

PROJETO DE TRABALHO DE CONCLUSÃO DE CURSO

BENJAMIN CZERNY ZUGMAN
GUSTAVO PERLIN FIOR
LEONARDO BAGGIO DE MELLO
THIAGO XIMENES

**PREVISÃO DE FALÊNCIAS DE EMPRESAS BRASILEIRAS UTILIZANDO
VARIÁVEIS MACROECONÔMICAS E MODELOS DE INTELIGÊNCIA
ARTIFICIAL**

Curitiba

2023

BENJAMIN CZERNY ZUGMAN
GUSTAVO PERLIN FIOR
LEONARDO BAGGIO DE MELLO
THIAGO XIMENES

**PREVISÃO DE FALÊNCIAS DE EMPRESAS BRASILEIRAS UTILIZANDO
VARIÁVEIS MACROECONÔMICAS E MODELOS DE INTELIGÊNCIA
ARTIFICIAL**

Trabalho de graduação apresentado
ao Curso de Administração da FAE
Centro Universitário.

Área de conhecimento: Economia,
Finanças e Modelagem.

Orientador: Eduardo da Silva Mattos.

CURITIBA

2023

RESUMO

Este artigo tem como objetivo analisar e prever o número de falências de empresas brasileiras com base na utilização de modelos de inteligência artificial, como **Baggin**, Random Forest e **XG Boost**. Vale ressaltar que houve uma alteração jurídica em 2005 criando a lei número 11.101, lei que regula a **recuperação judicial**, por esse motivo não utilizamos dados anteriores a este período.

A base de dados foi construída ao utilizar dados a partir de janeiro de 2006 até junho de 2023. **Foram utilizadas 7** variáveis macroeconômicas para treinamento dos modelos, são elas: taxa de câmbio (Bacen), taxa de capital de giro PJ (Bacen), taxa de juros (Selic/Bacen), Expectativa do empresário (ICE/FGV), PIB real (PIB descontado da inflação/Bacen), M1 (liquidez/Bacen) e Ibovespa. Com relação a variável explicada será o número de falências requeridas no período, pois as falências decretadas possuem inconstâncias do judiciário.

O XGBoost foi o modelo que apresentou o melhor resultado. O resultado **foi o menor** Mean Squared Error (MSE) entre todos os modelos, ao finalizar com um valor de **143.7779** no modelo que teve como variável explicativa representando Taxa Média de Capital de Giro para Pessoas Jurídicas (CGPJ). Além disso, vale destacar que os outros modelos (Baggin e Random Forest) tiveram desempenhos parecidos e obtiveram um desempenho melhor que o XGBoost no modelo com taxa de juros (SELIC). Conclui-se que os modelos de inteligência artificial superaram o modelo tradicional (ARIMA) em todos os testes realizados.

Palavras-chave: Falência. Fatores Macroeconômicos. Modelagem.

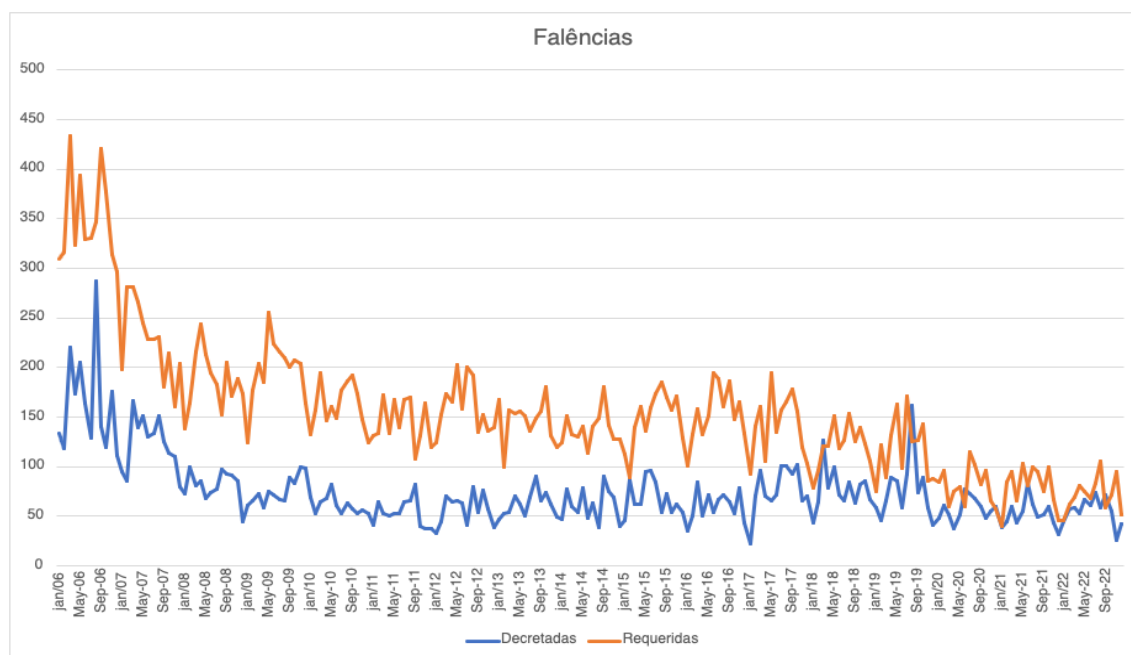
SUMÁRIO

1 INTRODUÇÃO.....	5
2 REVISÃO DE LITERATURA.....	7
2.1 DADOS.....	11
2.1.1 Taxa de câmbio.....	11
2.1.2 Produto Interno Bruto.....	12
2.1.3 Taxas de Juros.....	12
2.1.4 Liquidez.....	12
2.1.5 Expectativas.....	12
2.1.6 Ibovespa.....	12
2.1.7 Inflação.....	12
3 PROCEDIMENTOS METODOLÓGICOS.....	14
4 RESULTADOS.....	15
5 CONCLUSÃO.....	20
REFERÊNCIAS.....	21

