



Revista Universo Contábil, ISSN 1809-3337
Blumenau, 2022, v. 18: e2022102, p. 01-21

doi:10.4270/ruc.2022102

Disponível em www.furb.br/universocontabil



INSOLVÊNCIA EMPRESARIAL: Estudo bibliométrico nacional e internacional com suporte de mineração de informações textuais

BUSINESS INSOLVENCY: National and international bibliometric study supported by text mining

INSOLVENCIA EMPRESARIAL: Estudio bibliométrico nacional e internacional apoyado en Minería textual

Recebido em: 01-07-2021

Avaliado em: 25-03-2022

Reformulado em: 13-07-2022

Aceito para publicação em: 07-02-2023

Publicado em: 04-04-2023

Editor Responsável: Moacir Manoel Rodrigues Jr.

Lenio Vitor Oliveira Menezes¹
José Sérgio Casé de Oliveira²

RESUMO³

Este artigo teve como objetivo principal compreender e comparar os temas mais abordados na literatura nacional e internacional sobre a insolvência empresarial. Para tanto fez-se uso de uma adaptação da técnica para sistematização da seleção de trabalhos acadêmicos, chamada de *Knowledge Development Process-Constructivist* em conjunto com técnicas de mineração de informações auxiliada pelo *Software Iramuteq*. Foi identificada expansão na produção acadêmica nacional e internacional após o ano 2008, coincidindo com a crise do *subprime* norte americana com reflexos no mundo inteiro. Na análise final de 1.392 artigos internacionais e 154 nacionais foram identificadas divergências entre os temas abordados pela literatura nacional e internacional. A literatura nacional mostrou forte interesse em compreender o processo falimentar de pequenas e médias empresas, utilizando, principalmente, técnicas estatísticas e entrevistas, enquanto que a literatura internacional se mostrou bastante heterogênea nos segmentos estudados, fazendo uso de técnicas estatísticas, mas também mostrando interesse por técnicas de aprendizado de máquina. Como contribuição acadêmica, este trabalho sintetizou as principais linhas de estudo desenvolvidas em insolvência empresarial, uma área do conhecimento de interesse multidisciplinar e que ainda não possui uma teoria conclusiva. O trabalho contribui, também, como fonte para a identificação de lacunas de pesquisa, ou seja, na identificação de discussões ainda pouco exploradas no cenário nacional como o relacionamento entre gerenciamento de resultados e insolvência empresarial e a utilização, com maior consistência, de técnicas de *machine learn* como: Máquina de Vetores de Suporte e Redes Neurais Artificiais.

Palavras-chave: Insolvência empresarial. Bibliometria. ProKnow-C. Mineração de informações.

¹ Mestrando em Ciências Contábeis pela Universidade Federal da Bahia (UFBA); ORCID: <https://orcid.org/0000-0001-9061-9381>; E-mail: leniomenezes@gmail.com

² Doutor em Economia pela Universidade Federal da Paraíba; Professor da Faculdade de Ciências Contábeis da Universidade Federal da Bahia (UFBA); ORCID: <https://orcid.org/0000-0003-3031-4943>; E-mail: js_cdo@hotmail.com

³ O presente artigo foi desenvolvido com apoio da FAPESB através da concessão de bolsa de estudo de Mestrado, nº de processo 00017546048.

ABSTRACT

The main objective of this article was to understand and compare the most discussed topics in the national and international literature on business insolvency. It used a technique for systematizing the selection of academic papers, the Knowledge Development Process-Constructivist, and text mining techniques supported by the Iramuteq software. We identified an expansion in national and international academic production after 2008, coinciding with the North American subprime mortgage crisis reflected worldwide. We identified differences between the topics covered by national and international literature. The national literature showed a greater interest in understanding the bankruptcy process of small and medium enterprises, using mainly statistical techniques and interviews, while the international literature proved to be quite heterogeneous in the segments studied, using statistical techniques but also manifesting an interest in machine learning techniques. As an academic contribution, this work synthesized the main lines of study developed in business insolvency, an area of knowledge of multidisciplinary interest that still does not have a conclusive theory. The work also contributes as a source for the identification of research gaps, that is, in the identification of discussions still little explored in the national scenario, such as the discussion of the relationship between earnings management and corporate insolvency and the use, with greater consistency, of machine learning techniques such as Support Vector Machine and Artificial Neural Networks.

Keywords: Business insolvency. Bibliometry. ProKnow-C. Information mining.

RESUMEN

El objetivo principal de este artículo es comprender y comparar los temas más discutidos en la literatura nacional e internacional sobre la insolvencia empresarial. Para ello, se hizo uso de una técnica de sistematización de la selección de trabajos académicos, denominada Knowledge Development Process-Constructivist junto con técnicas de minería de información asistidas por el software Iramuteq. Se identificó una expansión en la producción académica nacional e internacional a partir de 2008, coincidiendo con la crisis subprime norteamericana que se reflejó a nivel mundial. Se identificaron diferencias entre los temas cubiertos por la literatura nacional e internacional. La literatura nacional mostró un fuerte interés en comprender el proceso concursal de las pequeñas y medianas empresas, utilizando principalmente técnicas estadísticas y entrevistas, mientras que la literatura internacional demostró ser bastante heterogénea en los segmentos estudiados, haciendo uso de técnicas estadísticas, pero también mostrando interés en las técnicas de aprendizaje automático. Como aporte académico, este trabajo sintetizó las principales líneas de estudio desarrolladas en la insolvencia empresarial, un área de conocimiento de interés multidisciplinario que aún no cuenta con una teoría concluyente. El trabajo también contribuye como fuente para la identificación de vacíos de investigación, o sea, en la identificación de discusiones aún poco exploradas en el escenario nacional, como la discusión de la relación entre la gestión de ganancias y la insolvencia empresarial y el uso, con mayor consistencia, de técnicas de aprendizaje automático como Support Vector Machine y Artificial Neural Networks.

Palabras-clave: Insolvencia empresarial. Bibliometría. ProKnow-C. Minería de información.

1 INTRODUÇÃO

A compreensão das condições financeiras de uma empresa é essencial tanto para os seus *stakeholders* quanto para a sociedade. Por *stakeholders* pode-se compreender as partes interessadas em um negócio, tais como funcionários, sócios, acionistas, governo entre outros grupos. Compreender

e avaliar o desenvolvimento científico em nível nacional e internacional desta área se mostra cada vez mais relevante à medida que pode ajudar a direcionar os esforços acadêmicos no sentido de atender as demandas informacionais da sociedade a respeito da falência empresarial.

Como afirmam Mansi, Maxwell e Zhang (2010) estudar a condição financeira de uma empresa é essencial para os investidores e credores, visto que eles poderiam determinar com maior confiabilidade os custos de capital diante dos riscos envolvidos caso a empresa entre em processo de falência.

Os primeiros estudos datam da década de 1930. Mas apesar do largo histórico de pesquisas acerca do processo de falência empresarial, Pinheiro e Pinheiro (2008) assim como, Wang, Ma e Yang (2014), afirmam não existir uma teoria positiva ou definitiva sobre a falência corporativa. Sendo assim, não existe um consenso a respeito da Insolvência empresarial e do processo falimentar, especialmente, por se tratar de um tema interdisciplinar no qual envolve os interesses de diversos atores e setores.

De acordo com dados divulgados pelo Serasa *Experian* (2022), no Brasil, foram requeridas 10.542 falências e 9.825 recuperações judiciais entre o período janeiro de 2015 até junho de 2022. Vale ressaltar que no contexto nacional 53,79% das falências foram solicitadas por micro e pequenas empresas e 46,21% solicitadas por empresas de médio e grande porte. Já as solicitações por recuperações judiciais são compostas em 61,01% de micro e pequenas empresas. Analisando exclusivamente a perspectiva dos empregos gerados, no ano de 2021, 78% dos empregos formais foram provenientes de pequenos negócios, de acordo com o Cadastro Geral de Empregados e Desempregados (Caged), de acordo com dados divulgados pela FENACON (2022). Diante desse contexto, é possível perceber o impacto social que a insolvência empresarial pode gerar. Contudo, cabe ressaltar que existem outras partes interessadas na manutenção da capacidade financeira dos negócios, como governos, fornecedores, investidores, entre outros.

Em torno da atividade empresarial se concentram diversos *stakeholders*, cada um deles com suas próprias necessidades. O grande número de partes interessadas nas atividades empresariais se traduz em uma infinidade de perspectivas de análise.

Sabendo que existe uma multiplicidade nas formas de analisar o fenômeno da insolvência empresarial e compreendendo que a literatura nacional pode não estar alinhada com as discussões e abordagens metodológicas apresentadas na literatura internacional desenvolveu-se a seguinte questão: Como estão estruturadas as pesquisas sobre insolvência empresarial a nível nacional e internacional? Com este questionamento em mente, traçou-se o objetivo de compreender quais são os temas mais abordados em nível nacional e internacional.

O presente estudo se justifica, inicialmente, pela necessidade de empreender esforços no intuito de consolidar o conhecimento científico a respeito da insolvência empresarial. Ademais, o artigo, ao aplicar técnicas estatísticas para a análise de conteúdo, se mostra como o primeiro trabalho bibliométrico nessa temática a propor extrapolar a consolidação da discussão e adentrar a análise de conteúdos de forma sistemática e não enviesada. Por fim, o artigo se propõe a comparar a literatura nacional e internacional para identificar pontos de convergência/divergência, permitindo, assim identificar oportunidades de pesquisa.

