



## **Modelagem de Insolvência Empresarial com Redes Neurais: Evidências Empíricas da Recuperação Judicial no Brasil**

Filipe Casellato Scabora  
Universidade de São Paulo  
[fcscabora@usp.br](mailto:fcscabora@usp.br)

Marcelo Botelho da Costa Moraes  
Universidade de São Paulo  
[mbotelho@usp.br](mailto:mbotelho@usp.br)

### **Resumo**

O número de pedidos de recuperação judicial no Brasil tem apresentado crescimento consistente desde a edição da Lei nº 11.101/2005, tendência intensificada após a crise financeira global de 2008 e observada em outros países. Os impactos da insolvência sobre credores, investidores e sociedade estimularam a criação de modelos de previsão, tradicionalmente baseados em técnicas estatísticas e indicadores financeiros. Entretanto, avanços recentes demonstram que abordagens de inteligência artificial, como as RNA (Redes Neurais Artificiais), quando combinadas a informações contábeis, de mercado e macroeconômicas, podem gerar previsões mais acuradas. Este estudo avalia a eficácia das RNA em relação a modelos de regressão linear e logística na previsão do pedido de recuperação judicial por empresas brasileiras, com base em dados primários e variáveis de fluxos de caixa, de mercado e macroeconômicas. Os resultados indicam desempenho superior das RNA, reforçando a relevância de integrar diferentes fontes de informação e técnicas de inteligência artificial para antecipar cenários de crise e apoiar decisões estratégicas de credores, gestores e reguladores.

**Linha temática:** Contabilidade Financeira e Finanças

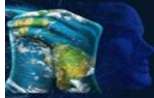
**Palavras-Chave:** recuperação judicial; redes neurais artificiais; insolvência.

### **1. Introdução**

Os efeitos da insolvência das organizações para os stakeholders levou inúmeros pesquisadores a investigarem as causas da falência, desenvolvendo modelos que buscassem prever a ocorrência desses eventos. Ao longo dos anos, e com o desenvolvimento de novas ferramentas e tecnologias no campo da estatística e informática, esses modelos foram submetidos a um processo de desenvolvimento histórico-cronológico, evoluindo a partir dos modelos teóricos (que focam nas causas qualitativas da falência), estatísticos (que focam nos sintomas da falência) e de inteligência artificial (que também focam nos sintomas da falência, mas com a capacidade de aprendizagem dos modelos).

Os chamados modelos estatísticos destacaram-se, desde cedo, como parte do mainstream da pesquisa em previsão de insolvência, o que levou a uma concentração das pesquisas na área, mesmo em tempos mais recentes, quando o desenvolvimento tecnológico demonstra que os modelos de inteligência artificial podem apresentar resultados melhores.

Muito além da controvérsia em torno da eficácia da grande maioria desses modelos para dados brasileiros, outros fatores justificam a realização de novos trabalhos na área: (i) o Brasil e o mundo experimentam, desde a crise mundial de 2008, um crescimento no número de falências e recuperações judiciais; (ii) no país, o advento da Lei nº. 11.101/05 (Lei de Recuperação e Falência – LRF) em 2005 modificou significativamente o mapa de falências



com a introdução da figura da recuperação judicial; e, por fim, (iii) as pesquisas na área, em sua maioria, além de adotarem modelos estatísticos clássicos, consideram como evento de interesse que pretendem observar, a ocorrência da falência ou a constatação da insolvência (técnica ou baseada em fluxos), quando, na verdade, existem inúmeros elementos que demonstram a supremacia da recuperação judicial como evento determinante a ser investigado.

Outro fator que distingue os modelos clássicos dos modelos de inteligência artificial diz respeito a seleção das variáveis, majoritariamente indicadores financeiros e de lucratividade extraídos do BP (Balanço Patrimonial) e da DRE (Demonstração do Resultado do Exercício), mesmo diante das evidências de que os fluxos de caixa têm impacto mais decisivo nos ciclos de vida das organizações (Dickinson, 2011), o que inclui a necessidade de buscarem ferramentas para superar períodos de crise, como a recuperação judicial e a falência.

As pesquisas em previsão de insolvência no Brasil também têm se limitado em razão da ausência de uma caracterização precisa do evento preditivo que pretendem ver analisado (insolvência). Isso porque a LRF estabelece uma distinção precisa entre os procedimentos de recuperação judicial e falência, que não é necessariamente encontrada em outros países, e pode mesmo representar um distanciamento temporal significativo.

Em sua grande maioria, os estudos sobre o tema no Brasil estão circunscritos a observar a capacidade preditiva de modelos até o ponto da insolvência técnica ou da falência, não demonstrando grandes preocupações em segregar a recuperação judicial como evento digno de nota.

A ausência dessa distinção é relevante, principalmente se considerarmos a existência de casos, no Brasil, em que o pedido de recuperação judicial e a decretação de falência estão separados por mais de uma década. Some-se também o fato de que a Lei nº. 14.112/20 introduziu uma série de mudanças na LRF, enquanto o número de recuperações judiciais no país não para de crescer.

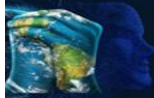
Diante disso, o presente artigo busca demonstrar o desempenho superior de abordagens baseadas em inteligência artificial – como as RNA – na previsão de insolvência, reforçando a relevância de integrar diferentes fontes de informação e técnicas de inteligência artificial para antecipar cenários de crise e apoiar decisões estratégicas de credores, gestores e reguladores.

## **2. Revisão da Literatura**

A recuperação judicial e a falência não são problemas recentes (Baird, 1991). O grande impacto desses eventos para o ambiente corporativo instigou, ao longo dos anos, um volume significativo de análises e modelos preditivos, principalmente a partir da década de 1960, com os trabalhos seminais de Altman (1968) e Beaver (1966).

Aziz e Dar (2006) propõe uma classificação tripartite para esses modelos de predição, que seriam, então, divididos em três categorias:

- (i) teóricos (Crouhy et al., 2000; Lucas, 1995; Aziz et al., 1988), voltados à explicação teórica da falência utilizando dados obtidos, inclusive, por meio de técnicas estatísticas, geralmente de análise multivariada, para fornecer elementos qualitativos que deem suporte à teoria adotada;
- (ii) estatísticos (Altman, 1968; Beaver, 1966), com foco nos sintomas da falência, utilizam dados obtidos a partir das informações contábeis das organizações, por meio de análise univariada ou multivariada, considerados o *mainstream* dos modelos de previsão de insolvência; e, finalmente,



- (iii) de inteligência artificial (Zhang et al., 1999; Brockett et al., 1994), também focados nos sintomas da falência, mas com um elemento tecnológico que faltava aos modelos puramente estatísticos, geralmente através de análises multivariadas em relação às informações contábeis das organizações.

