Applied Sciences (Switzerland)*, v. 11, n. 23, 2021. Disponível em: <[www.scopus.com](http://www.scopus.com)>. Citado na página 54.
- PINHEIRO, P. P. et al. *Direito Digital Aplicado 3.0*. [S.l.]: Revista dos Tribunais, 2018. v. 1. Citado na página 53.
- PROVOST, F.; FAWCETT, T. Robust classification systems for imprecise environments. In: *Proceedings of the National Conference on Artificial Intelligence*. [s.n.], 1998. p. 706–713. Cited By :79. Disponível em: <[www.scopus.com](http://www.scopus.com)>. Citado na página 63.
- RIMS. Risk management society. 2022. Disponível em: <<https://www.rims.org>>. Citado na página 30.
- SANTANDER. Demonstrações contábeis do santander de 2021, 4º trim. *SANTANDER*, São Paulo, SP, 2021. Disponível em: <<https://www.santander.com.br/ri/resultados>>. Citado na página 40.
- SANZ, J. L. C.; ZHU, Y. Toward scalable artificial intelligence in finance. In: *Proceedings - 2021 IEEE International Conference on Services Computing, SCC 2021*. [s.n.], 2021. p. 460–469. Disponível em: <[www.scopus.com](http://www.scopus.com)>. Citado 3 vezes nas páginas 53, 54 e 72.
- SOKOLOVA, M.; LAPALME, G. A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, v. 45, n. 4, p. 427–437, 2009. Cited By :2641. Disponível em: <[www.scopus.com](http://www.scopus.com)>. Citado na página 61.
- STANDARDIZATION, I. O. for. Iso 31000 - risk management principles and guidelines. *ISO*, 2009. Citado 5 vezes nas páginas 27, 30, 31, 32 e 71.
- STANDARDIZATION, I. O. for. Iso guide 73:2009 - risk management — vocabulary. *ISO*, 2009. Citado na página 27.
- SUPERVISION, B. C. on B. Basile ii - international convergence of capital measurement and capital standards. *IASB*, Switzerland, p. 144, 2004. Disponível em: <<https://www.bis.org/publ/bcbs128.pdf>>. Citado 2 vezes nas páginas 25 e 29.
- SUTTON, C. D. Classification and regression trees, bagging, and boosting. *Handbook of statistics*, Elsevier, v. 24, p. 303–329, 2005. Citado 2 vezes nas páginas 58 e 60.
- SWETS, J. A. *Signal detection theory and ROC analysis in psychology and diagnostics: Collected papers*. [S.l.]: Psychology Press, 2014. Citado na página 63.
- TEAM, A. V. C. Tree based algorithms: A complete tutorial from scratch (in r python). *Analytics Vidhya*, 2016. Citado na página 58.
- THOMPSON, C.; HOPKIN, P. *Fundamentals of Risk Management: Understanding, Evaluating and Implementing Effective Enterprise Risk Management*. [S.l.]: Kogan Page Publishers, 2021. Citado 2 vezes nas páginas 30 e 34.
- TSAI, C. .; HSU, Y. .; YEN, D. C. A comparative study of classifier ensembles for bankruptcy prediction. *Applied Soft Computing Journal*, v. 24, p. 977–984, 2014. Cited By :101. Disponível em: <[www.scopus.com](http://www.scopus.com)>. Citado na página 58.
- UK, H. G. F. *The Orange Book, Management of Risk, Principles and Concepts*. [S.l.]: Government Finance Function of United Kingdoms, 2020. Citado na página 30.
- WANG, G. et al. A comparative assessment of ensemble learning for credit scoring. *Expert Systems with Applications*, v. 38, n. 1, p. 223–230, 2011. Cited By :301. Disponível em: <[www.scopus.com](http://www.scopus.com)>. Citado na página 61.



# Apêndices





## APÊNDICE A – LEIAUTE DOS DADOS DAS VARIÁVEIS INDEPENDENTES DE CADASTRO

Variável	Tipo	Formato
1. Duração do processo	Número	0000
2. Unidade Jurídica	Número	1 - Matriz; 2 - AJ; 3 - BE; 4 - BH; 5 - BR; 6 - BU; 7 - CB; 8 - CG; 9 - CP; 10 - CT; 11 - FL; 12 - FO; 13 - GO; 14 - JP; 15 - ME; 16 - MN; 17 - NA; 18 - PO; 19 - PV; 20 - RE; 21 - RJ; 22 - SA; 23 - SL; 24 - SP; 25 - TE; 26 - VT
3. Área Judicial	Número	2 - Feitos diversos; 5 - Trabalhista; 6 - Habitacional
4. Assunto	Número	0000
5. Quantidade de PF	Número	0 - 0; 1 - 1; 2 - 2; 3 - 3; 4 - 4; 5 - 5 6 - 6; 7 - 7; 8 - 8; 9 - 9; 10 - 10; 11 - 11 a 20 12 - 21 a 30 13 - 31 a 50 14 - 51 a 100 15 - 101 a 200 16 - 201 a 500 17 - 501 a 1000 18 - > 1000
6. Quantidade de PJ	Número	0 - 0 ; 1 - 1 ; 2 - 2 ; 3 - 3 ; 4 - 4 5 - 5; 6 - 6 a 10; 7 - > 11
7. Base	Número	1 - DF; 2 - PE; 3 - RJ; 4 - RS; 5 - SP
8. Comarca	Número	0000
9. Vara	Número	0
10. Foro	Número	0
11. Valor da causa atualizado	Número	0 - 0 1 - R\$ 1 a R\$ 1000 2 - R\$ 1001 a R\$ 5000 3 - R\$ 5001 a R\$ 10000 4 - R\$ 10001 a R\$ 30000 5 - R\$ 30001 a R\$ 50000 6 - R\$ 50001 a R\$ 100000 7 - R\$ 100001 a R\$ 500000 8 > R\$ 500001
12. Tipo de ação	Número	000
13. Macro assunto	Número	0000



# APÊNDICE B – CÓDIGO FONTE DOS MODELOS NO PYTHON

```
# Carregar a base de dados
df = pd.read_csv("BD.csv")

# Analisar a base de dados
df.info()
df.head()
df.describe()

# Checar missing
df.isnull().sum()

# Analisar as correlações
correlacao = df.corr()

# Mapa de calor das correlações
plt.figure(figsize = (38,16))
sns.heatmap(df.corr(), annot = True)
plt.savefig('heatmap.png')
plt.show()

# Separar as variáveis (variável dependente e variáveis independentes)
X = df.iloc[:,df.columns != "HOUVE_PAGAMENTO_SN"]
y = df.HOUVE_PAGAMENTO_SN

# Checando as classes da variável dependente
print ("Houve desembolso financeiro? 0=NÃO; 1=SIM")

# Separando a base treino (80%) e base teste (20%)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=10)

# Estabelecer os Modelos
model1 = Logistic Regression
model2 = Random Forest Classifier
model3 = Bagging Classifier
model4 = Ada Boost Classifier
model5 = Decision Tree Classifier
model6 = Gradient Boosting Classifier
```

```
# REGRESSÃO LOGÍSTICA (default)
    model1 = LogisticRegression()

# Treinando o modelo
    model1.fit(X_train, y_train)

# Performance do modelo - Base treino
    train_acc1 = model1.score(X_train, y_train)
    print("Logistic Regression: a acurácia do treino é de %".format(train_acc1*100))

# Performance do modelo - Base teste
    y_pred1 = model1.predict(X_test)
    test_acc1 = accuracy_score(y_test, y_pred1)
    print("Logistic Regression: a acurácia dos testes é de %".format(test_acc1*100))

# matriz de confusao
    mc1 = confusion_matrix(y_test, y_pred1)
    print (mc1)

# auc
    pred_prob1 = model1.predict_proba(X_test)
    auc_score1 = roc_auc_score(y_test, pred_prob1[:,1])
    print("Logistic Regression: %".format(auc_score1*100))

# curva roc
    fpr1, tpr1, thresh1 = roc_curve(y_test, pred_prob1[:,1], pos_label=1)

# roc curve for tpr = fpr
    random_probs = [0 for i in range(len(y_test))]
    p_fpr, p_tpr, _ = roc_curve(y_test, random_probs, pos_label=1)
    plt.plot(fpr1, tpr1, linestyle='-', color='orange', label='Logistic Regression')
    plt.plot(p_fpr, p_tpr, linestyle='-', color='blue')
    plt.title('Curva ROC')
    plt.xlabel('Taxa de positivos falsos')
    plt.ylabel('Taxa de positivos verdadeiros')
    plt.legend(loc='best')
    plt.savefig('ROC',dpi=600)
    plt.show()

# Reporte de avaliações
    print(classification_report(y_test, y_pred1))
```

---

```
# REGRESSÃO LOGÍSTICA (tuning)

    model1 = LogisticRegression(l1_ratio=None, max_iter=50, n_jobs= 1, solver = 'liblinear',
penalty = 'l1', tol=0.0001, verbose=0, random_state = 10, fit_intercept=True)

# Treinando o modelo

    model1.fit(X_train, y_train)

# Performance do modelo - Base treino

    train_acc1 = model1.score(X_train, y_train)

    print("Logistic Regression: a acurácia do treino é de %".format(train_acc1*100))

# Performance do modelo - Base teste

    y_pred1 = model1.predict(X_test)

    test_acc1 = accuracy_score(y_test, y_pred1)

    print("Logistic Regression: a acurácia dos testes é de %".format(test_acc1*100))

# matriz de confusao

    mc1 = confusion_matrix(y_test, y_pred1)

    print (mc1)

# auc

    pred_prob1 = model1.predict_proba(X_test)

    auc_score1 = roc_auc_score(y_test, pred_prob1[:,1])

    print("Logistic Regression: %".format(auc_score1*100))

# curva roc

    fpr1, tpr1, thresh1 = roc_curve(y_test, pred_prob1[:,1], pos_label=1)

# roc curve for tpr = fpr

    random_probs = [0 for i in range(len(y_test))]

    p_fpr, p_tpr, _ = roc_curve(y_test, random_probs, pos_label=1)

    plt.plot(fpr1, tpr1, linestyle='-', color='orange', label='Logistic Regression')

    plt.plot(p_fpr, p_tpr, linestyle='-', color='blue')

    plt.title('Curva ROC')

    plt.xlabel('Taxa de positivos falsos')

    plt.ylabel('Taxa de positivos verdadeiros')

    plt.legend(loc='best')

    plt.savefig('ROC',dpi=600)

    plt.show()

# Reporte de avaliações

    print(classification_report(y_test, y_pred1))
```

```

# RANDOM FOREST (default)
model2 = RandomForestClassifier()

# Treinando o modelo
    model2.fit(X_train, y_train)

# Performance do modelo - Base treino
    train_acc2 = model2.score(X_train, y_train)
    print("Random Forest Classifier: a acurácia do treino é de %".format(train_acc2*100))

# Performance do modelo - Base teste
    y_pred2 = model2.predict(X_test)
    test_acc2 = accuracy_score(y_test, y_pred2)
    print("Random Forest Classifier: a acurácia dos testes é de %".format(test_acc2*100))

# matriz de confusao
    mc2 = confusion_matrix(y_test, y_pred2)
    print (mc2)

# auc
    pred_prob2 = model2.predict_proba(X_test)
    auc_score2 = roc_auc_score(y_test, pred_prob2[:,1])
    print("Random Forest Classifier: %".format(auc_score2*100))

# curva roc
    fpr2, tpr2, thresh2 = roc_curve(y_test, pred_prob2[:,1], pos_label=1)

# roc curve for tpr = fpr
    random_probs = [0 for i in range(len(y_test))]
    p_fpr, p_tpr, _ = roc_curve(y_test, random_probs, pos_label=1)
    plt.plot(fpr2, tpr2, linestyle='-', color='green', label='Random Forest Classifier')
    plt.plot(p_fpr, p_tpr, linestyle='-', color='blue')
    plt.title('Curva ROC')
    plt.xlabel('Taxa de positivos falsos')
    plt.ylabel('Taxa de positivos verdadeiros')
    plt.legend(loc='best')
    plt.savefig('ROC',dpi=600)
    plt.show()

# Reporte de avaliações
    print(classification_report(y_test, y_pred2))