Este artigo tem como estrutura além desta seção de introdução, a revisão de literatura apresentada na segunda seção, a metodologia utilizada na terceira seção, os resultados serão apresentados na quarta seção, na quinta seção serão apresentadas as discussões, as conclusões, na sexta seção e, por fim as referências utilizadas.

2 REFERENCIAL TEÓRICO

Diferentes definições para insolvência foram apresentadas por Beaver (1966), Altman (1968) e Janot (2001), entre outros autores. Beaver (1966) no clássico trabalho, “*Financial Ratios as Predictors of Failure*” trabalha com o conceito de *default*, onde o mesmo é definido como a incapacidade da empresa liquidar suas obrigações financeiras. O autor estabelece os eventos que

indicam o *default*, pontuando-os como: falência, inadimplência de títulos, uma conta bancária sacada ou não pagamento de dividendos à ações preferenciais.

Para Altman (1968), a insolvência ocorre quando a avaliação justa dos ativos é inferior ao total de passivos. Janot (2001) ao abordar a insolvência bancária, também defende a relação entre ativos e passivos, concluindo que é declarada insolvente a empresa que apresentar Patrimônio Líquido negativo, ou seja, que os Passivos totais superem os Ativos Totais.

De acordo com Wruck (1990), sob uma perspectiva econômico-contábil, é possível considerar que existem duas grandes frentes para conceituar insolvência. A primeira ligada a fluxos de caixa, e conseqüentemente, a capacidade empresarial de contrair dívidas e arcar com elas. E, a segunda ligada ao Balanço Patrimonial. Em síntese a insolvência pode ser dividida como insolvência técnica, também conhecida como insolvência baseada em fluxo, ou seja, aquela que ocorre pela incapacidade da empresa cumprir suas obrigações, portanto relacionada à capacidade de geração de caixa. E, a insolvência contábil ou insolvência de saldos, aquela que ocorre quando os registros contábeis apresentam um Patrimônio Líquido negativo.

Diante da inexistência de uma definição concreta e conclusiva sobre insolvência como afirmam Pinheiro e Pinheiro (2008) e Wang et al., (2014), e visando compreender as relações estabelecidas entre obras, surgem os trabalhos bibliométricos como tentativas de consolidar esta área do conhecimento.

Araújo (2006) define a bibliometria com o uma técnica quantitativa e estatística de medição de índices de produção e disseminação do conhecimento. Ensslin et al. (2010) complementam ao afirmar que a bibliometria é o processo de evidenciação quantitativa dos dados estatísticos de um conjunto definido de artigos com o objetivo de gestão da informação e do conhecimento científico a respeito de um assunto.

Prado et al. (2016) desenvolveram um estudo bibliométrico visando identificar a utilização de técnicas de análise multivariada de dados em pesquisas sobre risco de crédito e falência. O período estudado englobou os anos de 1968 a 2014. O levantamento feito pelos autores utilizou a base *Web of Science* da Thomson Reuters. Os autores observaram que houve expansão na produção acadêmica voltada para insolvência empresarial. Quando comparados os anos de 2008 e 2009 identifica-se a elevação de 16 para 40 publicações. Destaca-se que este volume de publicações mais elevado se mantém nos anos seguintes, merecendo destaque o ano de 2011 que apresentou o maior volume de publicações (54) dentro do período de estudo.

Führ et al. (2017) buscaram fazer uma revisão da literatura a respeito da classificação de crédito. Utilizando as bases *Web of Science* e *Scopus*, aliadas a técnica Proknow-C foram feitas as análises bibliométricas, dando especial atenção a técnicas classificação de Crédito para pequenas e médias empresas. Os autores também perceberam o aumento na produção após a crise de 2008, destacando que a revista com maior número de publicações foi a *Expert Systems with Applications*. Identificaram ainda que as técnicas mais utilizadas foram técnicas estatísticas como Regressão Logística e Análise discriminante, mas pontuam que as técnicas de aprendizado de máquina estão ganhando espaço na produção acadêmica internacional.

Kücher e Feldbauer-Durstmüller (2018) tiveram como objetivo obter uma visão geral, consolidar a pesquisa existente e visualizar os principais tópicos e conceitos a respeito da insolvência empresarial. Os artigos levantados foram sintetizados em 7 temáticas de pesquisa. Destas, foram destacadas 3, a primeira delas denominada *Organizational ecology and industrial organization*, caracterizada por apresentar uma literatura com uma perspectiva determinística; a segunda vertente denominada *Entrepreneurial failure*, abordou a temática relacionada aos efeitos posteriores da falência, tratando especialmente das conseqüências financeiras, sociais e psicológicas; por fim, a terceira temática foi denominada de *Organizational decline and strategic choice*, a qual caracterizou-se por adotar uma perspectiva voluntaristas abordando características psicológicas e organizacionais. Os autores concluíram que as classes 1 e 3 abordam as razões e processos que levaram a falência, já a classe 2 abordou as conseqüências e custos envolvidos.

Shi e Li (2019) desenvolveram um trabalho com o objetivo de explorar a aplicação de técnicas inteligentes nas previsões de falências. Para tanto, fizeram a análise de artigos desenvolvidos entre os anos de 1968 e 2018 localizados na *Web of Science*. Como resultado foi identificado que as revistas mais produtivas foram a *Expert Systems with Applications* seguida da *European Journal of Operational Research*. Os autores identificaram que os trabalhos mais citados são Altman (1968), Ohlson (1980) e Beaver (1966). Também foi percebido o crescimento no número de publicações após o ano de 2008. Cabe destacar que o estudo de Shi e Li (2019) consegue demonstrar que além das técnicas clássicas de regressão estatística, os estudos de insolvência empresarial têm aplicado técnicas recentes de *machine learning* para solucionar problemas de pesquisa.

Um ponto que deve ser ressaltado é que não foi identificado nos trabalhos, especialmente nos bibliométricos nacionais, o esforço no sentido de comparar o que é pesquisado no âmbito nacional e internacional.

É importante notar que os estudos bibliométricos contribuem para consolidar o conhecimento dentro de uma linha de pesquisa. Os estudos bibliométricos a respeito de insolvência empresarial são especialmente importantes, pois como afirma Kücher e Feldbauer-Durstmüller (2018), trata-se de uma linha de pesquisa fortemente fragmentada e com múltiplos interesses acadêmicos. No tocante aos resultados identificados pela literatura, é válido destacar a visibilidade que a temática ganhou especialmente após a crise dos *subprimes* norte americana.

3 METODOLOGIA

Para alcançar os objetivos propostos, utilizou-se uma adaptação da metodologia denominada *Knowledge Development Process-Constructivist* (ProKnow-C) com o intuito de auxiliar no levantamento, seleção e coleta dos artigos. Para o gerenciamento e organização dos artigos, os *softwares* escolhidos foram o Endnote, o Zotero e o Excel. Já para a análise estatísticas dos dados coletados foi utilizado o *software* Iramuteq.

O Endnote e Zotero são gerenciadores de referências bibliográficas desenvolvido, sendo o primeiro desenvolvido pela Thomson Reuters e o segundo um projeto desenvolvido pelo *Roy Rosenzweig Center for History and New Media da George Mason University*. A escolha por dois gerenciadores de referências se justifica pela eficiência na importação de referências apresentada pelo Endnote e posterior exportação para o Zotero. Este último utilizado por ser mais bem otimizado para exportação de informações para o Excel. O Excel é um editor de planilhas eletrônicas, enquanto que o Iramuteq é um *software* estatístico baseado na linguagem de programação R que permite a produção e interpretação de estatísticas textuais.

3.1 Metodologia ProKnow-C

O instrumento utilizado para coleta e seleção dos artigos foi uma adaptação do método ProKnow-C. Esta metodologia foi desenvolvida em 2005, pelo Laboratório de Metodologias Multicritério em Apoio à Decisão – LabMCDA, da Universidade Federal de Santa Catarina. O Proknow-C aplicado de forma completa segue o seguinte processo: 1) Seleção de portfólio a respeito do tema de pesquisa; 2) Análise bibliométrica do portfólio; 3) Análise sistemática do portfólio; 4) Estabelecimento dos objetivos da pesquisa. A metodologia completa tem como objetivo guiar o pesquisador até a formulação dos objetivos de pesquisa por meio de um levantamento e análise sistemática da literatura. Como o objetivo deste trabalho é compreender os temas abordados quando se discute a respeito da insolvência através de análises textuais, utilizou-se as etapas 1 e 2 por se referirem ao levantamento sistemático da bibliografia. Portanto, as etapas de análise sistemática do portfólio (3) e estabelecimento dos objetivos da pesquisa (4) não foram aplicadas neste trabalho.

A etapa de seleção de portfólio (1), para Ensslin, Ensslin e Pinto (2013), é desenvolvida em três fases: a primeira é a seleção de artigos científicos nos bancos de trabalhos acadêmicos como o objetivo de compor o Banco de Artigos Bruto; a segunda fase é caracterizada pela filtragem dos

artigos; já a terceira fase é o teste de representatividade do levantamento, conforme descritos a seguir. O resultado final deste processo é Portfólio Bibliográfico (BP).