Um panorama interessante sobre a acuidade desses modelos de previsão é apresentado por Jardim (2017), que traça um paralelo entre os principais trabalhos na área e o modelo conceitual adotado (teórico, estatístico ou de inteligência artificial).

A eficiência dos modelos de previsão também é abordada por Pereira e Martins (2016); os autores inovam ao conjugar informações relevantes acerca dos indicadores (como Jones et al., 2017), modelos de previsão, eficácia e cronologia utilizados, oferecendo um retrato significativo dos trabalhos realizados a partir da década de 1970 até os dias atuais, para dados brasileiros e estrangeiros.

Nada obstante o avanço na utilização de modelos de inteligência artificial a partir dos anos 1990 apontada por diversos autores (Yu et al., 2014), as conclusões apresentadas por Pereira e Martins (2016) e Aziz e Dar (2006) são semelhantes, ao indicarem tratar-se de uma abordagem ainda embrionária, apresentando pouca representatividade. Jardim e Séverin (2011) criticam os modelos teóricos e estatísticos adotados, não apenas em razão da abordagem, mas também ressaltam que tais modelos têm se tornado fonte de inúmeras críticas quanto à forma como as variáveis são selecionadas, a influências de fatores exógenos e a própria eficácia da previsão.

Para Jones et al. (2017), mesmo no campo da estatística os modelos de previsão de falência têm deixado de explorar inovações que poderiam trazer resultados interessantes, havendo evidências de que os modelos mais modernos podem suplantam significativamente a performance das abordagens mais tradicionais (Jones et al., 2015; Jones, 2017).

## **2.1. Falência e recuperação judicial**

No Brasil, até a edição da LRF, a incapacidade da organização em honrar seus contratos tinha como resultados possíveis a falência ou a concordata, ambas reguladas pelo DL (Decreto-Lei) nº. 7.661/45.

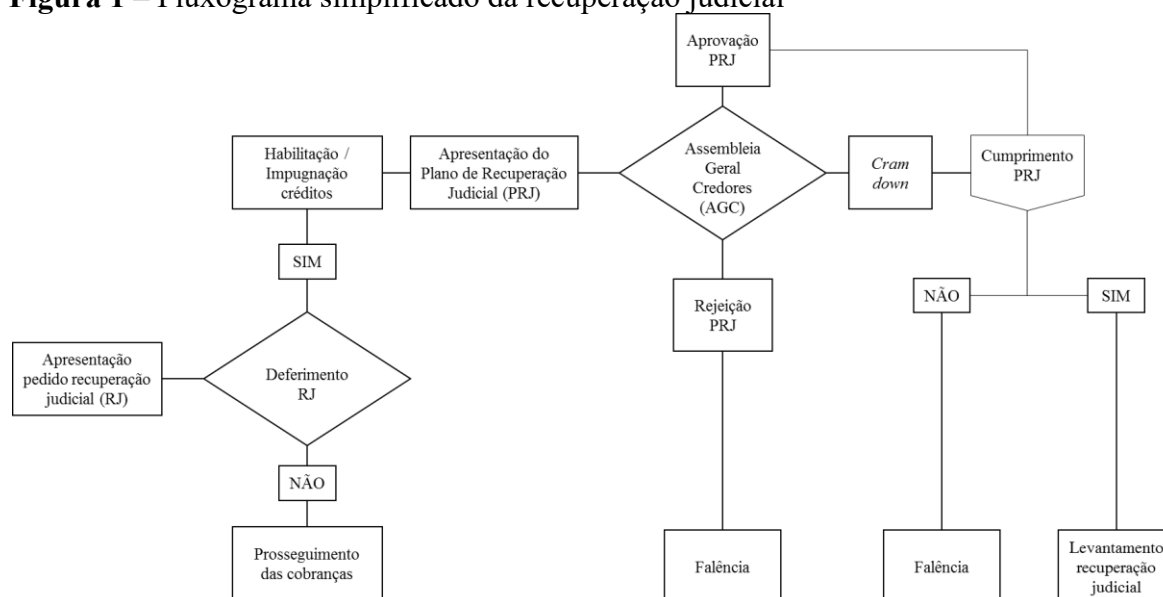
No regime do DL, a falência operava-se por meio da liquidação dos ativos da organização para adimplemento das obrigações assumidas, cessando as suas atividades, enquanto a concordata constituía uma espécie de moratória judicial requerida pelo devedor, que desobrigava o pagamento dos seus créditos por um determinado período, vedando a expropriação de bens para exigência das dívidas existentes, para permitir a renegociação junto aos credores.

Somente com o advento da LRF introduziu-se, na legislação brasileira, um instituto jurídico voltado para assegurar a reestruturação econômico-financeira das organizações com problemas de insolvência. A chamada recuperação judicial, prevista no artigo 47 da LRF (2005) “tem por objetivo viabilizar a superação da situação de crise econômico-financeira do devedor, a fim de permitir a manutenção da fonte produtora, do emprego dos trabalhadores e dos interesses dos credores, promovendo, assim, a preservação da empresa, sua função social e o estímulo à atividade econômica”.