```

---

```
# RANDOM FOREST (tuning)

model2 = RandomForestClassifier(n_estimators = 400, max_depth = 15, max_features = 'sqrt',
random_state = 18, criterion = 'gini', min_samples_split = 8).fit(X_train, y_train)

# Treinando o modelo

    model2.fit(X_train, y_train)

# Performance do modelo - Base treino

    train_acc2 = model2.score(X_train, y_train)

    print("Random Forest Classifier: a acurácia do treino é de %".format(train_acc2*100))

# Performance do modelo - Base teste

    y_pred2 = model2.predict(X_test)

    test_acc2 = accuracy_score(y_test, y_pred2)

    print("Random Forest Classifier: a acurácia dos testes é de %".format(test_acc2*100))

# matriz de confusao

    mc2 = confusion_matrix(y_test, y_pred2)

    print (mc2)

# auc

    pred_prob2 = model2.predict_proba(X_test)

    auc_score2 = roc_auc_score(y_test, pred_prob2[:,1])

    print("Random Forest Classifier: %".format(auc_score2*100))

# curva roc

    fpr2, tpr2, thresh2 = roc_curve(y_test, pred_prob2[:,1], pos_label=1)

# roc curve for tpr = fpr

    random_probs = [0 for i in range(len(y_test))]

    p_fpr, p_tpr, _ = roc_curve(y_test, random_probs, pos_label=1)

    plt.plot(fpr2, tpr2, linestyle='-', color='green', label='Random Forest Classifier')

    plt.plot(p_fpr, p_tpr, linestyle='-', color='blue')

    plt.title('Curva ROC')

    plt.xlabel('Taxa de positivos falsos')

    plt.ylabel('Taxa de positivos verdadeiros')

    plt.legend(loc='best')

    plt.savefig('ROC',dpi=600)

    plt.show()

# Reporte de avaliações

    print(classification_report(y_test, y_pred2))
```

```
# BAGGING (default)
    model3 = BaggingClassifier()
# Treinando o modelo
    model3.fit(X_train, y_train)
# Performance do modelo - Base treino
    train_acc3 = model3.score(X_train, y_train)
    print("Bagging: a acurácia do treino é de %".format(train_acc3*100))
# Performance do modelo - Base teste
    y_pred3 = model3.predict(X_test)
    test_acc3 = accuracy_score(y_test, y_pred3)
    print("Bagging: a acurácia dos testes é de %".format(test_acc3*100))

# matriz de confusao
    mc3 = confusion_matrix(y_test, y_pred3)
    print (mc3)

# auc
    pred_prob3= model3.predict_proba(X_test)
    auc_score3 = roc_auc_score(y_test, pred_prob3[:,1])
    print("Bagging: %".format(auc_score3*100))

# curva roc
    fpr3, tpr3, thresh3 = roc_curve(y_test, pred_prob3[:,1], pos_label=1)
# roc curve for tpr = fpr
    random_probs = [0 for i in range(len(y_test))]
    p_fpr, p_tpr, _ = roc_curve(y_test, random_probs, pos_label=1)
    plt.plot(fpr3, tpr3, linestyle='-', color='red', label='Bagging')
    plt.plot(p_fpr, p_tpr, linestyle='-', color='blue')
    plt.title('Curva ROC')
    plt.xlabel('Taxa de positivos falsos')
    plt.ylabel('Taxa de positivos verdadeiros')
    plt.legend(loc='best')
    plt.savefig('ROC',dpi=600)
    plt.show()

# Reporte de avaliações
    print(classification_report(y_test, y_pred3))
```



---

```
# BAGGING (tuning)
    model3 = BaggingClassifier( DecisionTreeClassifier (max_features=12, splitter="random",
max_leaf_nodes=22, random_state=10))

# Treinando o modelo
    model3.fit(X_train, y_train)

# Performance do modelo - Base treino
    train_acc3 = model3.score(X_train, y_train)
    print("Bagging: a acurácia do treino é de %".format(train_acc3*100))

# Performance do modelo - Base teste
    y_pred3 = model3.predict(X_test)
    test_acc3 = accuracy_score(y_test, y_pred3)
    print("Bagging: a acurácia dos testes é de %".format(test_acc3*100))

# matriz de confusao
    mc3 = confusion_matrix(y_test, y_pred3)
    print (mc3)

# auc
    pred_prob3= model3.predict_proba(X_test)
    auc_score3 = roc_auc_score(y_test, pred_prob3[:,1])
    print("Bagging: %".format(auc_score3*100))

# curva roc
    fpr3, tpr3, thresh3 = roc_curve(y_test, pred_prob3[:,1], pos_label=1)

# roc curve for tpr = fpr
    random_probs = [0 for i in range(len(y_test))]
    p_fpr, p_tpr, _ = roc_curve(y_test, random_probs, pos_label=1)
    plt.plot(fpr3, tpr3, linestyle='-', color='red', label='Bagging')
    plt.plot(p_fpr, p_tpr, linestyle='-', color='blue')
    plt.title('Curva ROC')
    plt.xlabel('Taxa de positivos falsos')
    plt.ylabel('Taxa de positivos verdadeiros')
    plt.legend(loc='best')
    plt.savefig('ROC',dpi=600)
    plt.show()

# Reporte de avaliações
    print(classification_report(y_test, y_pred3))
```

```
# ADA BOOST (default)
    model4 = AdaBoostClassifier()

# Treinando o modelo
    model4.fit(X_train, y_train)

# Performance do modelo - Base treino
    train_acc4 = model4.score(X_train, y_train)
    print("Ada Boost Classifier: a acurácia do treino é de %".format(train_acc4*100))

# Performance do modelo - Base teste
    y_pred4 = model4.predict(X_test)
    test_acc4 = accuracy_score(y_test, y_pred4)
    print("Ada Boost Classifier: a acurácia dos testes é de %".format(test_acc4*100))

# matriz de confusao
    mc4 = confusion_matrix(y_test, y_pred4)
    print (mc4)

# auc
    pred_prob4 = model4.predict_proba(X_test)
    auc_score4 = roc_auc_score(y_test, pred_prob4[:,1])
    print("Ada Boost Classifier: %".format(auc_score4*100))

# curva roc
    fpr4, tpr4, thresh4 = roc_curve(y_test, pred_prob4[:,1], pos_label=1)
    # roc curve for tpr = fpr
    random_probs = [0 for i in range(len(y_test))]
    p_fpr, p_tpr, _ = roc_curve(y_test, random_probs, pos_label=1)
    plt.plot(fpr4, tpr4, linestyle='-', color='yellow', label='Ada Boost Classifier')
    plt.plot(p_fpr, p_tpr, linestyle='-', color='blue')
    plt.title('Curva ROC')
    plt.xlabel('Taxa de positivos falsos')
    plt.ylabel('Taxa de positivos verdadeiros')
    plt.legend(loc='best')
    plt.savefig('ROC',dpi=600)
    plt.show()

# Reporte de avaliações
    print(classification_report(y_test, y_pred4))
```

---

```
# ADA BOOST (tuning)
    model4 = AdaBoostClassifier(learning_rate=1.0, n_estimators=300, random_state = 10,
algorithm='SAMME.R')

# Treinando o modelo
    model4.fit(X_train, y_train)

# Performance do modelo - Base treino
    train_acc4 = model4.score(X_train, y_train)
    print("Ada Boost Classifier: a acurácia do treino é de %".format(train_acc4*100))

# Performance do modelo - Base teste
    y_pred4 = model4.predict(X_test)
    test_acc4 = accuracy_score(y_test, y_pred4)
    print("Ada Boost Classifier: a acurácia dos testes é de %".format(test_acc4*100))

# matriz de confusao
    mc4 = confusion_matrix(y_test, y_pred4)
    print (mc4)

# auc
    pred_prob4 = model4.predict_proba(X_test)
    auc_score4 = roc_auc_score(y_test, pred_prob4[:,1])
    print("Ada Boost Classifier: %".format(auc_score4*100))

# curva roc
    fpr4, tpr4, thresh4 = roc_curve(y_test, pred_prob4[:,1], pos_label=1)

# roc curve for tpr = fpr
    random_probs = [0 for i in range(len(y_test))]
    p_fpr, p_tpr, _ = roc_curve(y_test, random_probs, pos_label=1)
    plt.plot(fpr4, tpr4, linestyle='-', color='yellow', label='Ada Boost Classifier')
    plt.plot(p_fpr, p_tpr, linestyle='-', color='blue')
    plt.title('Curva ROC')
    plt.xlabel('Taxa de positivos falsos')
    plt.ylabel('Taxa de positivos verdadeiros')
    plt.legend(loc='best')
    plt.savefig('ROC',dpi=600)
    plt.show()

# Reporte de avaliações
    print(classification_report(y_test, y_pred4))
```

```
# DECISION TREE (default)
    model5 = DecisionTreeClassifier()

# Treinando o modelo
    model5.fit(X_train, y_train)

# Performance do modelo - Base treino
    train_acc5 = model5.score(X_train, y_train)
    print("Decision Tree Classifier: a acurácia do treino é de %".format(train_acc5*100))

# Performance do modelo - Base teste
    y_pred5 = model5.predict(X_test)
    test_acc5 = accuracy_score(y_test, y_pred5)
    print("Decision Tree Classifier: a acurácia dos testes é de %".format(test_acc5*100))

# matriz de confusao
    mc5 = confusion_matrix(y_test, y_pred5)
    print (mc5)

# auc
    pred_prob5 = model5.predict_proba(X_test)
    auc_score5 = roc_auc_score(y_test, pred_prob5[:,1])
    print("Decision Tree Classifier: %".format(auc_score5*100))

# curva roc
    fpr5, tpr5, thresh5 = roc_curve(y_test, pred_prob5[:,1], pos_label=1)
    # roc curve for tpr = fpr
    random_probs = [0 for i in range(len(y_test))]
    p_fpr, p_tpr, _ = roc_curve(y_test, random_probs, pos_label=1)
    plt.plot(fpr5, tpr5, linestyle='-', color='purple', label='Decision Tree Classifier')
    plt.plot(p_fpr, p_tpr, linestyle='-', color='blue')
    plt.title('Curva ROC')
    plt.xlabel('Taxa de positivos falsos')
    plt.ylabel('Taxa de positivos verdadeiros')
    plt.legend(loc='best')
    plt.savefig('ROC',dpi=600)
    plt.show()

# Reporte de avaliações
    print(classification_report(y_test, y_pred5))
```