3.1.1 Seleção de Banco de Artigos Bruto

Em consonância com o objetivo deste artigo delineou-se apenas um único eixo de pesquisa para que os testes estatístico-textuais pudessem agrupar o eixo em subtemas. O instrumento ProKnow-C propõe quatro fases para a formulação do Banco de Artigos brutos: (a) seleção de palavras-chave; (b) seleção das bases de dados; (c) coleta dos artigos nas bases de busca; e (d) aplicação do teste de aderência das palavras-chave.

a) Seleção de palavras-chave

A seleção dos termos de busca dos artigos internacionais tomou como referência os termos utilizados no trabalho de Prado et al. (2016), diferenciando-se deste pela divisão dos termos *Logistic regression/Logit* em dois termos distintos. A opção pela segregação dos termos foi feita como forma de ampliar os termos de busca visando abarcar o maior quantitativo de artigos publicados. A pesquisa foi filtrada por cinco termos diferentes com base no título do artigo e dezessete termos com base no tópico (título, resumo ou palavras-chave especificadas pelo autor) abordado pelos artigos. Este levantamento resultou em 85 combinações de busca. Para os títulos foram utilizados os seguintes termos: *Bankruptcy, Business failure, Credit risk, Credit rating e Credit scoring*. Já para os tópicos utilizou-se os termos a seguir: *Conjoint analysis, Cluster analysis, Canonical correlation, Correspondence analysis, Discriminant analysis, Factor analysis, Principal componentes, Multivariate analysis of variance, Multivariate analysis of covariance, Multidimensional scaling, Perceptual mapping, Fuzzy logic, Linear probability models, Structural equation, Neural network, Logistic regression, Logit e Multiple regression*.

Para a seleção dos termos a serem empregados no levantamento dos artigos nacionais utilizou-se, inicialmente, a livre tradução dos termos utilizados nos trabalhos de Prado et al. (2016), Führ et al. (2017), Kücher e Feldbauer-Durstmüller (2018) e Shi e Li (2019). Portanto, além da tradução dos termos acima mencionados foram acrescentadas as traduções para os seguintes termos: *Credit Scoring, Bankruptcy Prediction, Default Prediction, Financial Failure, Financial Distress e Insolvency*. Esta seleção foi posteriormente incrementada com variações e sinônimos dos termos visando contemplar o refinamento linguístico das línguas de origem latina. Por exemplo, para o termo falência empresarial, buscou-se também os seguintes termos: falência de negócios, falência de companhias, falência corporativa, falência de empreendimentos e falência de firmas.

Como forma de complementar o levantamento foram acrescentados, ainda, as livres traduções e variações de termos encontrados em artigos relevantes tanto nacionais quanto internacionais na área, tais como: preditores de falha (Beaver, 1966); previsão insolvência (Horta, 2001; Janot, 2001; Matias & Siqueira, 1996; Silva, 1983); análise de insolvência (Bressan, Braga, & Lima, 2004); prever falência (Kanitz, 1974); modelo decisão de crédito (Elizabetsky, 1976); modelo de concessão de crédito (Matias, 1978); Análise discriminante bancos (Martins, & Samané, 1987); Procedimento decisão crédito (Securato, & Famá, 1997); insolvência sistema financeiro (Cabrera, 1998); Previsão solvência (Minussi, Damacena, & Ness, 2002); e Previsão concordata (Sanvicente, & Minardi, 1998). As preposições foram suprimidas com o objetivo de evitar que os resultados fossem restringidos por elas. Foi acrescentada ainda uma pesquisa livre pelo termo Modelo insolvência empresarial como forma de levantar artigos que não contivessem os termos citados anteriormente em seus títulos. Deste incremento resultou o total de 94 termos buscados exclusivamente pelo título, visto que a base de busca utilizada para os trabalhos nacionais não permite a busca por tópicos.

b) Seleção das bases de dados

Foram utilizadas duas bases de dados diferentes. Uma definida para os levantamentos de artigos nacionais e outra para o levantamento de artigos internacionais. Os artigos nacionais foram

levantados no banco de trabalhos acadêmicos livres Google Acadêmico – *Scholar*. A escolha desta plataforma de pesquisa se deu pela amplitude possibilitada na busca, visto que os resultados apresentados não ficariam restritos pela indexação a uma base específica.

Já para os artigos internacionais houve a necessidade de restringir a busca a uma base específica em função da grande quantidade de trabalhos identificados na plataforma *Scholar* que totalizou 38.619 artigos aptos a compor o banco de artigos bruto. Os artigos internacionais, portanto, foram coletados no banco de trabalhos acadêmicos da Elsevier/Scopus. A escolha deste banco de dados se deu pela necessidade de filtrar a grande quantidade de artigos por uma base confiável e que facilitasse a localização dos trabalhos internacionais.

c) Coleta dos artigos nas bases de busca

O levantamento inicial trouxe como resultado o total de 3.292 artigos internacionais e 1.208 artigos nacionais. Entretanto, neste quantitativo constam, ainda, trabalhos em duplicidade, referentes a outras áreas do conhecimento e especificamente na base nacional constam, também, trabalhos de conclusão de curso que foram retirados posteriormente na etapa de filtragem do banco de dados bruto. O banco de artigos bruto contou, inicialmente, com o total de 3.500 artigos nacionais e internacionais levantados. Destaca-se que banco bruto ainda não passou pelo processo de filtragem e de exclusão de trabalhos em duplicidade e conferência do alinhamento dos títulos com o tema que será detalhado a seguir.

d) Teste de aderência das palavras-chave

Em posse do Portfólio Bruto, foi efetuado o teste de aderência das palavras-chave com o objetivo de verificar a necessidade de inclusão de outros termos de busca. Foram escolhidos 347 artigos aleatoriamente para verificar se havia a necessidade de incluir novos termos de busca. Conclui-se que não havia a necessidade de fazer a inclusão de novos termos.

3.1.2 Filtragem do banco de dados bruto

A segunda etapa para a formulação do Portfólio Bibliográfico é a filtragem do Banco de Artigos bruto. Esta etapa é composta originalmente por 05 processos: a) exclusão de artigos em duplicidade; b) Conferência do alinhamento dos títulos com o tema; c) Relevância científica dos artigos; d) Conferência do alinhamento dos resumos com o tema; e e) Disponibilidade dos artigos completos.

Neste artigo, não foram utilizadas as etapas de análise da relevância científica e a disponibilidade dos artigos completos. O processo de análise da relevância científica dos artigos toma como base o número de citações em outras obras, apesar de recomendar o mínimo de 85% como ponto de corte, o ProKnow-C não indica expressamente percentuais a serem adotados como pontuou Castelli (2018). Diante disso e alinhado com o objetivo deste artigo, foi decidido que nenhum artigo que estivesse alinhado com o tema abordado seria excluído para que assim fosse possível traçar e comparar os temas mais trabalhados em âmbito nacional e internacional e não apenas aqueles mais referenciados e/ou relevantes.

A análise de disponibilidade dos artigos completos não se mostrou relevante para as análises bibliométricas efetuadas neste trabalho, visto que o arquivo textual a ser analisado foi composto pelos resumos dos artigos selecionados. Assim, foram excluídos apenas os artigos que não possuísem acesso aos seus respectivos resumos.

Após a exclusão de artigos em duplicidade, trabalhos não publicados e artigos sem aderência com o tema, a seleção trouxe como resultado 1.392 artigos internacionais e 154 nacionais aptos a compor o Portfólio Bibliográfico. Tendo aproveitamento de 42,28% para o banco internacional e 12,75%, para o banco nacional com relação ao Portfólio Bibliográfico Bruto.

3.2 Processo de mineração de informações

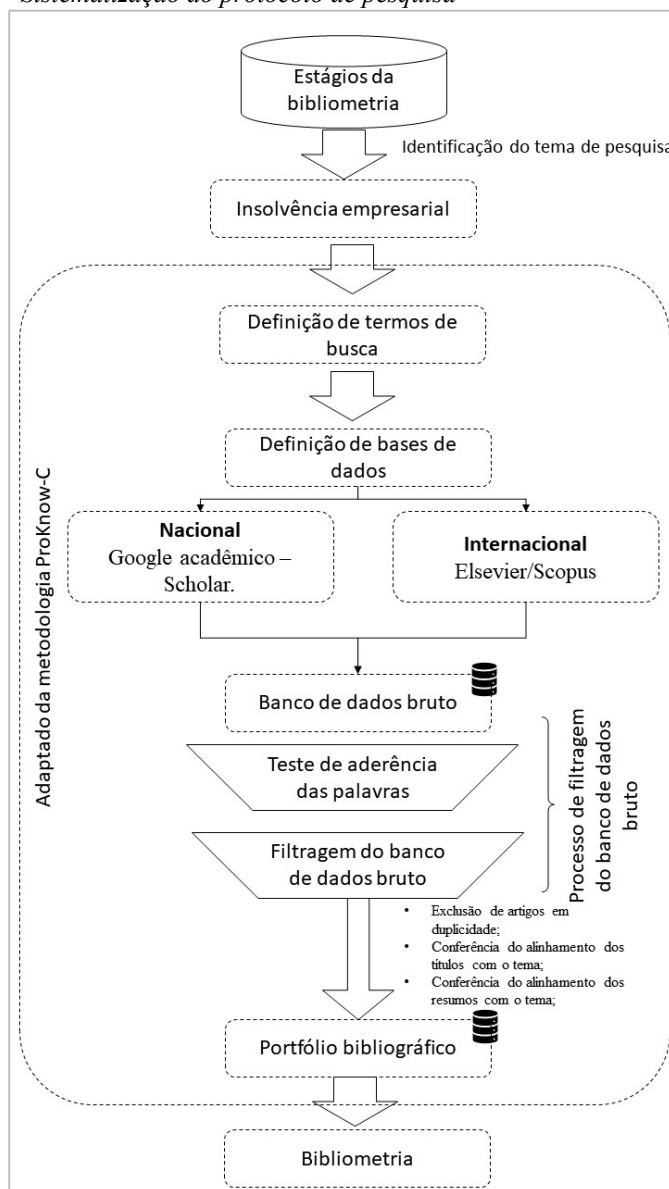
Para a execução da mineração das informações foi utilizado o *software* Iramuteq. O mesmo é capaz de fazer análises de estatísticas textuais tanto simples quanto complexas, para isso é utilizado o *software* estatístico R como ambiente de suporte. Podendo extrair as estatísticas textuais e a Classificação Hierárquica Descendente (CHD).