O dispositivo legal introduziu na legislação brasileira o “princípio da preservação da empresa”, valor básico que, objetivando a manutenção da atividade econômica e a função social da organização, assegura, sob determinadas condições, a concessão de estímulos para recuperação econômico-financeira da empresa, através de um processo complexo e com participação ativa dos credores:



**Figura 1** – Fluxograma simplificado da recuperação judicial



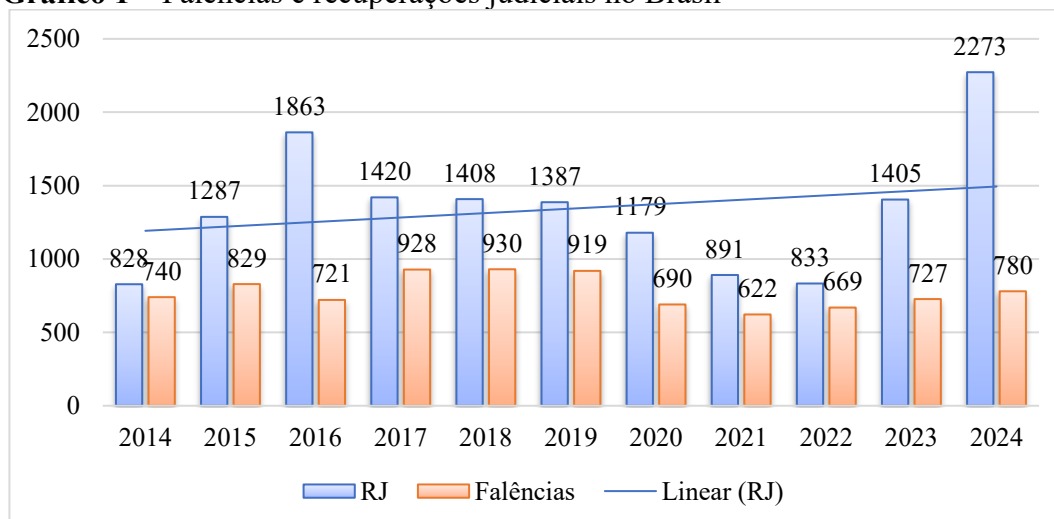
Fonte: elaborado pelos autores.

## 2.2. Recuperação judicial no Brasil

De acordo com o indicador Serasa Experian de Falências e Recuperações (Serasa, 2025), o número de pedidos de recuperação judicial atingiu o seu pico histórico em 2016, com 1.863 pedidos. Nos anos subsequentes, as recuperações judiciais diminuíram, mesmo durante a pandemia do Covid-19, em grande parte devido às medidas regulatórias e macroprudenciais implementadas pelo governo e pelas autoridades monetárias. Essas iniciativas, entre outros efeitos, permitiram que os bancos alongassem os prazos das dívidas corporativas, postergando a necessidade de as empresas recorrerem à recuperação judicial.

No entanto, em 2023, os pedidos voltaram a crescer, aumentando aproximadamente 69% em comparação com o mesmo período do ano anterior. Até o último trimestre de 2024, o volume de novos casos já havia superado o pico registrado em 2016, com 2.273 novas recuperações (Serasa, 2025).

**Gráfico 1** – Falências e recuperações judiciais no Brasil



Fonte: elaborado pelos autores.



Para Evans e Borders (2014), o crescimento no número de empresas buscando alternativas à insolvência é um fenômeno mundial, que pode ser observado também a partir dos dados brasileiros.

### 2.3. Fatores determinantes da insolvência

O estudo de falências e recuperações do Serasa Experian (2025) é construído a partir do levantamento mensal das estatísticas de falências (requeridas e decretadas) e das recuperações judiciais e extrajudiciais registradas mensalmente na sua base de dados, provenientes dos fóruns, varas de falências e dos Diários Oficiais e da Justiça dos Estados.

Para os economistas responsáveis pelo levantamento, o baixo dinamismo econômico somado a expressiva queda no consumo das famílias, determinada pela escalada do desemprego e da inflação, e as altas taxas de juros impactaram negativamente a atividade dos setores comercial, industrial e de serviços com reflexos diretos na saúde financeira das empresas (Serasa, 2025), provocando um crescimento acentuado nos pedidos de recuperação judicial.

Com efeito, ao tratar dos fundamentos da recuperação judicial, Teixeira (2012) esclarece que a crise de uma atividade econômica – assim entendida como o conjunto de fatores que leva uma companhia a buscar a recuperação judicial – podem ser de origem econômica (quando as vendas dos produtos ou a prestação de serviços não são realizadas em quantidade suficiente à manutenção do negócio), financeira (quando o empresário tem falta de fluxo de caixa, dinheiro ou recursos disponíveis para pagar suas prestações obrigacionais) ou patrimonial (quando o ativo do empresário é menor do que o seu passivo, logo, seus débitos superam os seus bens e direitos).

A classificação proposta por Teixeira (2012) não é diferente dos conceitos de insolvência apresentados por Wruck (1990) e Altman (1983), evidenciando, portanto, que na opinião desses autores, os eventos de crise que levam as empresas à recuperação judicial têm em comum um aspecto relevante no que diz respeito a capacidade da organização em gerar receitas e a administrar a sua alocação, refletida nos seus fluxos de caixa.

O trabalho de Dickinson (2011) apresenta inflexões importantes no que toca ao comportamento esperado dos fluxos de caixa nas diversas etapas dos ciclos de vida das organizações:

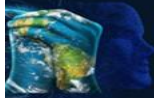
**Tabela 1** - Comportamento dos fluxos de caixa no ciclo de vida das organizações

| Natureza | Introdução | Crescimento | Maturidade | Oscilação | Declínio |
|----------|------------|-------------|------------|-----------|----------|
| FCO      | (-)        | (+)         | (+)        | (+/-)     | (-)      |
| FCI      | (-)        | (-)         | (-)        | (+/-)     | (+)      |
| FCF      | (+)        | (+)         | (-)        | (+/-)     | (+/-)    |

Fonte: adaptado de Dickinson (2011)

Nas fases de oscilação e declínio, o comportamento esperado do FCI (Fluxo de Caixa de Investimento) e do FCF (Fluxo de Caixa de Financiamento) é positivo, o que pode ser explicado pela necessidade da empresa em buscar recursos para fazer frente às suas obrigações nos períodos de arrefecimento do FCO (Fluxo de Caixa Operacional), seja por meio da alienação de ativos da companhia ou obtenção de empréstimos e financiamentos.

Vê-se, portanto, que a análise da DFC (Demonstração de Fluxo de Caixa) permite avaliar o comportamento desses fluxos de caixa ao longo do tempo, na medida em que evidencia as decisões quanto à origem, administração e aplicação de recursos (Ohlson & Aier, 2009), sendo possível buscar relações entre os fluxos evidenciados na DFC e o fato de as entidades buscarem a recuperação judicial como alternativa à insolvência.