---

```
# DECISION TREE (tuning)
    model5 = DecisionTreeClassifier( criterion = 'gini', max_depth = 20,min_samples_split=100,
random_state=10)

# Treinando o modelo
    model5.fit(X_train, y_train)

# Performance do modelo - Base treino
    train_acc5 = model5.score(X_train, y_train)
    print("Decision Tree Classifier: a acurácia do treino é de %".format(train_acc5*100))

# Performance do modelo - Base teste
    y_pred5 = model5.predict(X_test)
    test_acc5 = accuracy_score(y_test, y_pred5)
    print("Decision Tree Classifier: a acurácia dos testes é de %".format(test_acc5*100))

# matriz de confusao
    mc5 = confusion_matrix(y_test, y_pred5)
    print (mc5)

# auc
    pred_prob5 = model5.predict_proba(X_test)
    auc_score5 = roc_auc_score(y_test, pred_prob5[:,1])
    print("Decision Tree Classifier: %".format(auc_score5*100))

# curva roc
    fpr5, tpr5, thresh5 = roc_curve(y_test, pred_prob5[:,1], pos_label=1)

# roc curve for tpr = fpr
    random_probs = [0 for i in range(len(y_test))]
    p_fpr, p_tpr, _ = roc_curve(y_test, random_probs, pos_label=1)
    plt.plot(fpr5, tpr5, linestyle='-', color='purple', label='Decision Tree Classifier')
    plt.plot(p_fpr, p_tpr, linestyle='-', color='blue')
    plt.title('Curva ROC')
    plt.xlabel('Taxa de positivos falsos')
    plt.ylabel('Taxa de positivos verdadeiros')
    plt.legend(loc='best')
    plt.savefig('ROC',dpi=600)
    plt.show()

# Reporte de avaliações
    print(classification_report(y_test, y_pred5))
```

```
# GRADIENT BOOSTING (default)
    model6 = GradientBoostingClassifier()

# Treinando o modelo
    model6.fit(X_train, y_train)

# Performance do modelo - Base treino
    train_acc6 = model6.score(X_train, y_train)
    print("Gradient Boosting: a acurácia do treino é de %".format(train_acc6*100))

# Performance do modelo - Base teste
    y_pred6 = model6.predict(X_test)
    test_acc6 = accuracy_score(y_test, y_pred6)
    print("Gradient Boosting: a acurácia dos testes é de %".format(test_acc6*100))

# matriz de confusao
    mc6 = confusion_matrix(y_test, y_pred6)
    print (mc6)

# auc
    pred_prob6 = model6.predict_proba(X_test)
    auc_score6 = roc_auc_score(y_test, pred_prob6[:,1])
    print("Gradient Boosting: %".format(auc_score6*100))

# curva roc
    fpr6, tpr6, thresh6 = roc_curve(y_test, pred_prob6[:,1], pos_label=1)
    # roc curve for tpr = fpr
    random_probs = [0 for i in range(len(y_test))]
    p_fpr, p_tpr, _ = roc_curve(y_test, random_probs, pos_label=1)
    plt.plot(fpr6, tpr6, linestyle='-', color='pink', label='Gradient Boosting')
    plt.plot(p_fpr, p_tpr, linestyle='-', color='blue')
    plt.title('Curva ROC')
    plt.xlabel('Taxa de positivos falsos')
    plt.ylabel('Taxa de positivos verdadeiros')
    plt.legend(loc='best')
    plt.savefig('ROC',dpi=600)
    plt.show()

# Reporte de avaliações
    print(classification_report(y_test, y_pred6))
```

---

```

# GRADIENT BOOSTING (tuning)
    model6 = GradientBoostingClassifier (learning_rate=0.5, n_estimators=500, tol=0.1,
random_state=10)

# Treinando o modelo
    model6.fit(X_train, y_train)

# Performance do modelo - Base treino
    train_acc6 = model6.score(X_train, y_train)
    print("Gradient Boosting: a acurácia do treino é de %".format(train_acc6*100))

# Performance do modelo - Base teste
    y_pred6 = model6.predict(X_test)
    test_acc6 = accuracy_score(y_test, y_pred6)
    print("Gradient Boosting: a acurácia dos testes é de %".format(test_acc6*100))

# matriz de confusao
    mc6 = confusion_matrix(y_test, y_pred6)
    print (mc6)

# auc
    pred_prob6 = model6.predict_proba(X_test)
    auc_score6 = roc_auc_score(y_test, pred_prob6[:,1])
    print("Gradient Boosting: %".format(auc_score6*100))

# curva roc
    fpr6, tpr6, thresh6 = roc_curve(y_test, pred_prob6[:,1], pos_label=1)

# roc curve for tpr = fpr
    random_probs = [0 for i in range(len(y_test))]
    p_fpr, p_tpr, _ = roc_curve(y_test, random_probs, pos_label=1)
    plt.plot(fpr6, tpr6, linestyle='-', color='pink', label='Gradient Boosting')
    plt.plot(p_fpr, p_tpr, linestyle='-', color='blue')
    plt.title('Curva ROC')
    plt.xlabel('Taxa de positivos falsos')
    plt.ylabel('Taxa de positivos verdadeiros')
    plt.legend(loc='best')
    plt.savefig('ROC',dpi=600)
    plt.show()

# Reporte de avaliações
    print(classification_report(y_test, y_pred6))

```

**# IMPRESSÃO DE RESULTADOS (tuning)****# Acurácia**

```
acuracia_1 = metrics.accuracy_score(y_test, y_pred1)
acuracia_2 = metrics.accuracy_score(y_test, y_pred2)
acuracia_3 = metrics.accuracy_score(y_test, y_pred3)
acuracia_4 = metrics.accuracy_score(y_test, y_pred4)
acuracia_5 = metrics.accuracy_score(y_test, y_pred5)
acuracia_6 = metrics.accuracy_score(y_test, y_pred6)

print('acuracia_1:', acuracia_1)
print('acuracia_2:', acuracia_2)
print('acuracia_3:', acuracia_3)
print('acuracia_4:', acuracia_4)
print('acuracia_5:', acuracia_5)
print('acuracia_6:', acuracia_6)
```

**# Sensibilidade**

```
sensibilidade_1 = recall_score(y_test, y_pred1)
sensibilidade_2 = recall_score(y_test, y_pred2)
sensibilidade_3 = recall_score(y_test, y_pred3)
sensibilidade_4 = recall_score(y_test, y_pred4)
sensibilidade_5 = recall_score(y_test, y_pred5)
sensibilidade_6 = recall_score(y_test, y_pred6)

print('sensibilidade_1:', sensibilidade_1)
print('sensibilidade_2:', sensibilidade_2)
print('sensibilidade_3:', sensibilidade_3)
print('sensibilidade_4:', sensibilidade_4)
print('sensibilidade_5:', sensibilidade_5)
print('sensibilidade_6:', sensibilidade_6)
```

**#Especificidade —  $TN/(TN + FP)$** 

```
especificidade_1 = round(mc1[0,0]/(mc1[0,0]+mc1[0,1]),15)
especificidade_2 = round(mc2[0,0]/(mc2[0,0]+mc2[0,1]),15)
especificidade_3 = round(mc3[0,0]/(mc3[0,0]+mc3[0,1]),15)
especificidade_4 = round(mc4[0,0]/(mc4[0,0]+mc4[0,1]),15)
especificidade_5 = round(mc5[0,0]/(mc5[0,0]+mc5[0,1]),15)
especificidade_6 = round(mc6[0,0]/(mc6[0,0]+mc6[0,1]),15)
```



```
print('especificidade_1', especificidade_1 )
print('especificidade_2', especificidade_2 )
print('especificidade_3', especificidade_3 )
print('especificidade_4', especificidade_4 )
print('especificidade_5', especificidade_5 )
print('especificidade_6', especificidade_6 )
```

#### **# Precisão**

```
precision_1 = precision_score(y_test, y_pred1)
precision_2 = precision_score(y_test, y_pred2)
precision_3 = precision_score(y_test, y_pred3)
precision_4 = precision_score(y_test, y_pred4)
precision_5 = precision_score(y_test, y_pred5)
precision_6 = precision_score(y_test, y_pred6)
print ('precisao_1: ', precision_1)
print ('precisao_2: ', precision_2)
print ('precisao_3: ', precision_3)
print ('precisao_4: ', precision_4)
print ('precisao_5: ', precision_5)
print ('precisao_6: ', precision_6)
```

#### **# f1 score**

```
f1_1 = f1_score(y_test, y_pred1)
f1_2 = f1_score(y_test, y_pred2)
f1_3 = f1_score(y_test, y_pred3)
f1_4 = f1_score(y_test, y_pred4)
f1_5 = f1_score(y_test, y_pred5)
f1_6 = f1_score(y_test, y_pred6)
print ('f1_1:',f1_1)
print ('f1_2:',f1_2)
print ('f1_3:',f1_3)
print ('f1_4:',f1_4)
print ('f1_5:',f1_5)
print ('f1_6:',f1_6)
```

**# ROC e AUC**

```
roc_auc_1 = roc_auc_score(y_test, y_pred1)
roc_auc_2 = roc_auc_score(y_test, y_pred2)
roc_auc_3 = roc_auc_score(y_test, y_pred3)
roc_auc_4 = roc_auc_score(y_test, y_pred4)
roc_auc_5 = roc_auc_score(y_test, y_pred5)
roc_auc_6 = roc_auc_score(y_test, y_pred6)
print('roc_auc_1:', roc_auc_1)
print('roc_auc_2:', roc_auc_2)
print('roc_auc_3:', roc_auc_3)
print('roc_auc_4:', roc_auc_4)
print('roc_auc_5:', roc_auc_5)
print('roc_auc_6:', roc_auc_6)
```

## Anexos



# ANEXO A – DF 2021 - BANCO DO BRASIL



Valores expressos em milhares de Reais, exceto quando indicado

## Provisões para demandas cíveis, trabalhistas e fiscais – prováveis

O Banco constitui provisão para demandas cíveis, trabalhistas e fiscais com risco de perda “provável”, quantificada utilizando metodologia individualizada ou massificada (contempla os processos com probabilidade de êxito do autor igual a remoto, possível ou provável), de acordo com a natureza e/ou valor do processo.

As estimativas do desfecho e do efeito financeiro são determinadas pela natureza das ações, pelo julgamento da administração da entidade, por meio da opinião dos assessores jurídicos com base nos elementos do processo, complementadas pela complexidade e pela experiência de demandas semelhantes.