A Classificação Hierárquica Descendente permite a identificação de classes de termos que possuem relações estatísticas significantes entre si. Esta ferramenta foi utilizada neste artigo para identificar a quantidade de conteúdos abordados nos artigos de insolvência empresarial.

No Iramuteq o *Corpus* textual (conjunto dos textos a serem analisados) é fracionado em Segmentos de Texto (ST) como base em critério definidos pelo pesquisador. O *software* utiliza a lematização para fazer análises, ou seja, os termos são analisados com base em sua raiz, descartando conjugações, gênero e plural. A Figura 1 apresenta o protocolo de pesquisa adotado neste artigo.

Figura 1

Sistematização do protocolo de pesquisa



Para que pudesse ser feita a mineração das informações textuais foi feita a compilação de resumos dos artigos selecionados em dois arquivos de texto (*corpus* textual): o primeiro contendo os

resumos dos artigos nacionais; e o segundo, os resumos dos artigos internacionais, objetivando, posteriormente, efetuar as análises textuais. Nestes arquivos, foram feitas correções gramaticais; correções de partículas apassivadoras (-se); e união de grupos de palavras que apresentam apenas um sentido, tais como Redes Neurais, *Logistic Regression* e São Paulo que foram substituídas, respectivamente por Redes_Neurais, *Logistic_Regression* e São_Paulo.

4 RESULTADOS

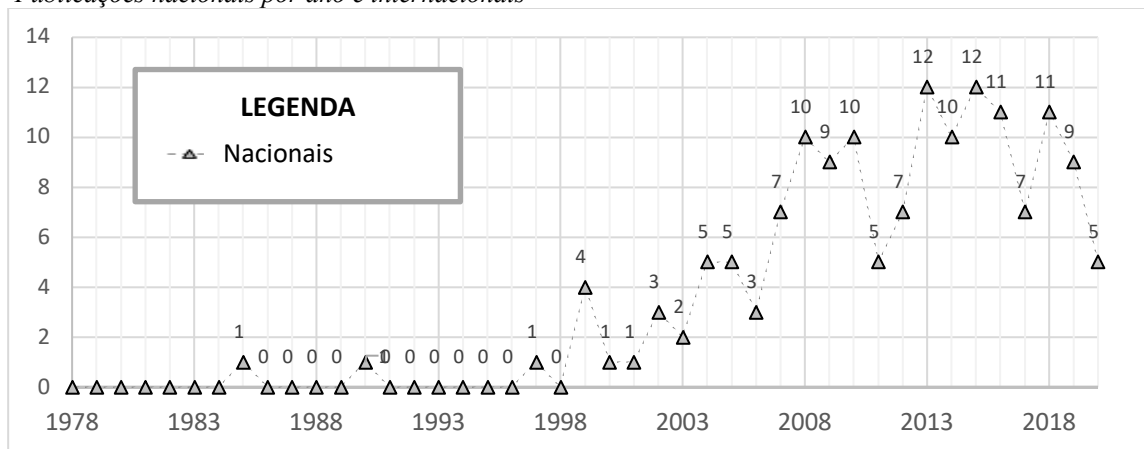
Em posse do Portfólio Bibliográfico nacional e internacional foi feita a análise dos trabalhos para identificar o volume de publicações por ano e a distribuição de publicações por periódico, bem como a mineração das informações contidas nos resumos dos artigos, de onde se obteve os resultados a seguir apresentados.

4.1 BIBLIOMETRIA NACIONAL

No Brasil, o levantamento mostrou como primeiro trabalho disponível o artigo “Insolvência de empresas e política macroeconômica”, datado de 1985 e escrito por Cláudio Roberto Contador. Depois houve um hiato de tempo de 05 anos sendo a próxima publicação apenas no ano 1990 com o artigo “Um modelo de análise de insucesso de empresas”, escrito por José Rigoberto Parada Daza. A produção nacional começa ter maior robustez a partir dos anos 2000, tendo uma elevação após a crise de 2008, conhecida como crise do *subprime*. Contudo, observa-se que após o ano de 2015 o volume de publicações se mostra em declínio, como pode ser observado na Figura 2.

Figura 2

Publicações nacionais por ano e internacionais



Três revistas assumem a dianteira no número de publicações nacionais, a Revista Contabilidade & Finanças, a Revista de Economia e Administração e Revista Universo Contábil. A listagem dos 20 periódicos que mais publicam na área é apresentada na Figura 3.

A análise textual dos trabalhos nacionais foi derivada de 154 textos, de onde foram analisados 729 segmentos de texto (ST) com aproveitamento de 614, totalizando 84,22%. Foram identificadas 25.582 ocorrências que podem ser definidas por formas, palavras ou vocábulos. Destas, 1.853 palavras aparecem uma única vez no *corpus* textual e foram identificadas 3.826 formas gramaticais distintas.

Para as análises textuais, foram selecionadas como formas ativas, ou seja, os grupos de palavras os quais o *software* Iramuteq apresentou as estatísticas textuais, as seguintes classes de palavras: substantivos; formas não reconhecidas; e verbos. As demais classes de palavras foram tratadas como suplementares. As formas não reconhecidas foram selecionadas para poder contemplar a união de grupos de palavras que possuem apenas um sentido e que foram substituídas, por exemplo, Redes Neurais e *Logistic regression*.

INSOLVÊNCIA EMPRESARIAL:

Estudo bibliométrico nacional e internacional com suporte de mineração de informações textuais

Figura 3

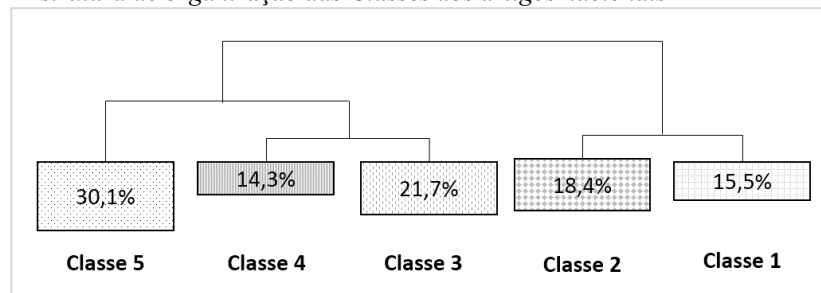
Periódicos com maior volume de publicações



No *corpus* textual foram identificadas 05 classes distintas. A Classe 1 contendo 95 ST totalizando 15,47% dos ST aproveitados; a Classe 2 possuindo 113 ST (18,4%); a Classe 3 com 133 ST (21,66%); a Classe 4 possui 88 ST (14,33%); e a Classe 5 representando 185 ST (30,13%), conforme apresentado na Figura 4.

Figura 4

Estrutura de organização das Classes dos artigos nacionais



Nas 05 classes identificadas é possível verificar o surgimento de dois conteúdos distintos. O primeiro composto pelos temas das Classes 5, 4, 3 e o segundo composto pelas Classes 2 e 1.

Na classe 5 emergem as palavras **modelo** ($\chi^2 = 59.89$ para $p < 0,0001$), **regressão logística** ($\chi^2 = 59.77$; $p < 0,0001$), **técnica** ($\chi^2 = 47.6$; $p < 0,0001$), **análise discriminante** ($\chi^2 = 34.27$; $p < 0,0001$) e com menos frequência **redes neurais** ($\chi^2 = 28.25$; $p < 0,0001$). Diante dos resultados apresentados pela classe 5, a mesma pode ser sintetizada sob a nomenclatura de “técnicas estatísticas”.

Os temas abordados pelas classes 4 e 3 se relacionam com os temas abordados na classe 5 e assim se forma o primeiro bloco de conteúdo abordado pelos trabalhos nacionais a respeito da insolvência empresarial. Na classe 4 surgem com maior força termos como **liquidez** ($\chi^2 = 134.41$; $p < 0,0001$), **capital** ($\chi^2 = 73.62$; $p < 0,0001$), **rentabilidade** ($\chi^2 = 71.91$; $p < 0,0001$), **risco** ($\chi^2 = 64.49$; $p < 0,0001$) e com menor força o termo **evidência** ($\chi^2 = 23.5$; $p < 0,0001$).