### 3. Materiais e Métodos

Embora a grande maioria das pesquisas na área estabeleça a decretação da falência ou o momento em que a empresa se torna insolvente, é com o pedido de recuperação judicial que a organização sinaliza mais fortemente aos stakeholders a sua situação econômico-financeira e a necessidade de buscar uma reorganização para manutenção dos contratos e obrigações.

Além disso, diversos outros motivos justificam a prevalência da recuperação judicial como evento de interesse, dentre os quais destaca-se a possibilidade de renegociação forçada dos créditos da organização, até mesmo à revelia de parte dos credores (*cram down*), a suspensão dos processos de cobrança durante o *stay period*, o efeito negativo da recuperação judicial na precificação dos créditos (*non performing loans*), a possibilidade de dedução imediata como perda para fins tributários (Lei nº. 9.430/96) e a própria interferência do Poder Judiciário nos negócios da organização (Gilson, 1991).

Por essas razões é que o pedido de recuperação judicial é o momento mais apropriado para caracterizar o evento de interesse que se pretende prever, ainda que a grande maioria das pesquisas sobre o tema não utilize a recuperação judicial como marco determinante dos modelos de previsão ou, pior ainda, não faça qualquer distinção entre esses eventos (falência, recuperação judicial e insolvência), tratando-os como sinônimos, apesar dos seus efeitos diversos para os credores e a própria organização.

#### 3.1. Seleção das variáveis

Embora a existência de relação entre as situações de insolvência e a aplicação dos fluxos de caixa seja evidenciada por inúmeros autores, a utilização de indicadores financeiros de BP e DRE tem grande prevalência nas pesquisas na área (Jones et al., 2017). Pereira e Martins (2016) apontam que os indicadores envolvendo fluxos de caixa não figuram entre as medidas mais utilizadas pela literatura nos modelos de previsão de insolvência, mesmo diante de inúmeras evidências de que a adoção de indicadores e métodos alternativos tem o potencial de apresentar resultados mais eficazes e promissores (Jones et al., 2017; Jones, 2017; Jones et al., 2015).

Tinoco e Wilson (2013) avaliam modelos de previsão com variáveis contábeis, macroeconômicas e de mercado, concluindo que a eficácia dos resultados obtidos é maior quando essas variáveis são tratadas de forma não excludente; tendo como objetivo a construção de um modelo de previsão mais acurado, além das variáveis de fluxo de caixa, foram selecionados ainda outros indicadores de mercado e macroeconômicos:

**Tabela 2** - Variáveis selecionadas

| Natureza                     | Variável                                | Sigla          |
|------------------------------|---|----------------|
| Dependente ( <i>output</i> ) | Tempo                                   | T              |
|                              | Recuperação Judicial                    | RJ             |
| Contábeis / Empresa          | Fluxo sobre Lucro Líquido               | FCO sobre LL   |
|                              | Crescimento anual FCO                   | $\Delta$ FCO   |
|                              | Crescimento anual sobre Lucro Líquido   | $\Delta$ LL    |
|                              | Crescimento anual Receita Total         | $\Delta$ RT    |
|                              | Giro do Ativo                           | Giro Ativo     |
|                              | Margem Líquida                          | Margem Líquida |
|                              | Tamanho (LN do Total de Ativos)         | LNTam          |
|                              | Idade (LN da Idade)                     | LNIdade        |
| Macroeconômicas e de Mercado | Setor                                   | Setor          |
|                              | Taxa Selic                              | Selic          |
|                              | Inflação (IPCA)                         | IPCA           |
| Fluxo de Caixa               | Fluxo de Caixa Líquido de Investimento  | FCI            |
|                              | Fluxo de Caixa Líquido de Financiamento | FCF            |

Fonte: elaborado pelos autores.



### **3.2. Amostra e coleta de dados**

Pereira e Martins (2016), Jones et al. (2017) e Rosa e Gartner (2018) se debruçam sobre a controvérsia em torno do tamanho da amostra, uma vez que as evidências obtidas na literatura indicam variações em razão do modelo e das variáveis selecionadas.

Outras restrições também podem ser encontradas especificamente para o caso brasileiro, dentre as quais pode-se destacar: (i) o pequeno número de companhias listadas em recuperação judicial; (ii) a instituição tardia da DFC no país, apenas a partir da edição da Lei nº. 11.638/07; e, (iii) as limitações legais em torno da obrigatoriedade de publicação das demonstrações financeiras (art. 176, § 6º da Lei nº. 6.404/76 e art. 3º da Lei nº. 11.638/07) e da exigência de auditoria independente (artigo 177, §3º da Lei nº. 6.404/76).

Em razão desses fatores, fora necessário a coleta de dados primários, junto aos Diários Oficiais e Juntas Comerciais de 25 Estados brasileiros, adotando-se as seguintes providências:

- (i) levantamento das demonstrações contábeis de 50 companhias em recuperação judicial, por meio de consulta de dados primários (Poder Judiciário, Juntas Comerciais e bolsa de valores – B3), no período entre 31/12/2008 (quando a DFC se tornou obrigatória) e 31 de dezembro do ano-calendário anterior ao pedido de recuperação judicial;
- (ii) a estrutura da amostra do item (i) foi reproduzida para formar um grupo de controle, com o mesmo número de componentes, no período entre 31/12/2008 e 31/12/2017; e,
- (iii) os dados foram analisados e tratados, chegando-se a uma amostra final composta de 94 empresas, sendo metade (47) delas em recuperação judicial e a outra metade (47) solvente, para seleção das variáveis indicadas na Tabela 2, que resultou em 669 observações.

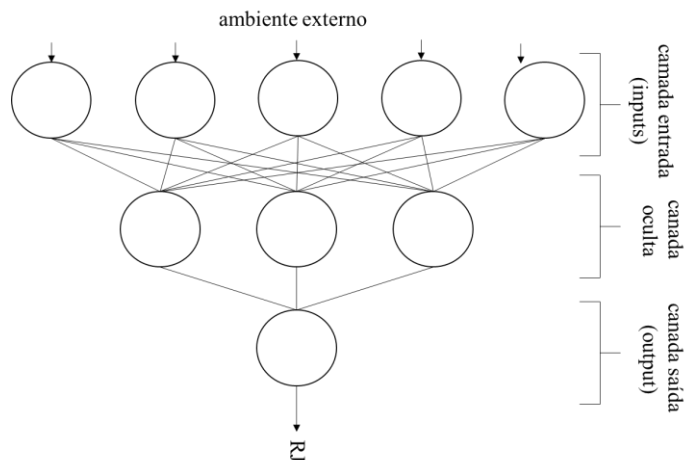
O número de empresas integrantes da amostra final encontra-se compatível com os levantamentos semelhantes apresentados por Pereira e Martins (2016), Jones et al. (2017) e Rosa e Gartner (2018), em se tratando de dados primários. A estrutura da amostra e do grupo de controle é bastante semelhante em termos geográficos, contendo observações de companhias de todas as regiões brasileiras e setoriais (indústria, comércio e serviços), além de terem sido submetidos a teses de estatística descritiva.