A Administração do Banco considera suficientes as provisões constituídas para atendimento às perdas decorrentes de demandas cíveis, trabalhistas e fiscais.

## Movimentações nas provisões para demandas cíveis, trabalhistas e fiscais, classificadas como prováveis

	BB Banco Múltiplo			BB Consolidado		
	2º Semestre/2021	Exercício/2021	Exercício/2020	2º Semestre/2021	Exercício/2021	Exercício/2020
<b>Demandas Cíveis</b>						
Saldo inicial	10.619.079	9.195.502	8.697.684	10.787.957	9.579.206	8.851.582
Constituição	2.408.895	5.252.536	3.274.000	2.444.735	5.339.495	3.547.535
Reversão da provisão	(130.505)	(190.840)	(290.614)	(147.065)	(475.140)	(326.014)
Baixa por pagamento	(1.788.463)	(3.284.522)	(2.710.405)	(1.875.587)	(3.372.074)	(2.714.146)
Atualização monetária e variação cambial	194.446	330.776	224.837	199.224	337.777	220.249
<b>Saldo final</b>	<b>11.303.452</b>	<b>11.303.452</b>	<b>9.195.502</b>	<b>11.409.264</b>	<b>11.409.264</b>	<b>9.579.206</b>
<b>Demandas Trabalhistas</b>						
Saldo inicial	4.534.838	4.442.070	3.610.333	4.593.447	4.499.477	3.663.769
Constituição	1.032.464	1.970.061	2.392.065	1.043.327	1.987.572	2.410.796
Reversão da provisão	(116.206)	(228.004)	(427.202)	(125.319)	(241.631)	(440.413)
Baixa por pagamento	(929.491)	(1.810.097)	(1.377.661)	(930.140)	(1.812.577)	(1.382.776)
Atualização monetária e variação cambial	163.302	310.877	244.535	165.604	314.078	248.101
<b>Saldo final</b>	<b>4.684.907</b>	<b>4.684.907</b>	<b>4.442.070</b>	<b>4.746.919</b>	<b>4.746.919</b>	<b>4.499.477</b>
<b>Demandas Fiscais</b>						
Saldo inicial	478.256	408.370	382.388	506.306	440.539	415.396
Constituição	179.390	284.623	192.791	185.474	295.824	211.523
Reversão da provisão	(30.182)	(44.047)	(118.376)	(32.601)	(54.141)	(136.246)
Baixa por pagamento	(99.122)	(123.714)	(55.049)	(99.122)	(123.714)	(56.628)
Atualização monetária e variação cambial	9.438	12.548	6.616	10.299	11.848	6.494
<b>Saldo final</b>	<b>537.780</b>	<b>537.780</b>	<b>408.370</b>	<b>570.356</b>	<b>570.356</b>	<b>440.539</b>
<b>Total das Demandas Cíveis, Trabalhistas e Fiscais</b>	<b>16.526.139</b>	<b>16.526.139</b>	<b>14.045.942</b>	<b>16.726.539</b>	<b>16.726.539</b>	<b>14.519.222</b>

## Cronograma esperado de desembolsos

	BB Banco Múltiplo			BB Consolidado		
	Cíveis	Trabalhistas	Fiscais	Cíveis	Trabalhistas	Fiscais
Até 5 anos	9.842.026	4.251.272	311.074	9.936.164	4.312.922	336.387
Acima de 5 anos	1.461.426	433.635	226.706	1.473.100	433.997	233.969
<b>Total</b>	<b>11.303.452</b>	<b>4.684.907</b>	<b>537.780</b>	<b>11.409.264</b>	<b>4.746.919</b>	<b>570.356</b>



Banco do Brasil S.A. – Demonstrações Contábeis Exercício/2021

Valores expressos em milhares de Reais, exceto quando indicado

O cenário de imprevisibilidade do tempo de duração dos processos, bem como a possibilidade de alterações na jurisprudência dos tribunais tornam incertos os valores e o cronograma esperado de saída.

#### Passivos contingentes – possíveis

As demandas cíveis, trabalhistas e fiscais são classificadas como passivos contingentes possíveis quando não há elementos seguros que permitam concluir o resultado final do processo e quando a probabilidade de perda é inferior à provável e superior à remota, ficando dispensadas de constituição de provisão.

#### Saldos dos passivos contingentes classificados como possíveis

	BB Banco Múltiplo		BB Consolidado	
	31.12.2021	31.12.2020	31.12.2021	31.12.2020
Demandas Fiscais <sup>(1)</sup>	12.192.801	13.115.848	13.021.018	14.000.019
Demandas Cíveis	2.248.542	2.351.648	2.530.401	2.527.303
Demandas Trabalhistas	87.689	98.276	124.622	155.729
<b>Total</b>	<b>14.529.032</b>	<b>15.565.772</b>	<b>15.676.041</b>	<b>16.683.051</b>

(1) As principais contingências têm origem em (i) autos de infração lavrados pelo INSS, visando o recolhimento de contribuições incidentes sobre verbas de transporte coletivo e utilização de veículo próprio por empregados do Banco do Brasil, no valor de R\$ 296.785 mil, e participações nos lucros e resultados de funcionários, correspondentes ao período de abril de 2001 a outubro de 2003, no valor de R\$ 1.292.963 mil e (ii) autos de infração lavrados pelas Fazendas Públicas dos Municípios visando a cobrança de ISSQN, no montante de R\$ 1.817.885 mil.

#### Depósitos em garantia de recursos

#### Saldos dos depósitos em garantia constituídos para as contingências

	BB Banco Múltiplo		BB Consolidado	
	31.12.2021	31.12.2020	31.12.2021	31.12.2020
Demandas Cíveis	18.558.641	18.643.924	18.621.596	18.687.428
Demandas Fiscais	8.989.897	8.852.638	9.566.584	9.411.760
Demandas Trabalhistas	6.734.105	6.242.101	6.765.433	6.275.600
<b>Total</b>	<b>34.282.643</b>	<b>33.738.663</b>	<b>34.953.613</b>	<b>34.374.788</b>

#### c) Obrigações legais

O Banco mantém registrado em Provisões – Cíveis, trabalhistas e fiscais e em Passivos Fiscais – Correntes o montante de R\$ 19.673.693 mil (R\$ 19.333.653 mil em 31.12.2020), relativo à seguinte ação:

Em 1998, o Banco pleiteou a compensação integral dos prejuízos fiscais acumulados de Imposto de Renda e das bases de cálculo negativas de Contribuição Social sobre o Lucro Líquido. Assim, passou a compensar integralmente prejuízos fiscais (até outubro/2005)



Valores expressos em milhares de Reais, exceto quando indicado

**d) Garantias financeiras prestadas**

	BB Banco Múltiplo				BB Consolidado			
	31.12.2021		31.12.2020		31.12.2021		31.12.2020	
	Valores Garantidos	Provisão	Valores Garantidos	Provisão	Valores Garantidos	Provisão	Valores Garantidos	Provisão
Outras garantias financeiras prestadas <sup>(1)</sup>	4.369.236	540.416	2.976.940	598.932	3.192.685	540.656	2.671.307	599.156
Aval ou fiança em processos judiciais e administrativos de natureza fiscal	537.419	63.507	456.289	55.680	408.354	63.507	327.223	55.680
Vinculadas a licitações, leilões, prestação de serviços ou execução de obras	1.213.208	1.379	906.105	89.646	1.213.208	1.379	906.105	89.646
Outras fianças bancárias	6.083.017	10.224	5.320.997	7.883	6.214.700	10.224	5.453.866	7.883
Vinculadas ao comércio internacional de mercadorias	167.249	372	46.442	55	167.249	372	46.442	55
Vinculadas ao fornecimento de mercadorias	57.733	20	161.941	53	57.733	20	161.941	53
Outros avals	--	--	--	--	2.288	--	2.308	--
<b>Total</b>	<b>12.427.862</b>	<b>615.918</b>	<b>9.868.714</b>	<b>752.249</b>	<b>11.256.217</b>	<b>616.158</b>	<b>9.569.192</b>	<b>752.473</b>

(1) Referem-se, principalmente, a garantias prestadas em moeda estrangeira.

As operações de garantias financeiras prestadas são avaliadas através dos modelos de classificação de risco de operações vigentes na instituição, no mesmo formato das operações de crédito, as quais seguem os preceitos das Resoluções CMN n.º 2.682 e n.º 2.697, divulgadas em 21.12.1999 e 24.02.2000, respectivamente, que estabelecem os critérios de classificação das operações de crédito e as regras para constituição de provisões para perdas associadas ao risco de crédito.

A classificação de risco das operações é realizada mediante a aplicação de metodologias desenvolvidas que consideram as características dos clientes, das operações e garantias. O resultado final da classificação é a atribuição de risco conforme escala constante na Resolução CMN n.º 2.682/1999, que define o percentual de provisão que deve ser alocada à operação.

**e) Resultado de provisões**

	BB Banco Múltiplo			BB Consolidado		
	2º Semestre/2021	Exercício/2021	Exercício/2020	2º Semestre/2021	Exercício/2021	Exercício/2020
<b>Cíveis, fiscais e trabalhistas</b>	<b>(2.971.624)</b>	<b>(6.302.931)</b>	<b>(3.870.486)</b>	<b>(3.020.764)</b>	<b>(6.391.020)</b>	<b>(4.176.118)</b>
Cíveis	(1.492.654)	(3.656.833)	(1.367.435)	(1.533.051)	(3.737.194)	(1.663.055)
Trabalhistas	(1.079.560)	(2.052.934)	(2.209.398)	(1.083.612)	(2.060.019)	(2.218.484)
Fiscais	(158.646)	(253.124)	(81.031)	(163.172)	(253.531)	(81.771)
Provisão para riscos fiscais (atualização do depósito)	(240.764)	(340.040)	(212.622)	(240.929)	(340.276)	(212.808)
<b>Outras</b>	<b>74.179</b>	<b>(65.346)</b>	<b>(905.038)</b>	<b>76.234</b>	<b>(63.291)</b>	<b>(905.049)</b>
Garantias financeiras prestadas	80.127	137.265	(317.008)	82.182	139.320	(317.019)
Outras	(5.948)	(202.611)	(588.030)	(5.948)	(202.611)	(588.030)
<b>Total</b>	<b>(2.897.445)</b>	<b>(6.368.277)</b>	<b>(4.775.524)</b>	<b>(2.944.530)</b>	<b>(6.454.311)</b>	<b>(5.081.167)</b>





## ANEXO B – DF 2021 - CAIXA

NOTAS EXPLICATIVAS ÀS DEMONSTRAÇÕES CONTÁBEIS INDIVIDUAIS E CONSOLIDADAS - 31 DE DEZEMBRO DE 2021  
Em milhares de reais, exceto quando indicado de outra forma.