1 INTRODUÇÃO

Esse artigo tem como objetivo analisar e prever a quantidade de falências requeridas por empresas brasileiras e como as variáveis macroeconômicas impactam neste número. Em 9 de Fevereiro de 2005, houve a promulgação da Lei número 11.101/2005, a qual ficou conhecida como lei da recuperação judicial (BRASIL, 2005). Esta lei firmou as bases legais para processos de recuperação de empresas no Brasil. Portanto, esta análise terá como início o ano de 2006. Ainda que os números de requerimentos tenham diminuído significativamente após esta data, como demonstrado pelo gráfico abaixo, ainda existem variações significativas e relações nebulosas entre variáveis macroeconômicas e seus impactos nos requerimentos de falências.

Figura 1- Falências decretadas e requeridas no Brasil no período de 2006 a 2022.



Fonte: Serasa Experian (2023).

Para exemplificar como essas relações ainda são, de certa forma, parcialmente incompreendidas, após a crise da pandemia de Covid-19 (2020), diversos especialistas em falências e recuperação judicial apontavam que os números iriam aumentar significativamente. Um dos estudos realizado pela Secretaria de Política Econômica (SPE) do Ministério da Economia, citado em

notícia no site G1, levantou que aproximadamente cerca de 3.500 companhias iriam pedir recuperação judicial ou entrar em falência nos próximos meses (FLOR, 2020).

Entretanto, percebe-se que houve uma diminuição no número de falências requeridas e de recuperação judicial requeridas no período analisado. No ano de 2019, houve 1417 falências requeridas, em 2020 o número diminuiu para 972, em 2021 houve mais uma diminuição para 950 e em 2022 uma nova queda para 866 empresas requerendo falência. Desse modo, percebe-se que houve uma grande distorção entre o que era indicado por especialistas e o que de fato aconteceu, ilustrando a dificuldade de antecipar cenários futuros. Deste modo, a utilização de novos modelos para entender como os fatores macroeconômicos afetam e criam uma crise empresarial é de grande interesse.

Luiz Rabi, economista do Serasa Experian, relatou para entrevista na revista Folha de Pernambuco ("Folha de Pernambuco", 2023), que a elevação de empresas inadimplentes no ano de 2022 foi em razão de um juro alto e crédito escasso. Atualmente, é notório que a economia mundial passa por um período de juros relativamente altos, onde a taxa base brasileira (taxa SELIC) saiu de 2% para 13,75% e a taxa base americana, que aumentou de 0,25% para 5,25% (dados extraídos, respectivamente, do BACEN e da plataforma TradingEconomics).

Desse modo, espera-se encontrar resultados para entender se esse componente de juros altos é um fator determinante para um maior número de empresas requerendo falência ou se outros componentes macroeconômicos têm mais relevância para este número. Com o objetivo de obter um panorama mais amplo sobre a questão dos juros e dessas outras variáveis, diferentes simulações foram realizadas, ao utilizar uma gama de dados macroeconômicos diferentes.

A primeira variável é a própria taxa SELIC e a taxa média mensal de juros das operações de crédito com recursos livres para pessoas jurídicas. Além da taxa base de juros, é possível ter como variáveis explicativas - explicando o contexto macroeconômico para a previsão - o PIB real da economia brasileira, a liquidez na economia (sendo contabilizada como o agregado monetário M1), a taxa de câmbio entre Real e Dólar americano (BRL x USD), um índice de

expectativas (ICE - Índice de Confiança Empresarial - FGV/IBRE) e, por fim, o índice IBOVESPA representando o mercado de ações brasileiro.

Vale ressaltar que a taxa de desemprego não foi inserida como variável neste estudo, pois houve uma alteração metodológica no índice em 2016, visto que era calculada pela Pesquisa Mensal de Emprego (PME), a qual foi descontinuada e substituída pela Pesquisa Nacional por Amostragem de Domicílio Contínua (PNAD). Além dessa variável, não foi utilizado nenhum índice de inflação, como o Índice de Preços ao Consumidor Amplo (IPCA) diretamente, visto que ele é utilizado indiretamente no cálculo do PIB Real e da liquidez real da economia (M1).

O trabalho divide-se em 5 capítulos: introdução, revisão de literatura, metodologia, resultados e conclusão. Na segunda seção, apresenta-se os estudos de maior relevância, tanto no Brasil quanto fora. Na seção seguinte, apresenta-se a explicação do modelo de aprendizado de máquina. Na quarta seção, são apresentados resultados do modelo e, por fim, as conclusões pertinentes.