### **3.3. Modelos utilizados**

As RNA são ferramentas de inteligência artificial capazes de se auto adaptar e ajustar para realizar uma atividade ou comportamento determinado. Seu modelo de processamento é baseado nos neurônios; de forma computacional, as informações interligam-se por uma rede na qual cada unidade recebe e combina uma série de entradas (inputs) numa única saída, que dá entrada a uma nova unidade até a saída final da rede (output) ou a resposta do problema (Bialoskorski, Nagano & Moraes, 2006).



Figura 2 - Representação de uma RNA *feed forward*



Fonte: Scabora (2023).

A escolha das RNA se justifica vez que a técnica é menos sensível a problemas quanto à distribuição e ao tamanho da amostra, que não têm reflexo tão significativo nos resultados quanto nos modelos estatísticos tradicionais (Trippi & Lee, 1996), além de serem uma ferramenta de inteligência que funciona bem com variáveis que se alteram no tempo, como é o caso dos fluxos de caixa. Para implementação das RNA foram observadas as etapas de definição, treinamento, utilização e manutenção da RNA.

Optou-se por uma RNA *feed forward*, em que os neurônios são ordenados em fase ou camadas, com cada camada se conectando a próxima fase da rede, sem conexões de retorno, isto é, sem referências circulares no raciocínio da rede (Azayte & Achchab, 2016). O algoritmo de *backpropagation* foi usado para treinamento da RNA; calculando o erro esperado, o algoritmo corrige os pesos atribuídos em cada uma das camadas, partindo da camada de saída até a camada de entrada, de forma a assegurar que o output seja o resultado dos inputs.

A utilização das redes baseou-se na função de classificação *Multilayer Perceptron*, através do software Weka 3.8.3, obedecendo os seguintes parâmetros para treinamento:

- (i) *learning rate* – variando entre 0,2 e 0,4, com alterações de 0,05; e,
- (ii) *training test* – variando entre 500 e 50000, com alterações escalonadas de 500, 1000 e 5000. O tempo de processamento e construção do modelo das RNA também variou entre 0,55 e 356,44 segundos e, ao final, foram testadas 96 RNA diferentes para identificação do modelo mais eficaz.

Para permitir o confronto dos resultados, as variáveis selecionadas na Tabela 2 também foram submetidas à análise por regressão linear e logística, uma das ferramentas estatísticas mais importantes na análise de dados, principalmente em se tratando de avaliar as relações entre as variáveis existentes num determinado conjunto.

Além disso, como demonstram Pereira e Martins (2016), a aplicação de técnicas de regressão são uma constante envolvendo pesquisas na área de previsão de insolvência, desde os trabalhos iniciais de Altman (1968), passando pelos principais modelos brasileiros (Elizabetsky, 1976; Kanitz, 1978; Matias, 1978; Altman, Baidya & Dias, 1979) até trabalhos mais recentes, como Tinoco e Wilson (2013).



#### 4. Resultados

A amostra não apresentou problemas de multicolinearidade ( $VIF < 2,0$  e  $1/VIF > 0,69$ ) e heterocedasticidade ( $p$ -valor = 0,1672), porém verificou-se a não normalidade das observações ( $p$ -valor = 0,0000 em todas as variáveis), prejudicando a generalização de resultados. Os dados foram então submetidos à regressão linear, em relação às variáveis independentes RJ e T, e à regressão logística (LOGIT), em relação à variável independente RJ, através do software Stata/SE 14.0.

Os testes permitiram observar que o modelo não tem grande capacidade explicativa em relação à ambas as variáveis independentes ( $R^2=0,0491$  para RJ e  $R^2=0,044$  para T) na regressão linear e o percentual de classificações corretas, na LOGIT, é de 64,13%, sendo que a maioria dos erros classificatórios se encontra no grupo de “falso negativo”, isto é, empresas classificadas como em recuperação judicial que não se encontram em recuperação judicial.

Em seguida, os dados foram submetidos a análise por meio de RNA; o tempo de processamento e construção do modelo das RNA variou entre 0,55 e 356,44 segundos e, ao final, foram testadas 96 RNA diferentes, para identificação do modelo de previsão mais adequado. Os resultados obtidos no teste de *cross-validation* encontram-se abaixo:

**Tabela 3 – Cross-validation test**

| Variável independente          | RJ     | T       |
|--------------------------------|--------|---------|
| Erro médio absoluto            | 0,3458 | 1,4005  |
| RMSE                           | 0,4746 | 4,3700  |
| Erro médio relativo            | 73,07% | 184,54% |
| Raiz do erro relativo quadrado | 97,58% | 159,93% |
| Observações                    | 669    | 669     |

Fonte: elaborado pelos autores a partir dos dados do Weka 3.8.3.

A Tabela 4 apresenta os percentuais de classificações corretas obtidos pelas melhores RNA no teste training set, em relação à cada uma das variáveis independentes:

**Tabela 4 – Melhores classificações – RJ e T**

| Variável | Learning Rate | 0,25   | RMSE    |
|----------|---------------|--------|---------|
| RJ       | 45000         | 88,19% | 32,30%  |
| T        | 30000         | 87,85% | 133,29% |

Fonte: elaborado pelos autores a partir dos dados do Weka 3.8.3.