- CNP (Holding XS1) no montante de R\$ 6.720.000 em cumprimento ao contrato de distribuição. O saldo é apropriado mensalmente pela vigência do contrato, até 2045.
- Tokio Marine (XS3) no montante de R\$ 1.444.000, CNP (XS5) no montante de R\$ 240.625, Icatu (XS4) no montante de R\$ 173.250 e Tempo (XS6) no montante de R\$ 28.500, em cumprimento ao contrato de distribuição pelo prazo de 20 anos com a Caixa Seguridade.
- FISERV, no montante de R\$ 154.000, buscando fortalecer sua atuação no mercado de meios de pagamento eletrônicos, prevê um contrato com vigência de 20 anos com a Caixa Cartões.
- VR BENEFÍCIOS e FLEETCOR ("VR-FLEETCOR"), no montante de R\$ 398.333, para atuação no segmento relacionado a meios de pagamento Pré-Pagos, pelo prazo de 20 anos, conforme contrato firmado com a Caixa Cartões.

### Nota 19 - Provisões

#### (a) Composição

Descrição	INDIVIDUAL		CONSOLIDADO
	31/12/2021	31/12/2020	31/12/2020
Trabalhistas (b.1)	6.991.311	6.233.115	6.233.118
Cíveis (b.1)	3.279.719	2.929.586	2.929.586
Fiscais (b.1)	378.813	409.275	409.275
Pré-pagamento FGTS (c)	1.259.083	1.006.982	1.006.982
Bolsa família (d)	288.479	276.797	276.797
Garantias financeiras prestadas (e)	49.418	26.395	26.395
Demais	38.427	38.826	54.613
<b>Total (1)</b>	<b>12.285.250</b>	<b>10.920.976</b>	<b>10.936.766</b>
Passivo circulante	5.068.749	4.399.878	4.415.668
Passivo não circulante	7.216.501	6.521.098	6.521.098

(1) Os montantes das provisões na demonstração individual e consolidada são coincidentes em 31/12/2021.

#### (b) Provisões para causas judiciais e obrigações legais

A CAIXA é parte em processos judiciais e administrativos de natureza tributária, cível e trabalhista, decorrentes do curso normal de seus negócios. Com base em pareceres dos seus advogados e levando em consideração que os procedimentos adotados pela CAIXA guardam conformidade com as previsões legais e regulamentares, a Administração entende que as provisões constituídas são suficientes para suportar os riscos de eventuais decisões desfavoráveis nesses processos.

Considerando o elevado número de processos administrativos e judiciais, a CAIXA utiliza a seguinte metodologia para calcular o valor provável de desembolso:

a) individualizada, na qual se estima o valor provável da condenação (valor provisionado); este cálculo parte da repercussão econômica dos pedidos feitos pelo autor e é ponderado com a situação do processo e a jurisprudência predominante em causas similares; tais ações são classificadas como de perda provável, possível ou remota;

b) massificada, em que o valor provisionado corresponde ao valor médio histórico de condenação pago em processos similares nos últimos 36 meses, multiplicado pelo total de processos provisionáveis ativos, e são classificadas como prováveis.

No 4º trimestre de 2021, foi implementada a metodologia massificada para cálculo do valor provável de desembolso para ações judiciais trabalhistas, que consiste na mensuração da probabilidade e do impacto do resultado desfavorável à instituição. A aplicação do novo modelo está em transição com previsão de conclusão em 2022.

NOTAS EXPLICATIVAS ÀS DEMONSTRAÇÕES CONTÁBEIS INDIVIDUAIS E CONSOLIDADAS - 31 DE DEZEMBRO DE 2021  
Em milhares de reais, exceto quando indicado de outra forma.



Abaixo, apresentamos as ações são agrupadas em demandas fiscais, cíveis e trabalhistas considerando a matéria tratada e a relevância econômica do grupo.

**(b.1) Movimentação das provisões para causas judiciais e obrigações legais – fiscais e previdenciárias**

Descrição	INDIVIDUAL						31/12/2021
	31/12/2020	Movimentação em 2021					
		Novas provisões	Atualização monetária	Adições de provisões existentes	Reversões de provisões existentes	Baixas por pagamento	
Trabalhistas (b.2) (1) (2)	6.233.115	562.018	267.487	1.759.263	(850.531)	(980.041)	6.991.311
Cíveis (b.3)	2.929.586	346.911	133.233	489.212	(164.281)	(454.942)	3.279.719
Feitos Diversos	954.584	243.139	64.091	169.477	(33.945)	(126.963)	1.270.383
Poupança - Planos Econômicos	1.069.853	17.956	18.587	225.988	(121.342)	(258.269)	952.773
Contingenciamento do FGTS	539.324	-	38.209	73.319	-	(9.740)	641.112
Habitacional	365.825	85.816	12.346	20.428	(8.994)	(59.970)	415.451
Fiscais (b.4)	409.275	39.714	6.388	22.340	(92.296)	(6.608)	378.813
ISSQN	255.941	33.417	5.056	11.209	(17.587)	(3.827)	284.209
INSS	56.930	710	644	4.864	(15.027)	(6)	48.115
IPTU	23.259	1.663	214	906	(6.585)	(2.367)	17.090
Outros	73.145	3.924	474	5.361	(53.097)	(408)	29.399
<b>Total</b>	<b>9.571.976</b>	<b>948.643</b>	<b>407.108</b>	<b>2.270.815</b>	<b>(1.107.108)</b>	<b>(1.441.591)</b>	<b>10.649.843</b>

(1) A movimentação das provisões são coincidentes na demonstração individual e consolidada, com exceção do item "Trabalhistas" que em 31/12/2020 apresentava na demonstração consolidada o montante de R\$ 6.233.118.

(2) A alteração na metodologia para cálculo do valor provável de desembolso para ações judiciais trabalhistas rotineiras e a individualização da provisão das ações trabalhistas em fase de execução, realizados no 4º trimestre de 2021, ensejaram na constituição líquida de provisão, no mesmo período, no montante de R\$ 891,62 milhões.

**(b.2) Ações trabalhistas**

A CAIXA é parte passiva em ações ajuizadas por empregados, ex-empregados próprios ou de prestadoras de serviços e sindicatos, relacionadas com a atividade laboral, planos de cargos, acordos coletivos, indenizações, benefícios, aposentadorias, subsidiariedade, entre outros.

Em 31/12/2021, constavam 56.754 processos trabalhistas provisionados, sendo 45.811 "rotineiros" e 10.943 "relevantes", cujos respectivos valores são corrigidos pelo IPCA-E.

Visando a reduzir o litígio judicial e diminuir os valores despendidos nos processos, a CAIXA continua executando sua política de conciliação judicial e extrajudicial, realiza o cumprimento espontâneo de certas decisões judiciais e faz a análise das perdas incorridas a fim de mitigar novos litígios de causas semelhantes. Em virtude disto, as ações relevantes não são individualmente divulgadas para não inviabilizar a realização de acordos.

**(b.3) Ações cíveis**

A CAIXA é parte passiva em ações cíveis de natureza indenizatória/contratual relativas a seus produtos, serviços e atendimento. Em 31/12/2021, constavam 413.504 processos cíveis provisionados, sendo 411.303 "rotineiros" e 2.201 "relevantes".

Destacam-se demandas que contestam o expurgo de indexadores de planos econômicos, como parte da política econômica do governo federal no combate aos índices inflacionários no passado, quando da correção de saldos em depósitos em cadernetas de poupança.

A CAIXA cumpriu a ordem legal vigente à época, entretanto, considerando as ações efetivamente notificadas e a análise da jurisprudência atual do Superior Tribunal de Justiça – STJ, em 31/12/2021, foi provisionado R\$ 952.773 para estes processos (31/12/2020 – R\$ 1.069.853).



Considerando-se o posicionamento jurisprudencial atual, o direito a novas postulações encontra-se prescrito, caracterizando a inexistência de passivo potencial representativo de novas ações judiciais. O Supremo Tribunal Federal – STF suspendeu a análise de todos os recursos interpostos até que sejam proferidas as decisões das matérias de repercussão geral relativas ao Planos Bresser e Verão (tema 264), de valores não bloqueados do Plano Collor I (tema 265), Plano Collor I (tema 284) e do Plano Collor II (tema 285) que terão efeitos vinculantes a todos os casos relacionados.

No final do exercício de 2017 FEBRABAN e CONSIF celebraram acordo com as principais entidades de defesa do consumidor para equacionamento do problema. Este acordo foi homologado pelo STF e a CAIXA aderiu a seus termos. Os pagamentos começaram em julho de 2018 e continuam sendo realizados em atendimento às adesões processadas no site <https://portalacordo.pagamentodapoupanca.com.br/> e mutirões de conciliação administrativos e judiciais.

O acordo teve prazo final em 12/03/2020, razão pela qual a FEBRABAN e as entidades de defesa do consumidor chegaram a bom termo para prorrogar o acordo anterior por mais 5 anos, tendo sido homologado pelo STF a prorrogação por 30 meses, podendo chegar a 60 meses, condicionado à análise dos resultados obtidos.

Também são significativos os processos que visam à reparação de danos envolvendo o contingenciamento de repasses de recursos do FGTS. O valor provisionado, em 31/12/2021, para estes processos é de R\$ 641.112 (31/12/2020 – R\$ 539.324).

As ações indenizatórias de perdas e danos se referem a eventuais problemas ocorridos no atendimento bancário, na prestação de serviços ou na aquisição/manutenção de algum produto.

Em 2021, a CAIXA continua executando sua política de conciliação judicial e extrajudicial, realizando o cumprimento espontâneo de decisões judiciais e analisando perdas incorridas a fim de mitigar novos litígios de causas semelhantes. Até 31/12/2021, realizou 65.545 acordos processuais (76% referente a ações indenizatórias e 24% referentes a recuperações de créditos) reduzindo o montante que seria despendido caso persistisse a condenação judicial e oportunizando ao cliente uma solução rápida para o problema.

#### **(b.4) Obrigações legais – fiscais e previdenciárias**

A CAIXA, como instituição cumpridora regular das obrigações fiscais e tributárias que repercutem sobre suas atividades e serviços, discute em sede judicial a legitimidade dos parâmetros de cobrança levados a efeito por órgãos fazendários dos diversos entes da Federação, de acordo com a especificidade própria de cada caso.

As provisões constituídas sob avaliação de risco provável com base em pareceres dos assessores jurídicos se referem a ações sobre tributos e contribuições. A CAIXA acompanha regularmente o ciclo das ações judiciais em andamento, as quais, a médio e longo prazo, poderão apresentar desdobramentos favoráveis à instituição com a reversão das respectivas provisões.

Destacam-se as atuações do Instituto Nacional do Seguro Social (INSS) para o recolhimento de contribuições previdenciárias sobre pagamentos a empregados da CAIXA, em que se discute o caráter indenizatório e não remuneratório de algumas verbas, tais como auxílio-alimentação, APIP e licença-prêmio, cujos valores, repositados para 31/12/2021, correspondem ao total de R\$ 2.233.577 (31/12/2020 – R\$ 2.341.635), para as quais a provisão constituída com base no histórico de êxito e do cenário jurisprudencial, considerados em recente análise técnica e jurídica acerca da matéria, é de R\$ 48.115 (31/12/2020 – R\$ 56.930).