2 REVISÃO DA LITERATURA

Um dos estudos mais relevantes sobre o impacto de fatores macroeconômicos na previsão de falências de empresas foi escrito por Altman (1968), no qual o autor desenvolve um conceito chamado de Z-score. A amostra de seu estudo utiliza uma pequena quantidade de empresas, 66, nos Estados Unidos entre os anos de 1946 e 1965. O modelo utilizava 5 variáveis:

$X_1 = \text{Capital de giro} / \text{Total de ativos}$

$X_2 = \text{Lucros} / \text{Total de ativos}$

$X_3 = \text{Ebit} / \text{Total de ativos}$

$X_4 = \text{Valor de mercado} / \text{Total de passivo}$

$X_5 = \text{Faturamento} / \text{Total de ativos}$

O modelo seria para prever uma falência com antecedência de 5 anos, a assertividade do modelo era de 95% para um ano de antecedência, 72% para dois anos e 48% para três, logo a assertividade para um prazo maior era muito baixa. O autor identificou que fatores externos também impactaram nos números de inadimplência.

Devido a isso, escreveu sobre as variáveis externas nos impactos na taxa de falência, Altman (1983). O autor realizou pesquisas analisando como 5 fatores macroeconômicos afetam a taxa de falência de empresas nos Estados Unidos agora entre os períodos de 1951 e 1978. Sendo as 5 variáveis escolhidas, crescimento econômico, condições de mercado de crédito, atividade de mercado de capitais, características de empresa de mercado e níveis de preço.

As conclusões de Altman (1983) foram que uma queda no PIB, uma menor oferta da moeda (redução de M2) e uma performance ruim do S&P500 (menor expectativa de crescimento dos investidores) levariam um maior número de empresas à falência.

Platt e Platt (1994) utilizaram dados de 1969 a 1982, nos Estados Unidos, para identificar quais fatores afetam a taxa de falência empresarial. Desfrutaram-se de indicadores econômicos com diferenças interestaduais, ao mesclar os modelos utilizados e séries de tempo. Os autores testaram taxa de lucro dos empresários, salário real, taxa de novas empresas e taxa de desemprego. Todos os fatores apresentaram impactos significativos, sendo lucros e desemprego correlacionados negativamente e, novas empresas e salário real, positivamente, com a taxa de falências.

Everett e Watson (1998), na Austrália, realizaram dois testes, analisando o impacto dos riscos não sistemáticos e dos riscos sistemáticos. Na primeira análise olharam para dados da empresa como idade da empresa e período. Enquanto na análise dos riscos externos analisaram a relação entre a taxa de falência, taxa de juros e taxa de desemprego.

Liu e Wilson (2002), no Reino Unido, coletaram dados de 1968 a 1998. No estudo, relataram que o número de falências (quebra estrutural) diminuiu no curto prazo, principalmente em razão da criação da lei da insolvência de 1986 (encontrar fonte).

Em um estudo mais recente, Zhang, Bessler e Leatham (2013), ao utilizar VAR e dados de empresas americanas, de 1980 a 2004, usufruíram de quatro variáveis econômicas: lucro corporativo agregado, inflação, taxa de juros e atividades no mercado de capitais. Concluíram que existe relação positiva entre taxa de juros e número de empresas em falência. As demais variáveis têm uma relação negativa com o número de empresas decretando fechamento.

Os pesquisadores Salman, Fuchs e Zampatti (2015) coletaram dados de empresas pequenas na Suécia e concluíram que a variação de câmbio só possui relação positiva em empresas de grande porte. Já as empresas de pequeno e médio porte não sofrem alteração nas taxas de falência.

No Brasil, Contador (1985) testou três modelos, cada um dos modelos possuía uma variável dependente, sendo as taxas de títulos protestados, falências requeridas e a taxa de insolvência. Os resultados nos modelos de falências requeridas e protestadas não foram satisfatórios, visto que houve problemas em capturar os dados de mão de obra e insumos. Contudo, o estudo concluiu que a política macroeconômica mais eficiente para uma boa saúde financeira da empresa é a com menor nível de inflação e a manutenção da taxa de juros em nível baixo.

Araújo e Funchal (2009), no Brasil, avaliaram os efeitos da lei n.11.101/2005. Essa análise testou algumas variáveis de crédito segmentada por setores para realizar as regressões controladas por PIB e câmbio (defasagem). No estudo, foi encontrado uma redução considerável no número de falências.

Por fim, um estudo mais recente, Mattos (2023) utilizou as variáveis câmbio, PIB, juros, liquidez e expectativas. O modelo usado foram o VECM e VAR, além da suavização de dados com médias trimestrais e defasagens. O estudo apontou correlação negativa entre câmbio e Liquidez com o número de empresas buscando insolvência, enquanto as demais variáveis tiveram uma correlação positiva com o número de empresas em insolvência.

Ao adentrar estudos que utilizaram aprendizado de máquina para prever insolvência, Barboza, Kimura e Altman (2017) realizaram algo similar para o contexto de uma empresa. Os pesquisadores utilizaram de diversos indicadores financeiros como Return on Equity (ROE), ativos, receitas e patrimônio líquido. Os modelos utilizados foram o Random Forest, SVM - Support Vector Machines (Linear e RBF), Boosting, Bagging, Neural Networks, Logit e MDA - Model Driven Architecture. No estudo, foi concluído que os modelos tradicionais possuíram uma acurácia menor que nos modelos de aprendizado de máquina, sendo respectivamente as taxas entre 52% e 77% para os modelos tradicionais e entre 71% e 87% para os modelos de inteligência artificial.

Almeida, Sales e Nunes (2023) realizaram um estudo realizando uma predição de insolvência de empresas por meio de inteligência artificial. No

estudo, ao invés de variáveis macroeconômicas foram utilizados indicadores financeiros da empresa como liquidez geral, liquidez seca, ativo circulante e não circulante, EBITDA, alavancagem financeira e patrimônio líquido. Foi concluído que a acurácia dos modelos de aprendizado de máquina foi viável e verificou-se um percentual de acertos de 70% a 85%.