Por fim, também foram analisadas as correlações entre cada uma das variáveis dependentes e a saída da rede (*output*), através de *Correlation based Feature Selection* (CFS), aliada à correlação linear de Pearson, uma vez que todas as variáveis são numéricas (ou foram transformadas em variáveis numéricas), conforme proposto por Hall (1998):

**Tabela 5 – Correlação dos atributos (RJ e T)**

| Classificação | RJ      | T              |
|---------------|---------|----------------|
| 2             | FCF     | Giro Ativo     |
| 3             | FCI     | LNIdade        |
| 4             | IPCA    | FCF            |
| 5             | Selic   | FCI            |
| 6             | LNTam   | LNTam          |
| 7             | Setor   | Margem Líquida |
| 8             | LNIdade | Setor          |

continua



**Tabela 5** – Correlação dos atributos (RJ e T)

| Classificação | RJ             | conclusão    |  |
|---------------|----------------|--------------|--|
|               |                | T            |  |
| 9             | Margem Líquida | IPCA         |  |
| 10            | Giro Ativo     | $\Delta$ LL  |  |
| 11            | $\Delta$ RT    | Selic        |  |
| 12            | $\Delta$ LL    | $\Delta$ RT  |  |
| 13            | $\Delta$ FCO   | $\Delta$ FCO |  |
| 14            | FCO sobre LL   | FCO sobre LL |  |

Fonte: elaborado pelos autores a partir dos dados do Weka 3.8.3.

Analisando os resultados obtidos, na Tabela 4 é possível identificar as RNA que apresentaram a maior capacidade de classificação: em relação à variável RJ o modelo foi capaz de classifica corretamente 88,19% das empresas que integram a amostra, apresentando ainda o melhor RMSE (32,30%). Já em relação à variável T, a RNA com melhor classificação conseguiu identificar corretamente 87,85% das empresas, com RMSE de 133,29%.

O RMSE – ou *Root Mean Square Error* – pode ser definido como a raiz do erro médio quadrático da diferença entre o valor observado e o valor da previsão; trata-se de uma métrica normalmente utilizada para expressar a acurácia de modelos, com a vantagem adicional de que apresenta os valores do erro nas mesmas dimensões da variável analisada (Takacs, 1985; Zumel & Mount, 2014).

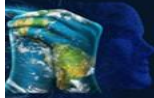
Nada obstante a superioridade da capacidade de classificação das RNA (88,19%) em comparação à LOGIT (64,13%), é necessário que a análise dos resultados ultrapasse a mera confrontação dos erros e acertos da previsão. Por essa razão, a matriz de confusão foi empregada para validar o processo de aprendizado supervisionado (Zumel & Mount, 2014) da RNA: como a amostra e o grupo de controle, contém, respectivamente, apenas empresas em recuperação judicial e organizações saudáveis, a matriz de confusão permitiu identificar os erros e acertos na saída final da rede em relação aos dois conjuntos (amostra e grupo de controle):

**Tabela 6** – Matriz de Confusão

| Classificação                      | LOGIT  |     |       | RNA    |     |       |
|------------------------------------|--------|-----|-------|--------|-----|-------|
|                                    | D      | ~D  | Total | D      | ~D  | Total |
| +                                  | 56     | 39  | 95    | 208    | 49  | 257   |
| -                                  | 201    | 373 | 574   | 30     | 382 | 412   |
| Total Geral                        | 257    | 412 | 669   | 238    | 431 | 669   |
| Sensibilidade                      | 21,79% |     |       | 87,40% |     |       |
| Especificidade                     | 90,53% |     |       | 88,63% |     |       |
| Valor preditivo positivo           | 58,95% |     |       | 80,93% |     |       |
| Valor preditivo negativo           | 64,98% |     |       | 92,72% |     |       |
| Falsa taxa + para ~D verd.         | 9,47%  |     |       | 11,37% |     |       |
| Falsa taxa - para D verd.          | 78,21% |     |       | 12,60% |     |       |
| Falsa taxa + para classif. +       | 41,05% |     |       | 19,07% |     |       |
| Falsa taxa - para classif. -       | 35,02% |     |       | 7,28%  |     |       |
| Classificação Correta (Eficiência) | 64,13% |     |       | 88,19% |     |       |

Fonte: elaborado pelos autores a partir dos dados do Weka 3.8.3 do Stata/SE 14.0.

Além da acurácia, outro dado importante que se extrai da matriz de confusão é a chamada eficiência – representada pela média aritmética da sensibilidade e especificidade. Enquanto a especificidade permite avaliar a capacidade da RNA em prever corretamente as empresas que não estão em recuperação judicial (*true negative*), a sensibilidade retorna a



proporção dos *true positives*, que representa a capacidade do modelo em prever corretamente as empresas em recuperação judicial.

A Tabela 6 também evidencia a eficiência superior das RNA (88,01%) em relação à LOGIT (56,16%) – média aritmética da sensibilidade e especificidade –, de forma a balancear os verdadeiros positivos e negativos do modelo, análise essencial uma vez que a acurácia, isoladamente, pode induzir a conclusões equivocadas, pois não leva em consideração os acertos positivos e negativos.

Demonstrada a superioridade da performance dos modelos propostos (RNA) em relação às modelagens mais tradicionais, o teste CFS (Tabela 5) demonstrou que dentre as variáveis mais correlacionadas com a saída (*output*) da rede encontram-se elementos de fluxo de caixa (FCF e FCI) e outros de natureza macroeconômica (IPCA e SELIC) e da empresa (Idade e Tamanho) em detrimento de indicadores financeiros, confirmado as evidências de que a conjugação de informações contábeis, macroeconômicas e de mercado (e modelos menos heterodoxos) produz melhores resultados no campo da previsão de insolvência (Tinoco & Wilson, 2013; Jones et al., 2015).

Por fim, a elevada correlação das variáveis de FCF e FCI com a saída da rede – em contraste com a baixa correlação dos indicadores financeiros – também corrobora as hipóteses levantadas, no sentido de que os fluxos de caixa são determinantes para a recuperação judicial das organizações; para Dickinson (2011) os fluxos de caixa de financiamento e investimento tem um papel decisivo nas fases de oscilação e declínio das organizações, como é o caso da recuperação judicial. As conclusões da autora também podem explicar a alta correlação do Giro do Ativo, na medida em que a alienação de bens e direitos não é um dos meios assegurados pela LRF para soerguimento da empresa. Já com relação ao FCO, o atributo se mostrou pouco correlacionado com o resultado da rede, ratificando as evidências de que o indicador, isoladamente, não se destaca enquanto variável na previsão de insolvência (Casey & Bartczack, 1985).