Em relação ao ISSQN, a CAIXA aplica as diretrizes da Lei Complementar Federal nº 116/2003, em adequação de seus sistemas e procedimentos para apuração da base de cálculo e recolhimento do tributo sobre serviços prestados.

Não obstante, as fiscalizações tributárias de municípios diversos da federação autuaram a instituição sob a alegação de falta de recolhimento ou recolhimento a menor, instaurando discussão a partir de interpretação distinta de aspectos como materialidade, alíquotas aplicáveis e local de incidência do tributo, cujo valor total, em 31/12/2021, corresponde ao montante de R\$ 1.127.117 (31/12/2020 – R\$ 1.028.824).

Em face do histórico de êxito e do cenário jurisprudencial, avaliados em análise técnica e jurídica acerca da matéria, na posição de 31/12/2021, o valor da provisão constituída é de R\$ 284.209 (31/12/2020 – R\$ 255.941).



Ainda, como destaque, a CAIXA vem discutindo a materialidade de débitos de CSLL, IRPJ e multa decorrentes de não homologações de Pedido Eletrônico de Restituição, Ressarcimento ou Reembolso e Declaração de Compensação – PER/DCOMP, os quais, em 31/12/2021, totalizam R\$ 17.843 (31/12/2020 – R\$ 65.143), em referência a questões procedimentais, no que, com base nos pronunciamentos jurisdicionais sobre cada matéria, a análise dos advogados foi pela constituição da provisão integral do valor.

#### (c) Provisão para pré-pagamentos do crédito imobiliário com recursos do FGTS

Os financiamentos habitacionais concedidos com recursos do FGTS e contemplados com subsídio para redução da prestação, remuneram o Agente Financeiro com pagamento total ou parcial do subsídio pelo FGTS, conforme definição descrita na Resolução do Conselho Curador do FGTS nº 702/2012 e suas atualizações.

Estes valores são repassados ao Agente Financeiro CAIXA no ato da contratação, a fim de fazer frente a todo período da operação. Alterações no fluxo inicialmente contratado, tais como liquidação antecipada, amortização extraordinária, amortização extraordinária com redução de prazo, transferência ou redução do prazo contratado do financiamento, exigem que a CAIXA devolva ao FGTS parte da remuneração recebida, proporcionalmente.

Para fazer frente à essa devolução de recursos ao FGTS é feita a provisão para devolução de receitas em caso de pré-pagamentos. Os valores provisionados são apurados com base nos fluxos médios de devolução de remuneração ocorridos, e seus respectivos impactos sobre o saldo de valores de remuneração do Agente Financeiro CAIXA.

#### (d) Provisão para devolução de recursos do Programa Bolsa Família

No âmbito da prestação de serviços regida por instrumento contratual, o então Ministério do Desenvolvimento Social e Combate à Fome (MDS), atual Ministério da Cidadania (MC) apresentou à CAIXA solicitação de restituição de valores referentes ao Programa Bolsa Família. Importante destacar que em dezembro de 2021 foi sancionada a Lei nº 14.284 que institui o Auxílio Brasil, programa de assistência financeira a famílias de baixa renda que substituiu o Bolsa Família.

#### (e) Garantias financeiras prestadas

INDIVIDUAL / CONSOLIDADO				
Exposição de Garantias Prestadas				
Carteira	31/12/2021		31/12/2020	
	Exposição	Provisão	Exposição	Provisão
Repasses - Entidades	30.979.459	15.490	1.988.382	597
Cessão de Crédito	9.369.737	-	12.183.318	-
FIES	259.253	33.928	300.588	25.798
PAR (Recursos FAR) (1)	15.173	8.009	20.784	10.586
<b>Total</b>	<b>40.623.622</b>	<b>57.427</b>	<b>14.493.072</b>	<b>36.981</b>

(1) Valor referente ao Programa de Arrendamento Residencial com recursos do FAR, classificado no ativo como perdas associadas ao risco de crédito.

A metodologia de provisão para garantias prestadas no Repasse de recursos do FGTS às outras Entidades foi revista no 2º semestre de 2021, com valores ajustados à nova metodologia no último trimestre do ano.

Para os contratos do FIES, a provisão é realizada com base na avaliação de risco de crédito da concessão e evolui conforme estabelecido na Resolução CMN nº 2.682/99.

O subgrupo Cessão de Crédito é composto por contratos SBPE, cedidos com cláusula de Ajuste no Valor de Cessão, cujo pagamento mensal, sempre que necessário, é efetuado pela CAIXA, na qualidade de cedente. Para esse grupo a exposição é caracterizada pela obrigação assumida pela CAIXA de cobrir eventual falta de recursos para pagamento do fluxo.

O estoque remanescente da carteira do PAR com recursos do FAR é provisionado pelo rating, tendo em vista a evolução do atraso, conforme Resolução CMN nº 2.682/99.



(f) Passivos contingentes classificados como perdas possíveis

Com base na Resolução CMN nº 3.823/2009, as contingências classificadas como de perdas possíveis são dispensadas de constituição de provisão: INDIVIDUAL / CONSOLIDADO

Descrição	31/12/2021	31/12/2020
Demandas fiscais (f.1)	8.966.993	8.713.865
Demandas cíveis (f.2)	1.985.291	1.792.905
<b>Total</b>	<b>10.952.284</b>	<b>10.506.770</b>

(f.1) Demandas fiscais

A CAIXA mantém acompanhamento de processos fiscais administrativos e judiciais em que figura como polo passivo ou ativo e, sob o amparo dos pareceres de suas unidades jurídicas, classificou como risco de perda possível processos que totalizam o valor de R\$ 8.966.993 em 31/12/2021 (31/12/2020 – R\$ 8.713.865), dentre os quais se destacam as seguintes demandas em razão dos valores em discussão:

a) Autuação de PIS/PASEP, no valor histórico de R\$ 4.053.509, pautada em insuficiência no recolhimento relativo ao período de janeiro de 1991 a dezembro de 1995, ao tempo da vigência dos Decretos-Leis nº 2.445/1988 e nº 2.449/1988, que alteraram a sistemática de cálculo da contribuição, e suposta compensação indevida de recolhimentos efetuados a maior no período de janeiro de 1992 a maio de 1993. Para a propositura da ação, foi efetuado depósito em garantia daquele valor em 30/12/2010, o qual, atualizado pela SELIC, na forma da Lei nº 9.703/98, totaliza R\$ 7.788.008 em 31/12/2021 (31/12/2020 – R\$ 7.636.407). Em 08/04/2019, o Excelentíssimo Sr. Juiz da 9ª Vara Federal da Seção Judiciária de Brasília/DF julgou procedente o pedido formulado pela CAIXA para declarar a extinção do crédito tributário em destaque, tendo em vista a comprovação da suficiência dos recolhimentos à época realizados. Em 10/06/2019, a União interpôs recurso de apelação pleiteando a reforma da sentença em tela pelo Egrégio Tribunal Regional Federal da 1ª Região, estando os autos com remessa ordenada para regular apreciação e julgamento;

b) Autuação de PIS/PASEP, no valor total de R\$ 233.332 em 31/12/2021 (31/12/2020 – R\$ 230.992), com base em apuração de diferenças de base de cálculo para recolhimento relativo ao período de janeiro de 1996 a dezembro de 1998, e janeiro a outubro de 1999, decorrente de exclusão ou não inclusão de receitas e cômputo de despesas consideradas indevidas e não dedutíveis sobre a base tributável, respectivamente;

c) CSLL, no valor de R\$ 187.303 em 31/12/2021 (31/12/2020 – R\$ 184.523), relativa a crédito oriundo de pagamento a maior declarado em DIPJ e compensado no exercício de 2003, havendo discussão quanto a questões procedimentais; e

d) Autuação de ICMS (Fazenda do Estado de São Paulo), no valor total de R\$ 281.626 em 31/12/2021 (31/12/2020 – R\$ 278.208), sobre a qual se discute a exigência do imposto diante da não retenção e recolhimento na fonte sobre serviços enquadrados pela fiscalização no conceito de “comunicação” para efeitos fiscais e tributários, e ainda, a determinação de sujeição passiva por responsabilidade tributária mediante convênio do Conselho Nacional de Política Fazendária (CONFAZ).

As matérias relacionadas aos processos contingentes em discussão são monitoradas sob a perspectiva de eventual sedimentação ou modificação do cenário jurisprudencial, possibilitando a manutenção como decorrência de avaliação contínua por parte da CAIXA das respectivas classificações de risco.

(f.2) Demandas cíveis

A CAIXA, com base na opinião de seus advogados, acompanha sistematicamente todos os processos com classificação de risco possível ou remoto.

O montante de R\$ 1.985.291 em 31/12/2021 (31/12/2020 – R\$ 1.792.905) refere-se a uma ação popular com probabilidade de perda possível, na qual se alega a prática de ilegalidade por parte da CAIXA na gestão de recursos provenientes do PREVHAB, quando da sucessão do BNH.



## ANEXO C – DF 2021 - ITAÚ

### I - Provisões Cíveis e Trabalhistas e Outros Riscos

Segue abaixo a movimentação das provisões cíveis, trabalhistas e outros riscos:

	31/12/2021			
	Cíveis	Trabalhistas	Outros Riscos	Total
<b>Saldo Inicial - 01/01</b>	<b>3.511</b>	<b>8.015</b>	<b>1.483</b>	<b>13.009</b>
(-) Provisões Garantidas por Cláusula de Indenização (Nota 2.4.n)	(216)	(950)	-	(1.166)
<b>Subtotal</b>	<b>3.295</b>	<b>7.065</b>	<b>1.483</b>	<b>11.843</b>
Atualização / Encargos (Nota 23)	221	155	-	376
<b>Movimentação do Período Refletida no Resultado (Nota 23)</b>	<b>820</b>	<b>2.652</b>	<b>85</b>	<b>3.557</b>
Constituição	1.176	2.888	119	4.183
Reversão	(356)	(236)	(34)	(626)
Pagamento	(1.244)	(2.532)	(10)	(3.786)
<b>Subtotal</b>	<b>3.092</b>	<b>7.340</b>	<b>1.558</b>	<b>11.990</b>
(+) Provisões Garantidas por Cláusula de Indenização (Nota 2.4.n)	225	879	-	1.104
<b>Saldo Final</b>	<b>3.317</b>	<b>8.219</b>	<b>1.558</b>	<b>13.094</b>
<b>Circulante</b>	<b>1.266</b>	<b>2.528</b>	<b>1.558</b>	<b>5.352</b>
<b>Não Circulante</b>	<b>2.051</b>	<b>5.691</b>	<b>-</b>	<b>7.742</b>

	31/12/2020			
	Cíveis	Trabalhistas	Outros Riscos	Total
<b>Saldo Inicial - 01/01</b>	<b>3.634</b>	<b>8.579</b>	<b>976</b>	<b>13.189</b>
(-) Provisões Garantidas por Cláusula de Indenização (Nota 2.4.n)	(216)	(980)	-	(1.196)
<b>Subtotal</b>	<b>3.418</b>	<b>7.599</b>	<b>976</b>	<b>11.993</b>
Atualização / Encargos (Nota 23)	191	482	-	673
<b>Movimentação do Período Refletida no Resultado (Nota 23)</b>	<b>889</b>	<b>2.110</b>	<b>547</b>	<b>3.546</b>
Constituição	1.179	2.296	550	4.025
Reversão	(290)	(186)	(3)	(479)
Pagamento	(1.203)	(3.126)	(40)	(4.369)
<b>Subtotal</b>	<b>3.295</b>	<b>7.065</b>	<b>1.483</b>	<b>11.843</b>
(+) Provisões Garantidas por Cláusula de Indenização (Nota 2.4.n)	216	950	-	1.166
<b>Saldo Final</b>	<b>3.511</b>	<b>8.015</b>	<b>1.483</b>	<b>13.009</b>
<b>Circulante</b>	<b>1.254</b>	<b>3.125</b>	<b>1.483</b>	<b>5.862</b>
<b>Não Circulante</b>	<b>2.257</b>	<b>4.890</b>	<b>-</b>	<b>7.147</b>

### II- Provisões Fiscais e Previdenciárias

As provisões fiscais e previdenciárias correspondem ao valor principal dos tributos envolvidos em discussões fiscais administrativas ou judiciais, objeto de lançamento de ofício, acrescido de juros, multa e, encargos, quando aplicável.