Abaixo segue um quadro, com o resumo dos estudos já realizados sobre o tema e as variáveis que foram analisadas em cada um deles.

QUADRO 1- Resumo dos estudos já realizados sobre o tema, em ordem de publicação.

<u>Artigo</u>	<u>Variável dependente</u>	<u>Variáveis explicativas</u>
Altman (1968)	Taxa de falência	Crescimento econômico, mercado de crédito, mercado de capitais e níveis de preço.
Contador (1985)	Taxa de falência	Juros, gastos da união, exportações, salário-mínimo e inflação.
Platt e Platt (1994)	Taxa de falência	Lucro dos empresários, salário, novas empresas e desemprego.
Everett e Watson (1998)	Taxa de falência	Desemprego, falências, vendas no varejo, inflação e taxa de juros.
Liu e Wilson (2002)	Taxa de falência	Crédito, inflação, lucros corporativos e taxa de juros.
Araújo e Funchal (2009)	Falências requeridas e decretadas	PIB, câmbio e variáveis crédito.
Zhang, Bessler e Leatham (2013)	Número de falências	Taxa de juros, S&P 500 e inflação.
Salman, Fuchs e Zampatti (2015)	Número de falências	Câmbio, produtividade e taxa de juros.

Barboza, Kimura e Altman (2017)	Taxa de falência	Ativos, receitas, alavancagem, patrimônio líquido e retorno.
Mattos (2023)	Taxa de Falências e recuperação judicial.	Câmbio, PIB, juros, liquidez e expectativas.
Almeida, Sales e Nunes (2023)	Taxa de falência	EBITDA, alavancagem, ativos, liquidez geral e liquidez seca.

Fonte: Os autores (2023).

2.1 DADOS

Os dados sobre falências requeridas foram extraídos da base fornecida pelo Serasa Experian. O período de coleta de dados se deu a partir de janeiro de 2006 até julho de 2023. Os anos anteriores a 2006 foram excluídos em razão da alteração de legislação a respeito de empresas inadimplentes, como já citado anteriormente.

As variáveis macroeconômicas utilizadas foram retiradas do Sistema Gerenciador de Séries Temporais do Banco Central do Brasil (SGS - Bacen), com exceção das variáveis de expectativas, retirada do Instituto Brasileiro de Economia da Fundação Getúlio Vargas (FGV IBRE) e o índice Ibovespa (IBOV) extraído da Bolsa de Valores do Brasil (B3). Como as variáveis explicativas passaram por um processo de defasagem para melhor compreensão dos seus efeitos, o período de coleta delas iniciou-se em 2005.

Utilizaram-se 8 variáveis macroeconômicas no estudo, sendo o IPCA indiretamente, visto que esta foi usada para cálculo do PIB real e liquidez (M1) real. As outras 7 variáveis escolhidas foram: taxa de câmbio, PIB, taxa de juros, taxa de juros para pessoas jurídicas, liquidez, expectativas (ICE) e Ibovespa.

2.1.1 Taxa de câmbio

É evidente que a taxa de câmbio pode afetar a estabilidade de empresas de maneira positiva ou negativa e que sua variação tem impactos consideráveis e variáveis dependendo do setor de atuação da empresa.

2.1.2 Produto Interno Bruto

Utilizou-se, primeiramente, o PIB nominal mensal, deflacionado pela variação do IPCA e utilizando o primeiro período como base. Desta forma, a variável final é o PIB real mensal.

2.1.3 Taxas de Juros

Duas versões para essa variável foram testadas, sendo a taxa média mensal para operações de crédito de pessoas jurídicas e a taxa média de juros (SELIC) em porcentagem anual. A primeira versão possui uma vantagem, principalmente, em razão da taxa estar mais próxima do que realmente acontece na prática nas operações das companhias.

2.1.4 Liquidez

Como a variável de taxa de juros representa o custo do crédito, utilizou-se a variável liquidez (M1) para representar as disponibilidades ou restrições de recursos para empréstimos. Dessa maneira, foi utilizado o agregado monetário M1, deflacionado pela variação do IPCA.

2.1.5 Expectativas

O Índice de Confiança Empresarial (ICE) agrega os dados das sondagens da Indústria de Transformação, Serviços, Comércio e Construção. Ele é calculado pelo Instituto Brasileiro de Economia da Fundação Getúlio Vargas (IBRE/FGV).

2.1.6 Ibovespa

O Ibovespa é o principal índice do mercado acionário brasileiro, o qual é composto de uma carteira com 86 empresas. É mantido e gerenciado pela bolsa de valores brasileira (B3).

2.1.7 Inflação

A variável IPCA não foi utilizada diretamente no treinamento dos modelos, porém foi utilizada para realizar o ajuste real das variáveis PIB e liquidez (M1). Os dados foram extraídos do SGS do Bacen.

QUADRO 2- Resumo das variáveis utilizadas no estudo.