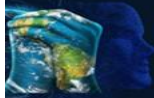
## 5. Conclusão

A conclusão deve retomar a relevância do problema de pesquisa e resumir, de forma concisa, o significado teórico e/ou prático dos resultados. Interprete-os à luz da literatura, destacando a contribuição do estudo. É importante abordar limitações da pesquisa e sugerir direções para trabalhos futuros.

Os dados do indicador Serasa Experian de Falências e Recuperações (2025) indicam que o número de pedidos de recuperação tem crescido, de forma sistemática, desde a edição da LFR, em 2005. Embora para Evans e Borders (2014) isso seja um fenômeno mundial que se acentuou com a crise financeira de 2008, os efeitos da insolvência das organizações têm estimulado a proliferação de modelos de previsão, principalmente a partir da década de 1960 (Altman, 1968), com os mais diversos resultados.

Khashei e Bijari (2010) sustentam que a integração de diferentes métodos e indicadores pode ser uma maneira efetiva de aperfeiçoar a capacidade preditiva desses modelos, principalmente em séries temporais. Nesse contexto, buscou-se desenvolver um modelo de previsão aprimorado, principalmente por meio da integração diferentes métodos – RNA– e indicadores pouco utilizados pela literatura, mormente em relação a dados brasileiros.

Além disso, em sua grande maioria, os estudos sobre o tema no País (Elizabetsky, 1976; Kanitz, 1978; Matias, 1978; Altman, Baidya & Dias, 1979) focam na previsão da insolvência ou da falência, não demonstrando grandes preocupações em segregar a recuperação como evento de interesse, apesar de inúmeros fatores que justificarem a sua prevalência, passando pela renegociação forçada dos créditos (arts. 50 e 58 da LRF), o efeito negativo no rating e na precificação dos créditos e até mesmo o tratamento tributário atribuído aos créditos submetidos



à RJ (art. 9º, §1º da Lei nº. 9.430/96) e influência (nem sempre positiva) do Judiciário na gestão/administração da organização durante a organização (Gilson, 1991).

Ao confrontar os resultados obtidos em relação à variável RJ identificou-se que o modelo proposto possui capacidade classificatória (88,1913%) significativamente superior do que as modelagens mais tradicionais testadas, como a LOGIT (64,13%). Contudo, a comparação dos resultados é um desafio; a RNA apenas ocasionalmente apresenta grande capacidade explicativa (Trippi & Lee, 1996), o que é muito mais comum nos modelos de análise discriminante e regressão consagrados para dados brasileiros.

Assim, foram analisadas ainda o RMSE e a matriz de confusão: os resultados obtidos também demonstram a superioridade do modelo proposto em relação à acurácia (0,323 x 0,47935), sensibilidade (87,40% x 21,79%) e eficiência (88,01% x 51,60%) – este último atributo (eficiência) também é superior àquela apresentada pela maioria dos modelos de análise multivariada no amplo levantamento bibliográfico realizado por Pereira e Martins (2016) no período entre 1930 e 2015.

A correlação dos atributos com a saída final da rede também evidenciou elementos de fluxo de caixa (FCF e FCI), macroeconômicos (IPCA e SELIC) e da empresa (Idade e Tamanho) em detrimento de indicadores financeiros mais usuais. Em função disso foi possível corroborar as hipóteses aventadas de que: (i) os fluxos de caixa em geral são fatores determinantes da recuperação judicial; e, (ii) modelos construídos por meio da conjugação de informações contábeis, de mercado e de dados macroeconômicos apresentam desempenho superior às abordagens tradicionais, o que inclui não apenas ferramentas de inteligência artificial mais sofisticadas, como também indicadores menos ortodoxos, como os fluxos de caixa.

Em se tratando de uma amostra viesada, cujas observações foram obtidas a partir de dados primários, os resultados não são abrangentes; outra dificuldade própria das RNA é a compreensão de como as diversas variáveis se comportam dentro da rede, que raramente tem grande capacidade explicativa, muito embora haja significativa contribuição metodológica (Trippi & Lee, 1996), pois teve-se grande preocupação na definição do evento de interesse.

Apesar dessas limitações, os resultados obtidos contam com inúmeras aplicações para o mercado e a academia, inclusive fomentando pesquisas futuras, na medida em que a recuperação judicial tem crescido não apenas no Brasil; novos trabalhos para continuar aprimorando técnicas de previsão, superando as lacunas apontadas, certamente encontrarão o seu lugar.

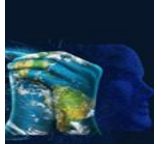
## Referências

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *The Journal of Finance*, 23(4), 589–609.

\_\_\_\_\_. (1983). *Corporate financial distress: a complete guide to predicting, avoiding and dealing with bankruptcy*. New York: John Wiley & Sons.

Altman, E. I., Baidya, T. K. N. & Dias, L. M. R. (1979). Previsão de problemas financeiros em empresas. *Revista de administração de empresas*, 19(1), 17-28.

Azayite, F. Z. & Achchab, S. (2016). Hybrid discriminant neural networks for bankruptcy prediction and risk scoring. *Procedia Computer Science*, 83, 670-674.



Aziz, A., Emanuel, D.C. & Lawson, G.H. (1988). Bankruptcy prediction - an investigation of cash flow-based models. *Journal of Management Studies*, 25(5), 419-437.

Aziz, M. A. & Dar, H. A. (2006). Predicting corporate bankruptcy: where we stand?. *Corporate Governance*, 6(1), 18-33.

Baird, D. G. (1991). The initiation problem in bankruptcy. *International Review of Law and Economics*, 11(2), 223-232.

Bialoskorski Neto, S. Nagano, M. S. & Moraes, M. B. C. (2006). Utilização de redes neurais artificiais para avaliação socioeconômica: uma aplicação em cooperativas. *Revista de Administração da USP*, 41(1), 59-68.

Brockett, P. L., Cooper, W. W., Golden, L. L. & Pitaktong, U. (1994). A neural network method for obtaining an early warning of insurer insolvency. *The Journal of Risk and Insurance*, 61(3), 402-424.

Casey, C. & Bartczak, N. (1985). Using operating cash flow to predict financial distress: some extensions. *Journal of Accounting Research*, 23(1), 384-401.

Crouhy, M., Galai, D. & Mark, R. (2000). A comparative analysis of current credit risk models. *Journal of Banking and Finance*, 24(1-2), 59-117.