0,0001). Os resultados apresentados pela classe 4 trazem indícios sobre a discussão a respeito as “variáveis influentes” na insolvência empresarial. Portanto, a classe será assim denominada.

Enquanto na classe 3 surgem com maior expressão os termos **crédito** ($\chi^2 = 62.83$; $p < 0,0001$), **concessão** ($\chi^2 = 54.09$; $p < 0,0001$), **simulação** ($\chi^2 = 33.03$; $p < 0,0001$), **instrumento** ($\chi^2 = 31.02$; $p < 0,0001$) e com menor força o termo **carteira** ($\chi^2 = 21.91$; $p < 0,0001$). Diante dos resultados apresentados pela classe 3, ela será denominada de “risco de crédito”.

O segundo grupo de conteúdos surge da concentração das classes 1 e 2. Na classe 1 se destacam com mais intensidade os termos **importância** ($\chi^2 = 88.43$; $p < 0,0001$), **falta** ($\chi^2 = 43.17$; $p < 0,0001$), **gestor** ($\chi^2 = 43.17$; $p < 0,0001$), **planejamento** ($\chi^2 = 32.26$; $p < 0,0001$), e **local** ($\chi^2 = 27.54$; $p < 0,0001$) e aparecem com menor força os termos **emprego** ($\chi^2 = 16.47$; $p < 0,0001$) e **contabilidade** ($\chi^2 = 17.18$; $p < 0,0001$). A classe 1 será denominada “gestão e planejamento”, visto que os resultados apresentam indícios desta discussão.

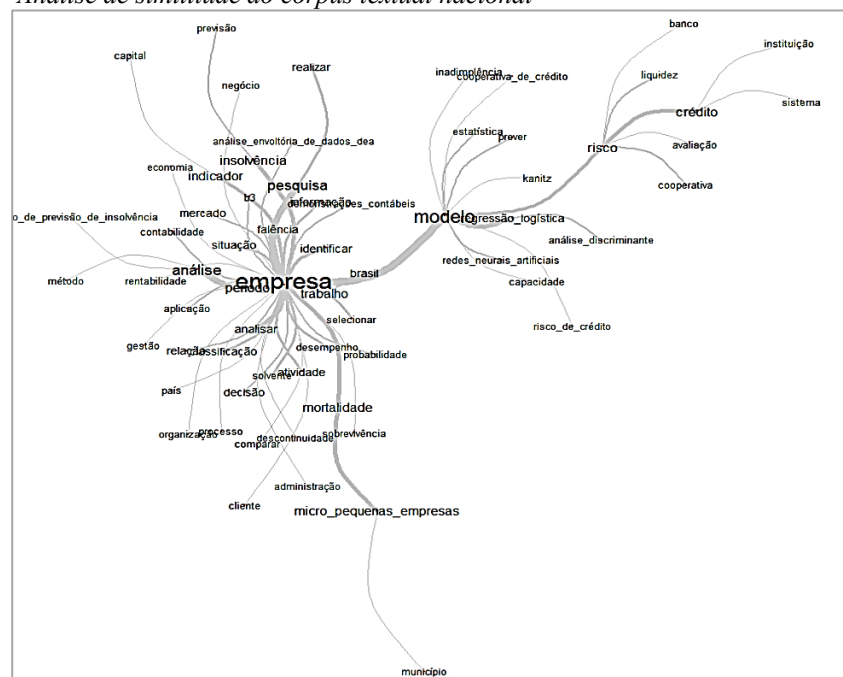
A classe 2 apresenta termos como **mortalidade** ($\chi^2 = 202.48$; $p < 0,0001$), **micro e pequenas empresas** ($\chi^2 = 79.87$; $p < 0,0001$) e **município** ($\chi^2 = 40.5$; $p < 0,0001$) e surgem com menor força os termos **região** ($\chi^2 = 18.02$; $p < 0,0001$), **fechamento** ($\chi^2 = 26.11$; $p < 0,0001$) e **entrevista** ($\chi^2 = 26.96$; $p < 0,0001$). “Micro e pequenos negócios” é a nomenclatura adotada para a classe 2 pois os resultados identificados apontam para uma discussão neste sentido.

Diante dos resultados estatísticos apresentados pelo *corpus* textual, os blocos de conteúdos foram denominados da nomeados da seguinte forma: bloco 1, composto das classes 3, 4 e 5: Variáveis influentes e técnicas estatísticas para prever insolvência empresarial e risco de crédito; e; bloco 2, composto das classes 1 e 2: Gestão e planejamento estratégico como forma de mitigar a insolvência empresarial de micro e pequenos negócios.

As formas como os diversos termos se organizam no *corpus* textual produzem as temáticas abordadas por um texto. Nos textos nacionais, observa-se que o termo central é **empresa**, de onde se ramificam os demais termos, especialmente, os termos **modelo** e **micro_pequenas_empresas**. O resultado indica que a produção e/ou teste de modelos e o estudo da situação das micro e pequenas empresas são duas vertentes fortes nos trabalhos nacionais, como apresentado na Figura 5.

Figura 5

Análise de similitude do corpus textual nacional



A ramificação onde se destaca o termo **modelo** merece ser destacada, principalmente, pelo volume de sub ramificações que a mesma apresentou. Destas sub-ramificações emergem termos como “banco”, “cooperativas_de_crédito”, “kanitz”, “análise_discriminante”, “regressão_logística” e “redes_neurais_artificiais”.

Assim, é possível inferir dois aspectos. O primeiro deles é que a modelagem matemática não é o polo central da discussão a respeito de insolvência, mas sim uma ramificação da linha de pesquisa. E que a modelagem matemática utilizada na literatura nacional é utilizada especialmente para estudar o setor bancário e o risco de crédito associado a esta atividade

A nuvem de palavras permite a rápida identificação de palavras que se destacam no *corpus* textual, apresentando com maior tamanho de fonte as palavras de maior destaque dentro do texto. A Figura 6 apresenta as palavras que mais se destacaram no *corpus* textual nacional.

Figura 6

Nuvem de palavras do corpus textual nacional

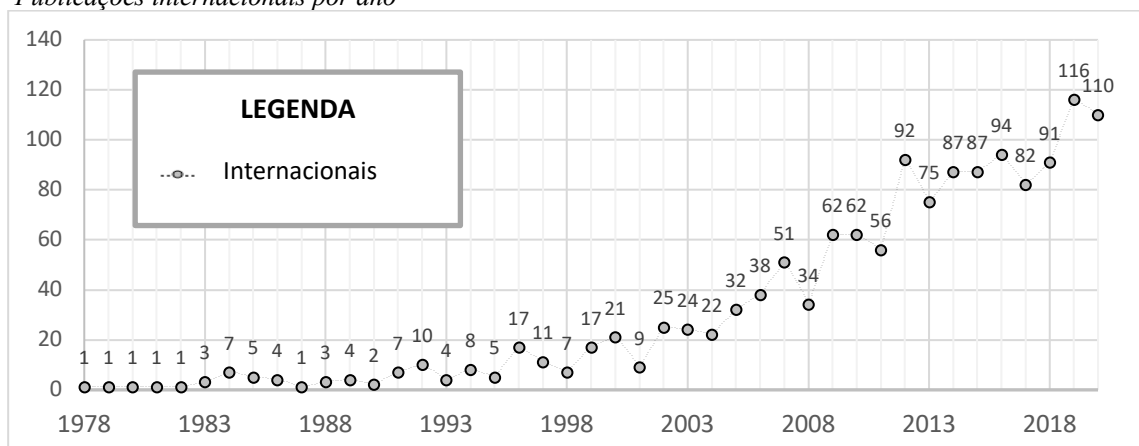


Assim como na análise de similitude no recurso gráfico na nuvem de palavras destacaram-se os termos **empresa, modelo, risco, insolvência**.