<u>Variável</u>	<u>Descrição (séries)</u>	<u>Fonte</u>
Taxa de Câmbio	Taxa de câmbio - Dólar americano (compra) - Média do período mensal (3698).	Bacen
PIB	PIB mensal - valores correntes (R\$ milhões) (4380)	Bacen
Taxa de Juros	Taxa de juros - Meta Selic definida pelo Copom (432)	Bacen
Taxa de capital de giro (PJ)	Taxa média mensal de juros das operações de crédito com recursos livres para pessoas jurídicas. (25444)	Bacen
Liquidez	Meios de pagamento - M1 (saldo em final de período) (27791)	Bacen
Expectativas	Índice de confiança empresarial - sem ajuste sazonal. (1428466)	FGV - IBRE
Ibovespa	Média mensal do fechamento do Ibovespa.	B3
IPCA	Média mensal do IPCA - Utilizado para realizar ajustes de liquidez e PIB.	Bacen
Falências requeridas e decretadas	Levantamento mensal sobre falências decretadas pelas empresas em atividade no Brasil.	Serasa Experian

Fonte: Os autores (2023).

3 PROCEDIMENTOS METODOLÓGICOS

Para o processo de coleta de dados, de modo a extrair variáveis econômicas, foram utilizadas as bases de dados do Sistema Gerenciador de Séries Temporais do Banco Central do Brasil (SGS - Bacen) e do Instituto Brasileiro de Economia da Fundação Getúlio Vargas (IBRE/FGV). Para a coleta do número de falências requeridas, utilizou-se a base de dados do Serasa Experian.

Após possuir todos os dados em uma base comum, as variáveis PIB e Liquidez (M1), que estavam em valores nominais, foram transformadas para valores reais, utilizando-se dos valores históricos do Índice Nacional de Preços ao Consumidor Amplo (IPCA) para remover a inflação durante o período. Outra adaptação feita foi a suavização da variável explicada (falências requeridas) utilizando uma média móvel de 4 meses.

Além disso, os períodos (com base nos dados de requerimento de falências) de treinamento e teste (previsão) dos modelos foram definidos da seguinte forma:

Treinamento: abril de 2006 até junho de 2020;

Teste (previsão): julho de 2020 até junho de 2023;

Finalmente, como último passo antes de os testes serem executados, definiu-se os modelos de Inteligência Artificial (I.A.) a serem utilizados nas previsões, sendo eles: Bagging (Bootstrap Aggregating), Random Forest e XGBoost (Extreme Gradient Boosting). Além disso, o modelo ARIMA foi utilizado para comparação dos resultados com os gerados com uma metodologia mais tradicional.

Para execução dos testes, foi realizado um processo de normalização das variáveis, ou seja, um processo de ajuste das variáveis para que possam ser comparadas. A normalização é uma etapa importante no pré-processamento de dados para muitos algoritmos de aprendizado de máquina, especialmente aqueles que são sensíveis à escala das variáveis. Além disso, com relação a hiper parâmetros dos modelos, foi utilizado um padrão de 100 (cem) árvores de decisão que serão construídas por cada um deles.

Após realizar os testes com as variáveis explicativas e explicadas de meses iguais (i.e. PIB, Liquidez, Taxa de Câmbio etc. do mesmo mês do número de falências requeridas) e obter resultados pouco significativos, optou-se por utilizar uma abordagem de defasagens (deslocamento de uma variável em relação ao tempo, analisou-se dados anteriores para melhorar as previsões). Com as variáveis explicativas, criou-se uma base de dados cujo possui 10 (dez) séries de cada variável, sendo a primeira defasada em 2 (dois) meses e a última em 12 (doze) meses, usando as datas dos dados de requerimento de falências como referência.

Dois testes com duas variáveis similares em bases diferentes foram realizados: SELIC e Taxa Média de Capital de Giro para Pessoas Jurídicas (CGPJ). Observou-se que os resultados mais assertivos foram utilizando CGPJ e, por esta razão, utilizou-se apenas esta variável na base de dados para análises posteriores aos resultados.

A metodologia de mensuração de efetividade do resultado produzido pelos modelos foi o Mean Squared Error (Erro Quadrático Médio) ou MSE, uma métrica comum para avaliar a qualidade das previsões em modelos de regressão.

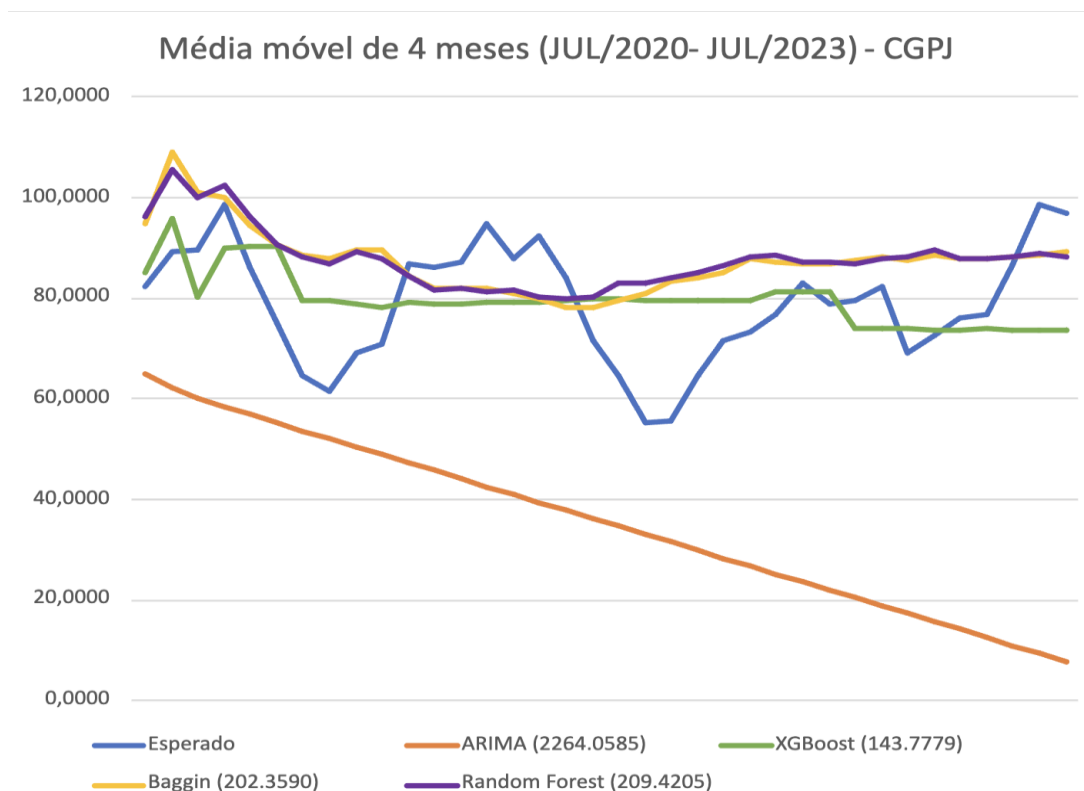
Neste endereço eletrônico pode ser encontrado a base utilizada para treinamento dos modelos em formato CSV (Comma Separated Values): <https://github.com/gustavo-fior/tcc/blob/master/cgpAA.csv>