Decreto-Lei n. 7.661 (1945, 21 de junho). Lei de Falências. Brasília, DF: Presidência da República.

Dickinson, V. (2011). Cash flow patterns as a proxy for firm life cycle. *The Accounting Review*, 86(6), 1969-1994.

Elizabetsky, R. (1976). *Um modelo matemático para decisões de crédito no banco comercial*. Dissertação de Mestrado, Escola Politécnica, Universidade de São Paulo, São Paulo, São Paulo.

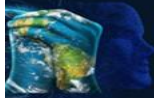
Evans, J. & Borders, A. L. (2014). Strategically surviving bankruptcy during a global financial crisis: the importance of understanding chapter 15. *Journal of Business Research*, 67(1), 2738-2742.

Gilson, S. C. (1991). Managing default: some evidence on how firms choose between workouts and Chapter 11. *Journal of Applied Corporate Finance*, 4(2), 62-70.

Hall, M. A. (1998). *Correlation-based Feature Subset Selection for Machine Learning*. Tese de Doutorado. University of Waikato, New Zealand.

Jardin, P. D. (2017). Dynamics of firm financial evolution and bankruptcy prediction. *Expert Systems with Applications*, 75(1), 25-43.

Jardin, P. D. & Séverin, E. (2011). Predicting corporate bankruptcy using a self-organizing map: an empirical study to improve the forecasting horizon of a financial failure model. *Decision Support Systems*, 51(3), 701-711.



Jones, S. (2017). Corporate bankruptcy prediction: a high dimensional analysis. *Review of Accounting Studies*, 22(3), 3, 1366-1442.

Jones, S., Johnstone, D. & Wilson, R. (2017). Predicting corporate bankruptcy: an evaluation of alternative statistical frameworks. *Journal of Business Finance & Accounting*, 44(1-2), 3–34.

\_\_\_\_\_. (2015). An empirical evaluation of the performance of binary classifiers in the prediction of credit ratings changes. *Journal of Banking and Finance*, 56, 72–85.

Kanitz, S. C. (1978). *Como prever falências*. São Paulo: McGraw do Brasil.

Khashei, M. & Bijari, M. (2010). An artificial neural network (p, d, q) model for timeseries forecasting. *Expert Systems with Applications*, 37(1), 479-489.

Lei n. 6.404, de 15 de dezembro de 1976. (1976, 15 de dezembro). Dispõe sobre as Sociedades por Ações. Diário Oficial da União, suplemento.

Lei n. 9.430, de 27 de dezembro de 1996. (1996, 27 de dezembro). Dispõe sobre a legislação tributária federal, as contribuições para a seguridade social, o processo administrativo de consulta e dá outras providências. Diário Oficial da União, edição extra.

Lei n. 11.101, de 09 de fevereiro de 2005. (2005, 09 de fevereiro). Regula a recuperação judicial, a extrajudicial e a falência do empresário e da sociedade empresária. Diário Oficial da União, edição extra.

Lei n. 11.638, de 28 de dezembro de 2007. (2007, 28 de dezembro). Altera e revoga dispositivos da Lei no 6.404, de 15 de dezembro de 1976, e da Lei no 6.385, de 7 de dezembro de 1976, e estende às sociedades de grande porte disposições relativas à elaboração e divulgação de demonstrações financeiras. Diário Oficial da União, edição extra.

Lei n. 14.112, de 24 de dezembro de 2020 (2020, 24 de dezembro). Altera as Leis n os 11.101, de 9 de fevereiro de 2005, 10.522, de 19 de julho de 2002, e 8.929, de 22 de agosto de 1994, para atualizar a legislação referente à recuperação judicial, à recuperação extrajudicial e à falência do empresário e da sociedade empresária. Diário Oficial da União, edição extra.

Lucas, D. J. (1995). Default correlation and credit analysis. *Journal of Fixed Income*, 76-87.

Matias, A. B. (1978). *Contribuição às técnicas de análise financeira: um modelo de concessão de crédito*. Dissertação de Mestrado, Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo.

Ohlson, J. A & Aier., J. K. (2009). On the analysis of firms' cash flows. *Contemporary Accounting Research*, 26(4), 1091-1114.

Pereira, V. S. & Martins, V. F. (2015). Estudos de previsão de falências – uma revisão das publicações internacionais e brasileiras de 1930 a 2015, *Revista de Contabilidade Contemporânea*, 12(26), 163-196.



Rosa, P. S. & Gartner, I. R. (2018). Financial distress in Brazilian banks: an early warning model. *Revista Contabilidade & Finanças*, 29(77), 312-331.

Scabora, Filipe Casellato. (2023). Recuperação Judicial e Inteligência Artificial: previsão de insolvência por meio de redes neurais artificiais. In Pleti, E. L. F. et al. (Orgs.). *Direito Digital e Desenvolvimento: olhares transdisciplinares sobre a efetivação de direitos*. Uberlândia: Marco Teórico.

Serasa. (2025). *Serasa Experian de Falências e Recuperações*. São Paulo: Autor.

Takacs, L. A. (1985). A two-step scheme for the advection equation with minimized dissipation and dispersion errors. *Monthly Weather Review*, 113, 1050-1065.

Teixeira, T. (2012). Recuperação judicial de empresas. *Revista da Faculdade de Direito da Universidade de São Paulo*, 106/107, 181-214.

Tinoco, M. H. & Wilson, N. (2103). Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables. *International Review of Financial Analysis*, 30, 394-419.

Trippi, R. R. & Lee, J. K. (1996). *Artificial intelligence in finance & investing: state-of-the-art technologies for securities selection and portfolio management*. Chicago: Irwin.

Wruck, K. (1990). Financial distress, reorganization, and organizational efficiency. *Journal of Financial Economics*, 27(2), 419-444.

Yu, Q., Miche, Y., Séverin, E. & Lendasse, A. (2014). Bankruptcy prediction using extreme learning machine and financial expertise. *Neurocomputing*, 128, 296-302.

Zhanga, G., Hub, M. Y., Patuwob, B. E. & Indro, D. C. (1999). Artificial neural networks in bankruptcy prediction: general framework and cross-validation analysis. *European Journal of Operational Research*, 116(1), 16-32.

Zumel, N. & Mount, J. (2014). *Practical data science with R*. New York: Manning.