Abaixo está demonstrada a movimentação das provisões:

	31/12/2021	31/12/2020
<b>Saldo Inicial - 01/01</b>	<b>6.810</b>	<b>8.266</b>
(-) Provisões Garantidas por Cláusula de Indenização (Nota 2.4 n)	(71)	(68)
<b>Subtotal</b>	<b>6.739</b>	<b>8.198</b>
Atualização / Encargos <sup>(*)</sup>	202	220
<b>Movimentação do Período Refletida no Resultado</b>	<b>8</b>	<b>56</b>
Constituição <sup>(*)</sup>	180	142
Reversão <sup>(*)</sup>	(172)	(86)
Pagamento	(523)	(1.735)
<b>Subtotal</b>	<b>6.426</b>	<b>6.739</b>
(+) Provisões Garantidas por Cláusula de Indenização (Nota 2.4 n)	72	71
<b>Saldo Final</b>	<b>6.498</b>	<b>6.810</b>
<b>Circulante</b>	<b>10</b>	<b>65</b>
<b>Não Circulante</b>	<b>6.488</b>	<b>6.745</b>

<sup>(\*)</sup> Os valores estão contemplados nas rubricas Despesas Tributárias, Despesas Gerais e Administrativas e em Imposto de Renda e Contribuição Social Correntes.

As principais discussões relativas a Ações Fiscais e Previdenciárias são descritas a seguir:

- INSS – Verbas não Remuneratórias – R\$ 1.823: defende-se a não incidência da contribuição previdenciária sobre verbas pagas a título de participação nos lucros. O saldo do depósito judicial totaliza R\$ 1.032;
- PIS e COFINS – Base de Cálculo – R\$ 641: defende-se a incidência de PIS e COFINS sobre o faturamento, devendo este ser entendido como a receita da venda de bens e serviços. O saldo do depósito judicial totaliza R\$ 628.

### III - Contingências não Provisionadas no Balanço

Os valores envolvidos em discussões administrativas e judiciais com risco estimado de perda possível não são objeto de provisão contábil e basicamente são compostas por:

#### Ações Cíveis e Trabalhistas

Nas Ações Cíveis de perda possível, o risco total estimado é de R\$ 4.903 (R\$ 4.470 em 31/12/2020), sendo que neste montante não existem valores decorrentes de participação em Entidades Controladas em Conjunto.

Para as Ações trabalhistas de perda possível, o risco estimado é de R\$ 448 (R\$ 389 em 31/12/2020).

#### Ações Fiscais e Previdenciárias

As Ações Fiscais e Previdenciárias de perda possível totalizam R\$ 35.855 (R\$ 31.330 em 31/12/2020), sendo as principais discussões descritas a seguir:

- INSS – Verbas não Remuneratórias – R\$ 8.119: defende-se a não incidência da contribuição sobre verbas, não remuneratórias, dentre as quais, destacam-se: participação nos lucros e plano para outorga de opções de ações;
- IRPJ, CSLL, PIS e COFINS – Despesas de Captação – R\$ 4.923: discute-se a dedutibilidade de despesas de captação (DI), relativas a recursos que foram capitalizados entre as empresas do Grupo;
- ISS – Atividades Bancárias – R\$ 4.680: discute-se a incidência e/ou local do recolhimento de ISS para determinadas receitas bancárias;
- IRPJ e CSLL – Ágio – Dedução – R\$ 3.479: discute-se a dedutibilidade do ágio com expectativa de rentabilidade futura na aquisição de investimentos;
- PIS e COFINS - Estorno da Receita de Superveniência de Depreciação – R\$ 2.428: discute-se o tratamento contábil e fiscal do PIS e da COFINS na liquidação das operações de arrendamento mercantil;
- IRPJ, CSLL, PIS e COFINS – Indeferimento de Pedido de Compensação – R\$ 1.517: casos em que são apreciadas a liquidez e a certeza do crédito compensado;
- IRPJ e CSLL – Glosa de Prejuízos – R\$ 1.233: discute-se o montante do prejuízo fiscal (IRPJ) e/ou base negativa de CSLL utilizados pela Receita Federal na lavratura de autos de infração, que ainda estão pendentes de decisão definitiva;



**c) Contas a Receber – Reembolso de Provisões**

O saldo de valores a receber relativo a reembolso de provisões totaliza R\$ 888 (R\$ 919 em 31/12/2020) (Nota 18a) e decorre, basicamente, da garantia estabelecida em 1997, no processo de privatização do Banco Banerj S.A., ocorrido em 1997, quando o Estado do Rio de Janeiro constituiu um fundo para garantir a recomposição patrimonial em Provisões Cíveis, Trabalhistas e Fiscais e Previdenciárias.

**d) Garantias de Contingências, Provisões e Obrigações Legais**

As garantias relativas a discussões que envolvem o ITAÚ UNIBANCO HOLDING são compostas, basicamente por:

	31/12/2021				31/12/2020
	Cíveis	Trabalhistas	Tributários	Total	Total
Depósitos em Garantia (Nota 18a)	1.427	1.990	8.847	<b>12.264</b>	12.693
Cotas de Fundos de Investimento	408	204	78	<b>690</b>	987
Fiança	70	48	3.997	<b>4.115</b>	4.012
Seguro Garantia	1.710	1.325	15.736	<b>18.771</b>	18.402
Garantia por Títulos Públicos	7	-	235	<b>242</b>	249
<b>Total</b>	<b>3.622</b>	<b>3.567</b>	<b>28.893</b>	<b>36.082</b>	<b>36.343</b>



## ANEXO D – DF 2021 - BRADESCO

### Demonstrações Contábeis Consolidadas

#### Notas Explicativas

- PIS e Cofins – R\$ 527.970 mil (Em dezembro de 2020 – R\$ 415.785 mil): pleiteia assegurar as empresas o direito de recolher as contribuições ao PIS e a Cofins pelo regime cumulativo (alíquota 3,65% sobre vendas de mercadorias/prestação de serviços);
- INSS - Contribuição ao SAT – R\$ 450.289 mil (Em dezembro de 2020 – R\$ 440.524 mil): em ação ordinária movida pela Federação Nacional dos Bancos - Febraban, desde abril de 2007, em nome de seus associados, é questionado o enquadramento dos bancos no grau de risco mais elevado, no que tange ao Risco de Acidentes de Trabalho - RAT, o que acabou por elevar a alíquota da respectiva contribuição de 1% para 3%, conforme Decreto nº 6.042/07; e
- INSS de Autônomos – R\$ 343.896 mil (Em dezembro de 2020 – R\$ 333.852 mil): discute a incidência da contribuição previdenciária sobre as remunerações pagas aos prestadores de serviços autônomos, instituída pela Lei Complementar nº 84/96, e regulamentações/alterações posteriores à alíquota de 20,0% e adicional de 2,5%, sob o argumento de que os serviços não são prestados às seguradoras, mas aos segurados, estando, dessa forma, fora do campo de incidência da contribuição prevista no inciso I, artigo 22, da Lei nº 8.212/91, com nova redação contida na Lei nº 9.876/99.

Em geral, as provisões referentes às ações judiciais são consideradas de longo prazo, devido à imprevisibilidade do tempo de duração dos processos no sistema judiciário brasileiro, razão pela qual não foi divulgada a estimativa com relação ao ano específico em que essas ações judiciais serão encerradas.

#### IV - Provisões segregadas por natureza

	R\$ mil	
	Em 31 de dezembro de 2021	Em 31 de dezembro de 2020
Processos trabalhistas	6.729.107	6.890.498
Processos cíveis	9.178.471	9.092.421
Provisão para riscos fiscais	8.072.037	8.271.112
<b>Total (Nota 22b)</b>	<b>23.979.615</b>	<b>24.254.031</b>

#### V - Movimentação das provisões – Despesas de provisões

	R\$ mil		
	Trabalhistas	Cíveis	Fiscais
<b>Saldo em 31 de dezembro de 2020</b>	<b>6.890.498</b>	<b>9.092.421</b>	<b>8.271.112</b>
Atualização monetária	799.803	484.516	176.903
Constituições líquidas de (reversões e baixas)	1.044.511	1.734.207	(351.476)
Pagamentos	(2.005.705)	(2.132.673)	(24.502)
<b>Saldo em 31 de dezembro de 2021</b>	<b>6.729.107</b>	<b>9.178.471</b>	<b>8.072.037</b>

#### c) Passivos contingentes classificados como perdas possíveis

A Organização mantém um sistema de acompanhamento para todos os processos administrativos e judiciais em que a Instituição figura como “autora” ou “ré” e, amparada na opinião dos assessores jurídicos, classifica as ações de acordo com a expectativa de insucesso. Periodicamente são realizadas análises sobre as tendências jurisprudenciais e efetivada, se necessária, a reclassificação dos riscos desses processos. Neste contexto, os processos contingentes avaliados como de risco de perda possível não são reconhecidos contabilmente e totalizaram, em 31 de dezembro de 2021, R\$ 7.979.276 mil (Em dezembro de 2020 – R\$ 7.222.015 mil) para os processos cíveis e R\$ 37.556.235 mil (Em dezembro de 2020 – R\$ 35.761.167 mil) para os processos fiscais.



# ANEXO E – DF 2021 - SANTANDER

<b>Relatório da Administração</b>	<b>Relatório dos Auditores Independentes</b>	<b>Demonstrações Financeiras</b>	<b>Notas Explicativas</b>	<b>Declarações dos Diretores</b>
-----------------------------------	--	----------------------------------	---------------------------	----------------------------------

\*Valores expressos em milhares, exceto quando indicado.