Por fim, o arquivo do código fonte na linguagem de programação Python, utilizado para a execução dos modelos pode ser encontrado no seguinte endereço eletrônico: <https://github.com/gustavo-fior/tcc/blob/master/code.p>

4 RESULTADOS

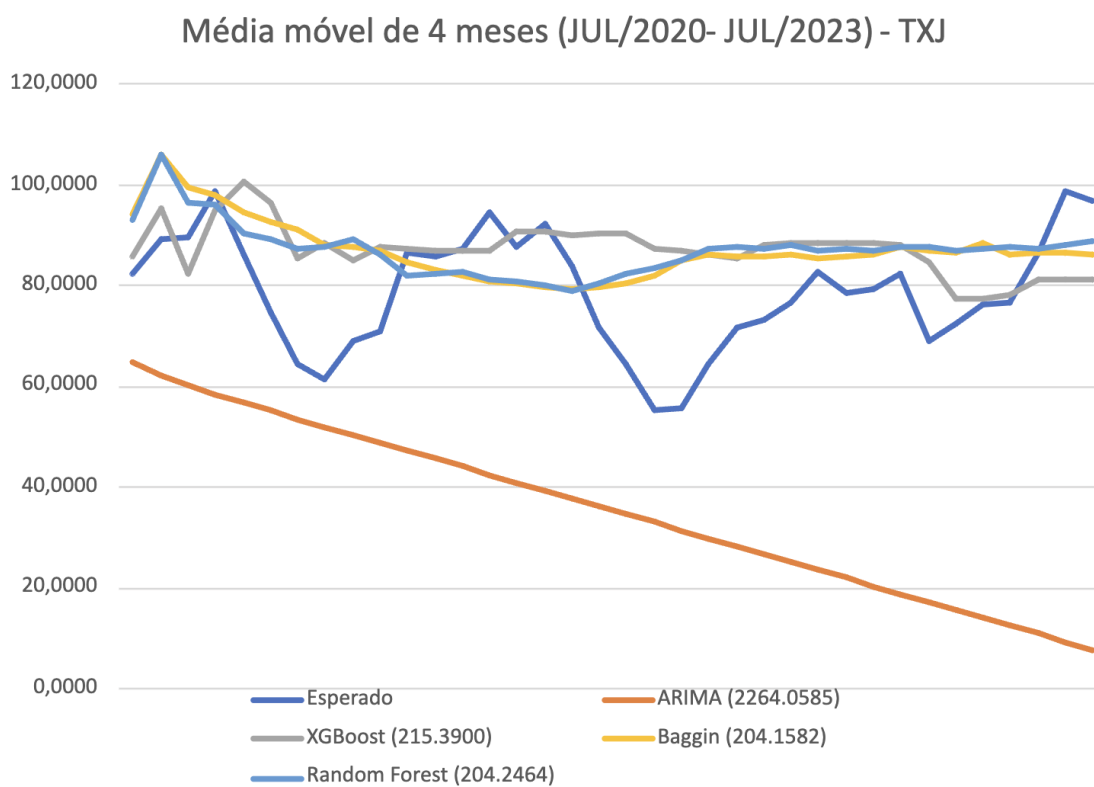
Destaca-se os resultados dos modelos nas figuras abaixo discriminadas por qual variável de juros foi utilizada nos modelos.

Figura 2- Resultados utilizando CGPJ.



Fonte: Mean Squared Error dos modelos entre parênteses.

Figura 3- Resultados utilizando SELIC.



Fonte: Mean Squared Error dos modelos entre parênteses (2023).

Como pode ser observado, o XGBoost foi o modelo que apresentou o melhor resultado entre os 8 (oito) testes - 4 (quatro) com CGPJ e 4 (quatro) com taxa de juros (SELIC). Ele obteve o menor Mean Squared Error (MSE) entre todos os modelos, finalizando com um valor de 143.7779 no modelo que teve como variável explicativa, representando juros a CGPJ. Além disso, vale destacar que os outros 2 (dois) modelos (Baggin e Random Forest) tiveram desempenhos parecidos e ainda melhor que o XGBoost, no modelo com taxa de juros (SELIC). Por fim, vale ressaltar que o modelo ARIMA performou consideravelmente pior que os modelos de Inteligência Artificial.