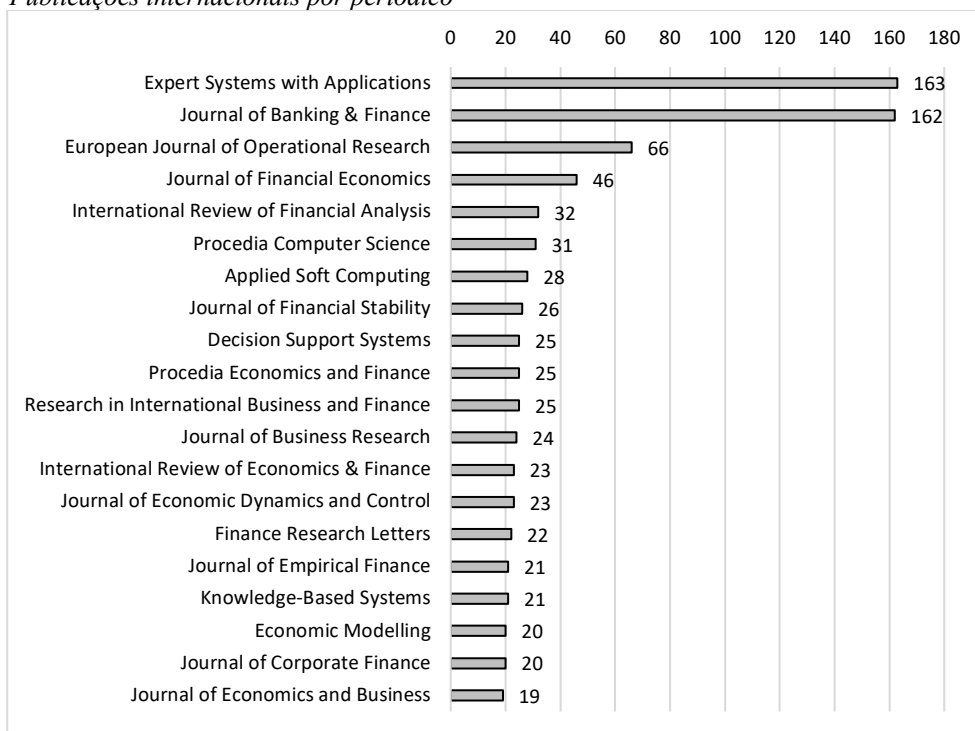
4.2 BIBLIOMETRIA INTERNACIONAL

No levantamento feito para a produção científica internacional verifica-se que as duas primeiras publicações datam de 1977, com os trabalhos “*Bankruptcy, absolute priority, and the pricing of risky debt claims*” de Jerold Warner e “*ZETATM analysis A new model to identify bankruptcy risk of Corporations*”, de Edward Altman, Robert Haldeman, e Narayanan. Dois momentos merecem ser destacados. O primeiro é a década de 1990 e o segundo é o período após 2008, onde se observou dois momentos de expansão na quantidade publicações.

Assim como foi observado na literatura nacional, a produção internacional passou a ganhar volume a partir dos anos 2000, sendo o ano de 2008, o momento em que a temática ganhou maior visibilidade internacional como apresentado na Figura 7.

Figura 7*Publicações internacionais por ano*

Merecem destaque dois periódicos pela quantidade de publicações, o *Expert Systems with Applications* e o *Journal of Banking & Finance*, os quais contaram com 163 e 162 publicações respectivamente, como observado na Figura 8.

Figura 8*Publicações internacionais por periódico*

A análise textual derivou de 1.392 publicações das quais foram analisados 5.800 segmentos de texto e foram aproveitados na análise textual 5.716 (98,55%). Foram identificadas 206.404 ocorrências contendo 9.950 formas distintas e 3.624 palavras aparecem uma única vez nos resumos.

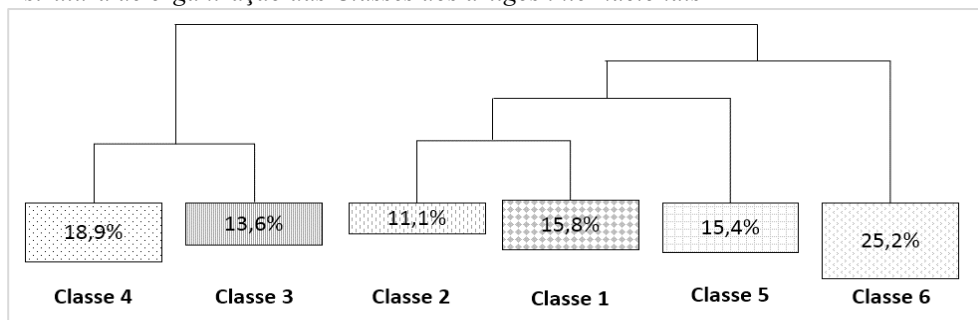
A análise dos resumos internacionais foi feita utilizando como formas ativas as seguintes classes de palavras: adjetivos, substantivos, formas não reconhecidas e verbos. Assim como na análise dos trabalhos nacionais, as demais classes de palavras foram utilizadas de forma suplementar.

Foram identificadas 06 classes de temas: a Classe 1 contou com 900 ST possuindo aproveitamento de 15,75%; a Classe 2 contendo 636 ST (11,13%); a Classe 3 com 777 ST (13,59%), a Classe 4 possui 1.083 ST (18,95%); a Classe 5 contem 882 ST (15,43%); e a Classe 6 representada por 1.438 ST (25,16%).

As 06 classes apresentadas estão divididas em dois conteúdos. O primeiro deles é composto pelas classes 1, 2, 5 e 6. Já o segundo conteúdo é composto pelas classes 3 e 4, como apresentado na Figura 9.

Figura 9

Estrutura de organização das Classes dos artigos internacionais



Da classe 1 surgem termos como *bussiness* ($\chi^2 = 181.11$; $p < 0,0001$), *manegement* ($\chi^2 = 153.89$; $p < 0,0001$), *board* ($\chi^2 = 94.39$; $p < 0,0001$) e com menos frequência os termos *social* ($\chi^2 = 56.39$; $p < 0,0001$), *audit* ($\chi^2 = 58.05$; $p < 0,0001$) e *government* ($\chi^2 = 64.8$; $p < 0,0001$). Os resultados apresentados pela classe 1 indicam para uma discussão a respeito da “gestão e gerenciamento de resultados”, portanto, a classe será assim denominada.

A classe 1 possui forte ligação com a classe 2 da qual surgem com maior força os termos *bankruptcy* ($\chi^2 = 972.02$; $p < 0,0001$), *reorganization* ($\chi^2 = 325.58$; $p < 0,0001$), *creditor* ($\chi^2 = 196.63$; $p < 0,0001$) e com menor força os termos *cost* ($\chi^2 = 103.2$; $p < 0,0001$) e *auction* ($\chi^2 = 112.1$; $p < 0,0001$). “Custos da insolvência empresarial” será a nomenclatura adotada para a classe 2 pois os resultados identificados apontam para uma discussão neste sentido.

Pode-se observar que as classes 1 e 2 possuem relação com a classe 5. Esta classe apresenta termos *rating* ($\chi^2 = 673.57 < 0,0001$), *agency* ($\chi^2 = 663.04$; $p < 0,0001$), *credit rating* ($\chi^2 = 345.05$; $p < 0,0001$) e em menor escala *evidence* ($\chi^2 = 117.85$; $p < 0,0001$) e *crisis* ($\chi^2 = 154.67$; $p < 0,0001$). Diante dos resultados apresentados pela classe 5, a mesma pode denominada de “Agências de crédito”.

O primeiro grupo de conteúdos surgiu da associação dos temas abordados nas classes 1, 2 e 5 com a classe 6. Nesta classe podem ser mencionados com maior força termos como *default* ($\chi^2 = 655.39$; $p < 0,0001$), *risk* ($\chi^2 = 519.7$; $p < 0,0001$), *price* ($\chi^2 = 306.51$; $p < 0,0001$) e com menor força termos como *credit* ($\chi^2 = 116.97$; $p < 0,0001$) e *structural* ($\chi^2 = 132.18$; $p < 0,0001$). A classe 6 será denominada de “risco de *default*”.

O segundo grupo é composto pela ligação das classes 3 e 4. A classe 4 possui como termos de maior força *classifier* ($\chi^2 = 526.21$; $p < 0,0001$), *algorithm* ($\chi^2 = 525.12$; $p < 0,0001$), *method* ($\chi^2 = 435.92$; $p < 0,0001$) e com menor força os termos *Support Vector Machine* ($\chi^2 = 223.64$; $p < 0,0001$) e *Neural Network* ($\chi^2 = 340.96$; $p < 0,0001$). A classe 4 será denominada de “Algoritmo de classificação”, visto que os resultados apresentam indícios sobre uma discussão nesse sentido.

A classe 3 apresenta como termos mais relevantes *statistical* ($\chi^2 = 349.24$; $p < 0,0001$), *Credit Score* ($\chi^2 = 324.1$; $p < 0,0001$) e *machine learn* ($\chi^2 = 315.36$; $p < 0,0001$) e com menor

percepção do microempresário sobre os motivos que levaram a falência de seu negócio, como observado no trecho a seguir.

Foram entrevistados os sócios das empresas, gerentes ou pessoas que ocuparam cargos de chefia nas empresas estudadas, que encerraram suas atividades com até três anos de operação, para que se tivesse uma visão mais ampliada e detalhada sobre os motivos do encerramento. (Ferreira et al., 2012, p. 816)

Assim, o primeiro grupo de conteúdos abordou uma perspectiva voltada para empreendimentos locais fazendo uso de técnicas não estatísticas demonstrando uma forte preocupação com a influência dos cenários da economia e da gestão nestes negócios.

O segundo grupo de conteúdos (Variáveis influentes e técnicas estatísticas para prever insolvência empresarial e risco de crédito) foi composto pelas classes 3, 4 e 5. Foi possível inferir que na classe 5 são abordadas temáticas relacionadas as técnicas utilizadas para previsão de insolvência empresarial, especialmente técnicas estatísticas, como a regressão logística e análise discriminante. Observa-se que o tema possui forte influência do trabalho desenvolvido por Stephen Kanitz em 1974.

As classes 3 e 4 apresentaram forte ligação entre elas. Na classe 3 observou-se conteúdos voltados para a análise da carteira de crédito do setor financeiro, estudando com bastante intensidade os riscos envolvidos na concessão destes créditos. Já na classe 4 foi possível perceber uma temática voltada essencialmente para a escolha de variáveis que indiquem a incapacidade financeira, explicando sua forte relação como a Classe 3.