## a) Provisão para Garantias Financeiras Prestadas

A classificação das operações de garantias prestadas para constituição de provisionamento é baseada na estimativa do risco envolvido. Decorre do processo de avaliação da qualidade dos clientes e operações, por modelo estatístico baseado em informações quantitativas e qualitativas ou por um analista de crédito especializado, que permite classificá-las em função de sua probabilidade de default, baseado em variáveis objetivas internas e de mercado (bureaus), previamente identificadas como preditivas da probabilidade de default. Após essa avaliação, as operações são classificadas de acordo com os ratings de provisionamento, tendo como referência a Resolução CMN nº 2.682/1999. Através desta análise, são registrados os valores de provisão para a cobertura de cada operação, considerando o tipo da garantia prestada, de acordo com o requerido na Resolução CMN nº 4.512/2016.

Tipo de Garantia Financeira	Banco/Consolidado			
	31/12/2021		31/12/2020	
	Saldo Garantias Prestadas	Provisão	Saldo Garantias Prestadas	Provisão
Vinculadas ao Comércio Internacional de Mercadorias	6.244.755	28.506	1.813.620	4.121
Vinculadas a Licitações, Leilões, Prestação de Serviços ou Execução de Obras	6.796.175	4.198	5.602.995	5.403
Vinculadas ao Fornecimento de Mercadorias	1.698.518	2.442	1.361.792	1.846
Aval ou Fiança em Processos Judiciais e Administrativos de Natureza Fiscal	11.823.964	243.235	12.082.480	175.443
Outros Avais	2.748.497	1.897	335.281	1.689
Outras Fianças Bancárias	19.525.773	36.489	16.532.462	33.055
Outras Garantias Financeiras Prestadas	88.388	7.960	5.047.032	33.622
<b>Total</b>	<b>48.926.070</b>	<b>324.727</b>	<b>42.775.662</b>	<b>255.179</b>

## Movimentação da Provisão para Garantias Financeiras Prestadas

	Banco/Consolidado	
	01/01 a 31/12/2021	01/01 a 31/12/2020
<b>Saldo Inicial</b>	<b>255.179</b>	<b>166.105</b>
Constituição	95.431	99.035
Reversão (1)	(25.883)	(9.961)
<b>Saldo</b>	<b>324.727</b>	<b>255.179</b>

(1) Corresponde a fianças honradas, mudança de rating ou provisão constituída na linha de Provisão para Perdas Esperadas Associadas ao Risco de Crédito.

## 20. Provisões, Passivos Contingentes, Ativos Contingentes e Obrigações Legais - Fiscais e Previdenciárias

### a) Ativos Contingentes

No Banco e no Consolidado, em 31 de dezembro de 2021 e 31 de dezembro de 2020, não foram reconhecidos contabilmente ativos contingentes.

### b) Saldos Patrimoniais das Provisões para Processos Judiciais e Administrativos e Obrigações Legais por Natureza

	Banco		Consolidado	
	31/12/2021	31/12/2020	31/12/2021	31/12/2020
<b>Provisão para Riscos Fiscais e Obrigações Legais (Nota 19)</b>	<b>4.312.234</b>	<b>4.249.744</b>	<b>6.748.684</b>	<b>6.707.292</b>
<b>Provisão para Processos Judiciais e Administrativos - Ações Trabalhistas e Cíveis (Nota 19)</b>	<b>5.033.675</b>	<b>5.921.882</b>	<b>5.325.716</b>	<b>6.342.280</b>
Ações Trabalhistas	1.941.169	2.656.098	2.084.247	2.900.835
Ações Cíveis	3.092.507	3.265.784	3.241.469	3.441.445
<b>Total</b>	<b>9.345.909</b>	<b>10.171.626</b>	<b>12.074.400</b>	<b>13.049.572</b>

<b>Relatório da Administração</b>	<b>Relatório dos Auditores Independentes</b>	<b>Demonstrações Financeiras</b>	<b>Notas Explicativas</b>	<b>Declarações dos Diretores</b>
-----------------------------------	--	----------------------------------	---------------------------	----------------------------------

\*Valores expressos em milhares, exceto quando indicado.

#### **e) Passivos Contingentes Fiscais e Previdenciárias, Trabalhistas e Cíveis Classificados como Risco de Perda Possível**

São processos judiciais e administrativos de natureza fiscal e previdenciária, trabalhista e cível classificados, com base na opinião dos assessores jurídicos, como risco de perda possível, não sendo, portanto, provisionados.

As ações de natureza fiscal com classificação de perda possível, totalizaram R\$29.726 milhões no Consolidado, sendo os principais processos os seguintes:

**INSS sobre Participação nos Lucros ou Resultados (PLR)** - o Banco e as empresas controladas possuem processos judiciais e administrativos decorrentes de questionamentos das autoridades fiscais, a respeito da cobrança de contribuição previdenciária sobre os pagamentos efetuados a título de participação nos lucros e resultados. Em 31 de dezembro de 2021, o valor era de aproximadamente R\$7.341 milhões.

**Imposto sobre Serviços (ISS) - Instituições Financeiras** - o Banco Santander e as empresas controladas discutem administrativa e judicialmente a exigência, por vários municípios, do pagamento de ISS sobre diversas receitas decorrentes de operações que usualmente não se classificam como prestação de serviços. Em 31 de dezembro de 2021, o valor era de aproximadamente R\$4.146 milhões.

**Compensação Não Homologada** - o Banco e suas coligadas discutem administrativa e judicialmente com a Receita Federal a não homologação de compensações de tributos com créditos decorrentes de pagamento a maior ou indevido. Em 31 de dezembro de 2021, o valor era de aproximadamente R\$5.355 milhões.

**Amortização do Ágio do Banco Real** - a Receita Federal do Brasil emitiu auto de infração contra o Banco para exigir os pagamentos de IRPJ e CSLL, incluindo os encargos moratórios, referentes ao período-base de 2009. As Autoridades Fiscais consideraram que o ágio referente à aquisição do Banco Real, amortizado contabilmente antes da sua incorporação, não poderia ser deduzido pelo Banco Santander para fins fiscais. O auto de infração foi devidamente impugnado e atualmente, aguardamos julgamento perante o CARF. Em 31 de dezembro de 2021, o valor era de aproximadamente R\$1.466 milhões.

**Perdas em Operações de Crédito** - o Banco e as empresas controladas contestaram os lançamentos fiscais emitidos pela Receita Federal do Brasil alegando a dedução indevida de perdas em operações de crédito das bases de cálculo do IRPJ e da CSLL por supostamente não atenderem às exigências das leis aplicáveis. Em 31 de dezembro de 2021, o valor era de aproximadamente R\$1.176 milhões.

**Utilização de Prejuízo Fiscal e de Base Negativa da CSLL** - Autos de infração lavrados pela Receita Federal do Brasil no exercício de 2009 por supostas compensações indevidas de prejuízo fiscal e de base negativa da CSLL, como consequência de autuações fiscais lavradas em períodos anteriores. Aguarda-se julgamento na esfera administrativa. Em 31 de dezembro de 2021, o valor era de aproximadamente R\$1.093 milhões.

**Amortização do Ágio do Banco Sudameris** - as autoridades fiscais lavraram autos de infração para exigir os pagamentos de IRPJ e CSLL, incluindo os encargos moratórios, referentes à dedução fiscal da amortização do ágio pago na aquisição do Banco Sudameris, referentes ao período base de 2007 a 2012. O Banco Santander apresentou as respectivas defesas administrativas, as quais foram julgadas desfavoravelmente. Atualmente, os processos aguardam julgamento no CARF. Em 31 de dezembro de 2021, o valor era de aproximadamente R\$659 milhões.

**IRPJ e CSLL - Ganho de Capital** - a Secretaria da Receita Federal do Brasil emitiu um auto de infração contra a Santander Seguros (sucessora legal da ABN AMRO Brasil Dois Participações S.A. (AAB Dois Par) cobrando imposto de renda e contribuição social relacionados ao exercício fiscal de 2005. A Receita Federal do Brasil alega que o ganho de capital na venda das ações da Real Seguros S.A. e da Real Vida e Previdência S.A pela AAB Dois Par deve ser tributado a uma alíquota de 34,0% ao invés de 15,0%. O lançamento foi contestado administrativamente com base no entendimento que o tratamento fiscal adotado na transação estava em conformidade com a legislação tributária vigente e o ganho de capital foi devidamente tributado. O processo administrativo encerrou desfavoravelmente à Companhia. Em julho de 2020, a Companhia ajuizou ação visando anular o débito. A ação judicial aguarda julgamento. O Banco Santander é responsável por qualquer resultado adverso nesse processo como ex-controlador da Zurich Santander Brasil Seguros e Previdência S.A. Em 31 de dezembro de 2021, o valor era de aproximadamente R\$496 milhões.

As ações de natureza trabalhista com classificação de perda possível totalizaram R\$267 milhões no Consolidado, excluindo o processo abaixo:

**Reajuste das Complementações de Aposentadoria do Banesprev pelo IGPDI** - ação ajuizada em 2002 na Justiça Federal pela Associação de Funcionários Aposentados do Banco do Estado de São Paulo requerendo o reajuste da complementação de aposentadoria pelo IGPDI para aposentados do Banespa que tenham sido admitidos até 22 de maio de 1975. A sentença deferiu a correção, mas apenas nos períodos em que não houve a aplicação de nenhuma outra forma de reajuste. O Banco e o Banesprev recorreram dessa decisão e os Recursos ainda estão pendentes de julgamento. Em Execução Provisória foram apresentados cálculos

<b>Relatório da Administração</b>	<b>Relatório dos Auditores Independentes</b>	<b>Demonstrações Financeiras</b>	<b>Notas Explicativas</b>	<b>Declarações dos Diretores</b>
---------------------------------------	--	--------------------------------------	---------------------------	--------------------------------------

*\*Valores expressos em milhares, exceto quando indicado.*

pelo Banco e Banesprev em razão da exclusão de participantes que, entre outros motivos, constam como autores em outras ações ou já tiveram algum tipo de reajuste. O valor envolvido não é divulgado em razão da atual fase processual do caso e de potencialmente poder afetar o andamento da ação.

Os passivos relacionados a ações cíveis com risco de perda possível totalizaram R\$2.380 milhões no Consolidado, tendo como principais processos:

**Ação Indenizatória Oriunda do Banco Bandepe** - relacionada ao contrato de mútuo em fase de recurso pelo Superior Tribunal de Justiça (STJ).

**Ação Indenizatória Referente à de Serviços de Custódia** - prestados pelo Banco Santander em fase inicial e ainda sem sentença proferida.

**Ação Oriunda de Disputa Contratual** - na aquisição do Banco Geral do Comércio S.A. em fase de recurso pelo Tribunal de Justiça do Estado de São Paulo (TJSP).

#### **f) Outras Ações Judiciais de Responsabilidade de Ex-Controladores**

Referem-se a ações de natureza cível no montante de R\$496 (31/12/2020 – R\$496) no Banco e no Consolidado, registrados em outros passivos (Nota 19) de responsabilidade dos ex-controladores de Bancos e empresas adquiridas. Com base nos contratos firmados, estas ações possuem garantias de ressarcimento integral por parte dos ex-controladores, cujos respectivos direitos foram contabilizados em outros ativos (Nota 12).