QUADRO 3- Definição das variáveis.

Código da variável	Descrição
i2 - i12	Taxa de Juros CGPJ (i)
x2 - x12	Taxa de câmbio (x)
b2 - b12	SELIC (b)
e2 - 12	ICE - (expectativas) (e)
l2 - l12	M1 (liquidez) (l)
g2 - g12	PIB real (g)
ib2 - ib12	IBOVESPA (ib)

Fonte: Os autores (2023).

Abaixo segue tabela com as variáveis de maior influência nos modelos de previsão.

Figura 4- As variáveis de maior influência nos modelos de previsão.

MAIORES PESOS - CGPJ - PESO ENTRE PARÊNTES ^1000		
XGBoost	Baggin	Random Forest
IBOVESPA - 8 Defasagens (258,3631873)	Liquidez (M1) - 10 Defasagens (132,5145146)	Liquidez (M1) - 11 Defasagens (119,6839667)
Liquidez (M1) - 10 Defasagens (254,6762824)	Liquidez (M1) - 11 Defasagens (123,3745977)	Liquidez (M1) - 12 Defasagens (101,9509005)
CGPJ - 10 Defasagens (136,0767484)	Liquidez (M1) - 3 Defasagens (90,2878429)	Liquidez (M1) - 9 Defasagens (97,05052909)
CGPJ - 11 Defasagens (97,37367928)	Liquidez (M1) - 2 Defasagens (77,00562369)	Liquidez (M1) - 10 Defasagens (94,07890957)
PIB - 8 Defasagens (91,7269662)	Liquidez (M1) - 4 Defasagens (61,38250287)	Liquidez (M1) - 2 Defasagens (84,98911118)

Fonte: Os autores (2023).

Por fim, esses foram os pesos de cada variável na análise de cada modelo:

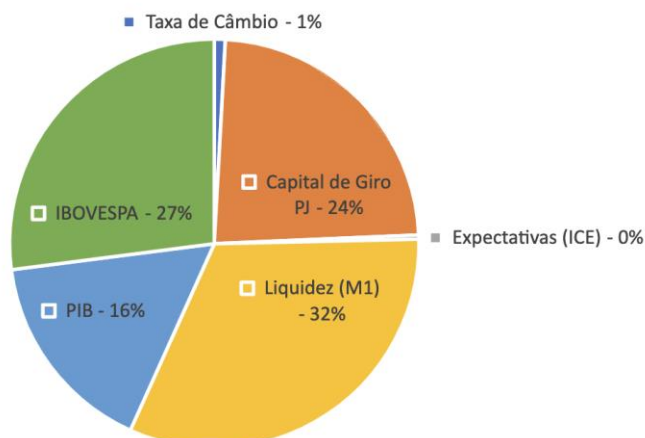
Figura 5- Peso geral das variáveis no modelo Baggin.



Fonte: Os autores (2023).

Figura 6- Peso geral das variáveis no modelo XGBoost.

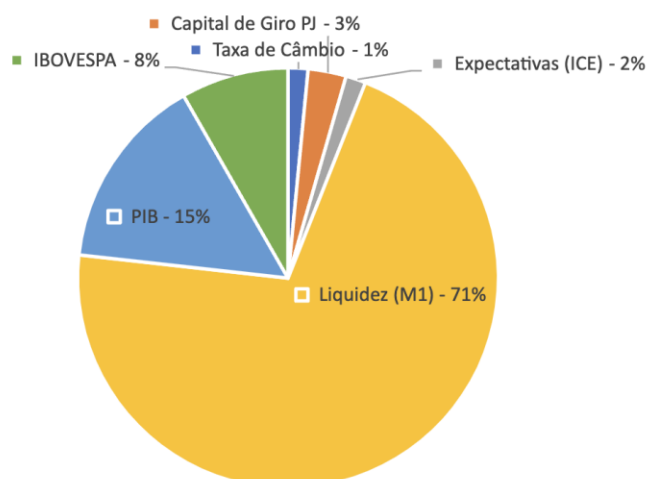
PESOS XGBOOST - CGPJ



Fonte: Os autores (2023).

Figura 7: Peso geral das variáveis no modelo Random Forest.

PESOS RANDOM FOREST- CGPJ



Fonte: Os autores (2023).

A partir da análise feita, percebe-se que a importância dos pesos das variáveis diverge entre os modelos, enquanto o Random Forest e o Bagging tiveram como variável mais importante para projeção sendo M1. Já o XGBoost, o modelo que melhor performou, realizou uma análise mais equilibrada, utilizando de 4 (quatro) variáveis como as mais importantes, sendo o PIB, M1, CGPJ e Ibovespa. Conclui-se que a variável liquidez (M1) foi a mais importante entre os modelos para a projeção de empresas em insolvência.

5 CONCLUSÃO

Com base nos resultados obtidos neste estudo, conclui-se que o modelo XGBoost se destacou como a escolha mais promissora para previsão de falências, demonstrando um desempenho superior em comparação com os demais modelos de inteligência artificial avaliados. O XGBoost exibiu o menor Mean Squared Error (MSE) e a maior constância de resultados e acurácia nos testes realizados, reforçando a capacidade de oferecer previsões mais precisas e confiáveis.