O trecho a seguir expõe as principais características da Classe 4 de conteúdos, ao demonstrar a preocupação em identificar as variáveis que influenciam o risco de crédito e fazer a seleção destas variáveis por meio de técnicas estatísticas.

Os resultados encontrados também mostram que os **créditos concedidos** sem garantias, como os adiantamentos feitos aos cooperados, tendem a deteriorar a saúde financeira das cooperativas, pois, como apresentado no **modelo de regressão**, os indicadores que utilizam a **conta provisões de crédito de liquidação duvidosa**, constituída por atrasos e inadimplência por parte dos cooperados, apresentaram impactos negativos sobre a liquidez dessas cooperativas. (Gonçalves & Braga, 2006, p. 217, grifo do autor)

Analisando a discussão internacional observam-se, também, dois grupos de conteúdos, porém compostos por seis classes distintas. O primeiro grupo é composto pelas classes 1, 2, 5 e 6, já o segundo, composto pelas Classes 3 e 4. O primeiro grupo foi denominado de “Gestão, gerenciamento de resultados, custos da insolvência e risco de *default* mensurado pelas agências de *rating*” enquanto o segundo foi identificado como “Algoritmo de classificação de insolvência e *machine learn*”.

Os assuntos abordados na Classe 1 se preocupam com aspectos como gestão e gerenciamento de resultados apontando, também, a relevância da opinião da auditoria para os usuários da informação contábil, como pode ser percebido no trecho a seguir classificado como um segmento de texto típico para esta classe.

The effect of earnings management to bankruptcy risk is essential for external stakeholders, such as investors and creditors, to assess bankruptcy risk, financial capability, and credit worthiness of a firm, while business strategy effect on bankruptcy risk benefits internal stakeholders, such as managers, in formulating strategies to deal with going concern issues. (Agustia, Muhammad, & Permatasari, 2020 p.1, grifo do autor)

A Classe 2 tende a abordar temáticas relacionadas aos custos que envolvem a insolvência empresarial, mostrando preocupação com os *Stakeholder*. Especialmente, com os credores e trazendo forte abordagem da legislação no processo falimentar sendo bastante citado o capítulo 11 da Lei de Falências Norte Americana.

Na Classe 5 são discutidos assuntos relacionados a risco de crédito avaliados pelas agências de classificação de crédito, como a *Moody's* e *Standard & Poor's*. Trabalha, também, com os impactos gerados pela redução ou elevação na classificação proposta por estas agências. Como pode ser observado no trecho a seguir.

The economic relevance of defaults is tremendous: Moody's reported defaulted bonds with a face value of around USD 390 billion between 1997 and 2004 (Moody's, 2004). Therefore, investors and regulators are growing concerned about the quality of external credit ratings given the rating agency's business model, i.e. the fact that they are paid by the issuers, and given the oligopolistic structure of the market for external credit ratings. (Güttler & Wahrenburg, 2007, p. 752)

Na Classe 6 emergem assuntos ligados ao risco de *default* em contratos créditos. Portanto, esta é uma classe que irá adentrar assuntos ligados a carteira de crédito bancária, *swap* e *spread* bancário. Vale salienta que a Classe 6 se mostrou como a mais significativa do grupo detendo 25,2% dos Segmentos de Texto analisados. Como exemplo de texto dessa classe pode ser citado o trecho do trabalho de Claußen et al. (2017, p.1) “*In summary, specific systematic risk variables such as the Credit Market Climate, the Cross-market Correlation and the Market Volatility may play a major role in pricing credit default swaps*”.

Observa-se que este grupo de conteúdos demonstra forte preocupação com os riscos associados ao processo de insolvência empresarial, levantando assuntos como gestão empresarial, opinião da auditoria e também, dos *stakeholders*. Esse é um grupo de conteúdos que se mostrou bastante heterogêneo à medida que levanta, também, questões ligadas ao risco soberano e ao mercado acionário.

O segundo grupo de conteúdos é composto pelas Classes 3 e 4. Na Classe 3 é possível perceber o interesse na comparação entre técnicas estatísticas e de inteligência de máquina aplicando principalmente no segmento de classificação de crédito. Desta classe emergem, também, técnicas voltadas para o trabalho conjunto de análises estatísticas e de aprendizado de máquina. Já na classe 4 podemos observar a importância que as técnicas de aprendizado de máquina têm recebido nos estudos internacionais, destacando-se as técnicas de Máquina de Vetores de Suporte e Redes Neurais. Cabe ressaltar que estes resultados corroboram com os obtidos por Führ et al. (2017), que verificou a expansão das técnicas de aprendizado de máquina, apesar, da predominância das técnicas estatísticas.

Na comparação dos conteúdos abordados nos trabalhos nacionais e internacionais há forte diferença entre o porte das empresas estudadas. Os artigos nacionais se voltam para atender as demandas das pequenas e médias empresas, adotando inclusive entrevistas como técnica de coleta de informações. O interesse por estes assuntos pode ser justificado pela tentativa de replicação dos modelos de insolvência tradicionais à realidade socioeconômica brasileira. Na prática, observa-se que, em muitas situações, a contabilidade é utilizada em micro e pequenas empresas para cumprir apenas exigências fiscais, e que, portanto, não representa fidedignamente a situação econômico-financeira da entidade, justificando a utilização de entrevistas como forma de obter informações ou de refinar a informação contábil. A seguir é apresentado a TabelaTabela 1 que traz um comparativo dos resultados identificados na bibliometria nacional e internacional.

Tabela 1

Comparativo entre a bibliometria nacional e internacional

	Bibliometria nacional	Bibliometria internacional
Porte do empreendimento	Empreendimentos locais	Empreendimentos de grande porte
Principais discussões	Gestão e planejamento estratégico Percepção de microempresário sobre a insolvência - -	Efeitos do gerenciamento de resultados no risco de insolvência; Os custos sociais que envolvem o processo de insolvência empresarial; Impacto da avaliação das agências de classificação de crédito, como a Moody's e Standard & Poor's; Risco de <i>default</i> em contratos de crédito
Técnicas de coleta de informações	Entrevistas para refino da informação contábil	Dados secundários
Técnicas estatísticas predominantes	Análise discriminante e regressão logística	Técnicas de <i>Machine Learn</i> e <i>Machine Learn</i> associado a técnicas estatísticas

Com relação à análise das técnicas utilizadas, houve também forte diferença. Nos trabalhos nacionais, é possível identificar a influência das técnicas estatísticas, como análise discriminante e regressão logística, com leve tendência a aplicação de técnicas de aprendizado de máquinas. Enquanto que nos trabalhos internacionais observa-se maior discussão desta última, inclusive, com aplicação de técnicas que agregam aprendizado de máquina e estatística emergindo como uma classe de temas exclusiva. Apesar das críticas que as técnicas computacionais recebem, especialmente, relacionadas à baixa replicabilidade, é importante notar que se trata uma metodologia mais recente e que se mostra bastante promissora.

6 CONCLUSÕES

Este trabalho teve como objetivo: compreender as linhas de pesquisa desenvolvidas no âmbito dos estudos de insolvência empresarial em nível nacional e internacional. Para tanto, utilizou-se da técnica ProKnow-C para sistematização da seleção de dados. Já para o processo de mineração dos dados foi utilizado o *software* Iramuteq.

Verificou-se que após o ano de 2008 os assuntos ligados à insolvência empresarial receberam maior interesse entre os pesquisadores, demonstrando aumento significativo na produção acadêmica tanto nacional quanto internacional.

Os resultados obtidos mostraram diferenças nas linhas de pesquisas adotadas por pesquisadores nacionais e internacionais, mostrando divergências entre conteúdos e técnicas utilizadas. Sendo os trabalhos nacionais voltados para atender a realidade socioeconômica local ao abordar temáticas relacionadas a microempresas. Enquanto que os trabalhos internacionais apresentaram abordagens mais heterogêneas. Com relação às técnicas observa-se: a predominância de técnicas estatísticas nos trabalhos nacionais; enquanto nos internacionais, observou-se a expansão de técnicas de aprendizado de máquina e, também, de técnicas mistas.

Para as pesquisas de insolvência empresarial, este trabalho se mostra relevante ao sintetizar as principais linhas de pesquisas desenvolvidas em uma área do conhecimento de interesse multidisciplinar e que ainda não possui uma teoria conclusiva.

Como limitações do presente trabalho, podem ser mencionadas a utilização de uma base de busca para a literatura nacional e outra para a internacional de modo que os motores de busca não apresentavam as mesmas variáveis para seleção e filtragem dos artigos buscados. Outro fator limitante é a agregação utilizada na análise da literatura internacional, o que não permite a identificação do desenvolvimento da linha de pesquisa entre os diferentes países. Por fim, é válido destacar a que a riqueza de variações permitidas pela língua portuguesa para expressar um mesmo

sentido pode de alguma forma influenciar a qualidade informacional dos resumos utilizados no *corpus* nacional.

Salienta-se que o objetivo deste artigo não é esgotar a discussão e, portanto, deixa-se como sugestão para trabalhos futuros a expansão das palavras-chave utilizadas, a utilização de outras técnicas de mineração de informação, a aplicação da análise em todo o conteúdo textual dos artigos selecionados. Assim como, produzir análises para os resultados internacionais estratificados por país.