No entanto, é importante ressaltar que os modelos Baggin e Random Forest também apresentaram resultados significativos, superando, inclusive, o XGBoost em cenários específicos, como no caso da variável representando a taxa de juros (SELIC). Com relação às variáveis, a variável mais importante entre os modelos foi a Liquidez (M1).

Por fim, a análise revelou uma clara superioridade dos modelos de inteligência artificial em comparação com o modelo ARIMA, o que evidencia a capacidade desses modelos de aprendizado de máquina em lidar com a previsão de empresas em insolvência.

REFERÊNCIAS

ALTMAN, Edward I. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. **The journal of finance**, v. 23, n. 4, p. 589-609, 1968.

ALTMAN, Edward I. Why businesses fail. **Journal of Business Strategy**, v. 3, n. 4, p. 15-21, 1983.

ARAUJO, Aloisio; FUNCHAL, Bruno. A nova lei de falências brasileira: primeiros impactos. **Brazilian Journal of Political Economy/Revista de Economia Política**, v. 29, n. 3, p., 191-212, 2009.

BARBOZA, F; KIMURA, H; ALTMAN E. Machine learning models and bankruptcy prediction. **Expert Systems with Applications**, v. 83, 2017, p.405-417, ISSN 0957-4174. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0957417417302415>> Acesso em: 25 abr. 2023.

CONTADOR, C. R. Insolvência de empresas e política macroeconômica. **Revista de Administração da Universidade de São Paulo**, v. 20, n. 2, 1985.

EVERETT, Jim; WATSON, John. Small business failure and external risk factors. **Small Business Economics**, v. 11, n. 4, p. 371-390, 1998.

FLOR, A. **Pandemia pode levar 3,5 mil empresas à recuperação judicial e à falência, diz estudo**. Disponível em: <<https://g1.globo.com/economia/blog/ana-flor/post/2020/07/17/pandemia-pode-levar-35-mil-empresas-a-recuperacao-judicial-e-a-falencia-diz-estudo.ghtml>>. Acesso em: 24 out. 2023.

GOV BR. **Painéis do Mapa de Empresas**. gov.br. Disponível em: <<https://www.gov.br/empresas-e-negocios/pt-br/mapa-de-empresas/painel-mapadeempresas>> Acesso em: 25 abr. 2023.

Levantamento indica recorde histórico de empresas inadimplentes no Brasil em junho. **Folha de Pernambuco**, 2023. Disponível em: <<https://www.folhape.com.br/economia/levantamento-indica-recorde-historico-de-empresas-inadimplentes-no/283073/>>. Acesso em: 24 out. 2023.

LIU, Jia; WILSON, Nick. Corporate failure rates and the impact of the 1986 insolvency act: An econometric analysis. **Managerial Finance**, v. 28, n. 6, p. 61-71, 2002.

MESSA, Alexandre. **IMPACTO DA TAXA DE CÂMBIO SOBRE A COMPETITIVIDADE DA INDÚSTRIA BRASILEIRA**. Repositório. Disponível em: <<https://repositorio.ipea.gov.br/bitstream/11058/8723/1/Impacto%20da%20taxa.pdf>> Acesso em: 25 abr. 2023.

MOLITERNO, Danilo. **Pedidos de falência e recuperação judicial aumentam em cenário de juro alto e crédito escasso**. CNN Brasil. Disponível em: <<https://www.cnnbrasil.com.br/economia/pedidos-de-falencia-e-recuperacao-judicial-aumentam-em-cenario-de-juro-alto-e-credito-escasso/>> Acesso em: 25 abr. 2023.

NOGUEIRA, Ana Cecília. **PIB: o que é e quais seus impactos na economia?**. Blog Banco Inter. Disponível em: <https://blog.bancointer.com.br/o-pib-e-seus-impactos>. Acesso em: 23 abr. 2023.

PILAR, Ana Flávia. **Falências de empresas saltam 80% em dois anos, diz Serasa Experian**. Abrasel. Disponível em: <<https://abrasei.com.br/noticias/noticias/falencias-de-empresas-saltam-80-em-dois-anos-diz-serasa-experian/>> Acesso em: 23 abr. 2023.

PLATT, Harlan D.; PLATT, Marjorie B. Business cycle effects on state corporate failure rates. **Journal of Economics and Business**, v. 46, n. 2, p. 113-127, 1994.

RODRIGUEZ, Rodrigo. **A relação de variáveis macroeconômicas com a taxa de falência de empresas brasileiras**. Repositório Insper Edu. Disponível em: <http://repositorio.insper.edu.br/bitstream/11224/2297/3/Rodrigo%20Zalli%20Rodriguez_Trabalho.pdf> Acesso em: 23 abr. 2023.

SALMAN, A. Khalik; FUCHS, Matthias; ZAMPATTI, Davide. Assessing risk factors of business failure in manufacturing sector: a count data approach from Sweden. **International Journal of Economics, Commerce and Management**, v. 3, n. 9, p. 42-62, 2015.

SOUSA, Nathália. **Em 2022, 833 empresas pediram recuperações judiciais no Brasil**. Sampi. Disponível em: <<https://sampi.net.br/jundiai/noticias/2731078/jundiai/2023/01/em-2022-833-empresas-pediram-recuperacoes-judiciais-no-brasil>> Acesso em: 23 abr. 2023.

ZHANG, Jin; BESSLER, David A.; LEATHAM, David J. Aggregate business failures and macroeconomic conditions: a var look at the US between 1980 and 2004. **Journal of Applied Economics**, v. 16, n. 1, p. 179-202, 2013.