REFERÊNCIAS

- Agustia, D., Muhammad, N. P. A., & Permatasari, Y. (2020). Earnings management, business strategy, and bankruptcy risk: evidence from Indonesia. *Heliyon*, 6(2), 9. <https://doi.org/10.1016/j.heliyon.2020.e03317>
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589–609. <https://doi.org/10.1111/j.1540-6261.1968.tb00843.x>
- Altman, E. I., Haldeman, R. G., & Narayanan, P. (1977). ZETATM analysis A new model to identify bankruptcy risk of corporations. *Journal of banking & finance*, 1(1), 29-54.
- Araújo, C. A. (2006). Bibliometria: evolução histórica e questões atuais. *Em Questão*, 12(1), 11–32. <https://doi.org/10.19132/1808-5245121>
- Beaver, W. H. (1966). Financial Ratios As Predictors of Failure. *Journal of Accounting Research*, 4, 71. <https://doi.org/10.2307/2490171>
- Bressan, V. G. F., Braga, M. J., & Lima, J. E. de. (2004). Análise de insolvência das cooperativas de crédito rural do estado de Minas Gerais. *Estudos Econômicos* (São Paulo), 34(3), 553–585. <https://doi.org/10.1590/s0101-41612004000300006>
- Cabrera, G. A. S. (1998). *Um modelo neural para previsão de insolvência no sistema financeiro* (Dissertação de Mestrado) Pontifícia Universidade Católica do Rio de Janeiro. Rio de Janeiro. Brasil. <https://doi.org/10.17771/PUCRio.acad.7493>
- Castelli, T. M. (2018). *Análise da metodologia Knowledge Development Process–Constructivist (Proknow-C) e suas contribuições à avaliação de desempenho organizacional: um estudo à luz do apoio à decisão* (Dissertação de mestrado) Universidade Federal do Rio Grande. Rio Grande do Sul. Brasil. <https://repositorio.furg.br/handle/1/7721>
- Claußen, A., Löhr, S., Rösch, D., & Scheule, H. (2017). Valuation of systematic risk in the cross-section of credit default swap spreads. *The Quarterly Review of Economics and Finance*, 64, 183–195. <https://doi.org/10.1016/j.qref.2016.06.007>
- Contador, C. R. (1985). Insolvência de empresas e política macroeconômica. *Revista de Administração*, 20(2), 15-27.
- Daza, J. R. P. (1990). Um modelo de análise de insucesso de empresas. *Revista Brasileira de Economia*, 44(4), 603-624.
- Elizabetsky, R. (1976). *Um modelo matemático para decisões de crédito no banco comercial*. 1976. (Dissertação de mestrado, Escola Politécnica da Universidade de São Paulo). São Paulo-SP).
- Ensslin, L., Ensslin, S. R., & Pinto, H. de M. (2013). Processo de investigação e análise bibliométrica: avaliação da qualidade dos serviços bancários. *Revista de Administração Contemporânea*, 17(3), 325–349. <https://doi.org/10.1590/S1415-65552013000300005>
- Ensslin, L., Ensslin, S. R., Lacerda, R. T. de O., & Tasca, J. (2010). *Processo de análise bibliométrica. processo técnico com patente de registro pendente junto ao INPI*. Brasil.
- Fenacon. (2022, 4 fevereiro). *Pequenos negócios geraram 78% dos empregos formais criados em 2021, aponta Sebrae*. FENACON. <https://fenacon.org.br/noticias/pequenos-negocios-geraram-78-dos-empregos-formais-criados-em-2021-aponta-sebrae/>
- Ferreira, L. F. F., Oliva, F. L., Santos, S. A. dos, Grisi, C. C. de H. e, & Lima, A. C. (2012). Análise quantitativa sobre a mortalidade precoce de micro e pequenas empresas da cidade de São Paulo. *Gestão & Produção*, 19(4), 811–823. <https://doi.org/10.1590/S0104-530X2012000400011>
- Für, F., Lima, J. D., & Schenatto, F. J. A. (2017). Uma revisão sistemática da literatura sobre credit scoring. In *VII Congresso Brasileiro de Engenharia de Produção* (pp. 1-12). Rio de Janeiro: ABREPRO.

- Gonçalves, R. M. L., & Braga, M. J. (2006). Condicionantes do risco de liquidez em cooperativas de economia e crédito mútuo de Minas Gerais. *Revista de Economia e Administração*, 5(2), 201–220. <https://doi.org/10.11132/rea.2002.122>
- Güttler, A., & Wahrenburg, M. (2007). The adjustment of credit ratings in advance of defaults. *Journal of Banking & Finance*, 31(3), 751–767. <https://doi.org/10.1016/j.jbankfin.2006.05.014>
- Horta, R. A. M. *Utilização de indicadores contábeis na previsão de insolvência: Análise empírica de uma amostra de empresas comerciais e industriais brasileiras*. (Dissertação Mestrado) Faculdade de Ciências Contábeis da Universidade do Estado do Rio de Janeiro. Rio de Janeiro. Brasil.
- Janot, M. M. (2001). Modelos de Previsão de Insolvência Bancária no Brasil. *Working Paper Series Do Banco Central*, 13, 1–41.
- Kanitz, S. C. (1974). Como prever falências. *Revista de Negócios Em Exame*, 95–102.
- Kücher, A., & Feldbauer-Durstmüller, B. (2019). Organizational failure and decline – A bibliometric study of the scientific frontend. *Journal of Business Research*, 98, 503–516. <https://doi.org/10.1016/j.jbusres.2018.05.017>
- Mansi, S., Maxwell, W. F., & Zhang, A. (Jianzhong). (2010). Bankruptcy Prediction Models and the Cost of Debt. *SSRN Electronic Journal*, 21. <https://doi.org/10.2139/ssrn.1622407>
- Martins, M., & Samané, C. (1987). Análise Discriminante e os Bancos Comerciais Privados Brasileiros: uma Aplicação. *Revista Brasileira de Mercado de Capitais*, 13(39), 123–142.
- Matias, A. B. (1978). *Contribuição às técnicas de análise financeira: um modelo de concessão de crédito. 1978. 106 f* (Dissertação de Mestrado em Administração, Faculdade de Economia e Administração da USP. São Paulo).
- Matias, A. B., & Siqueira, J. de O. (1996). Risco bancário: modelo de previsão de insolvência de bancos no Brasil. *Revista de Administração - RAUSP*, 31(abr./jul 1996), 19–28.
- Minussi, J. A., Damacena, C., & Ness Jr, W. L. (2002). Um modelo de previsão de solvência utilizando regressão logística. *Revista de Administração Contemporânea*, 6(3), 109–128. <https://doi.org/10.1590/S1415-6552002000300007>
- Neves, J. A. D., & Pessoa, R. W. A. (2006). Causas da mortalidade de micros e pequenas empresas: o caso das lojas de um shopping center. *Organizações Em Contexto*, 2(4), 143–164. <https://doi.org/10.15603/1982-8756/roc.v2n4p%20165%20-%20195>
- Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18(1), 109–131. <https://doi.org/10.2307/2490395>
- Pinheiro, L. E. T., & Pinheiro, J. L. (2008). Modelos de Evaluación del Riesgo de Insolvencia de Empresas Españolas Cotizadas. *Revista Contabilidade Vista e Revista*, 19(3), 95–121.
- Prado, J. W., Castro Alcântara, V., Melo Carvalho, F., Vieira, K. C., Machado, L. K. C., & Tonelli, D. F. (2016). Multivariate analysis of credit risk and bankruptcy research data: a bibliometric study involving different knowledge fields (1968–2014). *Scientometrics*, 106(3), 1007–1029. <https://doi.org/10.1007/s11192-015-1829-6>
- Sanvicente, A. Z., & Minardi, A. M. A. F. (1998). Identificação de indicadores contábeis significativos para a previsão de concordata de empresas. *Working Paper. Instituto Brasileiro de Mercado de Capitais*.
- Securato, J. R., & Famá, R. (1997). Um procedimento para a decisão de crédito pelos bancos. *Revista de Administração Contemporânea*, 1(1), 101–119. <https://doi.org/10.1590/S1415-65551997000100006>
- Serasa Experian (2022). *Indicadores econômicos*. <https://www.serasaexperian.com.br/conteudos/indicadores-economicos/>
- Shi, Y., & Li, X. (2019). A bibliometric study on intelligent techniques of bankruptcy prediction for corporate firms. *Heliyon*, 5(12), e02997. <https://doi.org/10.1016/j.heliyon.2019.e02997>
- Silva, J. P. da. (1983). *Administração de Crédito e Previsão de Insolvência* (1st ed.). Atlas.
- Wang, G., Ma, J., & Yang, S. (2014). An improved boosting based on feature selection for corporate bankruptcy prediction. *Expert Systems with Applications*, 41(5), 2353–2361. <https://doi.org/10.1016/j.eswa.2013.09.033>
- Warner, J. B. (1977). Bankruptcy, absolute priority, and the pricing of risky debt claims. *Journal of financial economics*, 4(3), 239–276.
- Wruck, K. H. (1990). Financial distress, reorganization, and organizational efficiency. *Journal of Financial Economics*, 27(2), 419–444. [https://doi.org/10.1016/0304-405X\(90\)90063-6](https://doi.org/10.1016/0304-405X(90)90063-6)