



**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO PIAUÍ
CAMPUS FLORIANO
CURSO TECNOLOGIA EM ANÁLISE E DESENVOLVIMENTO DE
SISTEMAS**

ALEXÂNDER ARAÚJO REIS

**MODELO PREDITOR DE LICENCIAMENTO DE PATENTES ACADÊMICAS
BRASILEIRAS NOS CAMPOS TECNOLÓGICOS BIOTECNOLOGIA E
FARMACÊUTICA**

FLORIANO

2022

ALEXÂNDER ARAÚJO REIS

MODELO PREDITOR DE LICENCIAMENTO DE PATENTES ACADÊMICAS
BRASILEIRAS NOS CAMPOS TECNOLÓGICOS BIOTECNOLOGIA E
FARMACÊUTICA

Trabalho de Conclusão de Curso (Artigo científico) apresentado como exigência parcial para obtenção do diploma do Curso de Tecnologia em Análise e Desenvolvimento de Sistemas do Instituto Federal de Educação, Ciência e Tecnologia do Piauí, Campus Floriano.

Orientador: Prof. O Dr. Rafael Angelo Santos Leite.

FLORIANO

2022

MODELO PREDITOR DE LICENCIAMENTO DE PATENTES ACADÊMICAS BRASILEIRAS NOS CAMPOS TECNOLÓGICOS BIOTECNOLOGIA E FARMACÊUTICA

Alexander Araujo Reis¹

Rafael Angelo Santos Leite²

RESUMO

Compreender quais aspectos norteiam a determinação potencial de licenciamento de uma patente é uma questão central para as universidades empreendedoras que buscam criar valor para a sociedade por meio da transferência de tecnologia. Este artigo propõe uma abordagem de *Random Forest* (RF) para prever o potencial de licenciamento de patentes acadêmicas brasileiras nos campos tecnológicos biotecnologia e farmacêutica a partir da mineração de dados contidos em documentos de patentes licenciadas e não licenciadas. Primeiro, construiu-se duas bases de dados: a primeira com as patentes licenciadas das 8 universidades que mais licenciaram patentes nos campos tecnológicos farmacêutica e biotecnologia, e a segunda com as demais características de patentes licenciadas e não licenciadas. Em seguida, três algoritmos de aprendizado de máquina foram avaliados nos dados, sendo que o algoritmo de *Random Forest* alcançou o melhor desempenho relativo. Por isso, um modelo em RF foi desenvolvido para avaliar o potencial de licenciamento dessas patentes. Ao final, o desempenho do modelo foi avaliado com métricas de accuracy, precision e recall. O resultado é um modelo em RF para previsão do potencial de licenciamento de patentes a partir da determinação das variáveis mais influentes.

Palavras-chave: random forest; mineração de dados; licenciamento de patentes acadêmicas; biotecnologia; farmacêutica.

ABSTRACT

Understanding which aspects guide the potential licensing determination of a patent is a central issue for entrepreneurial universities seeking to create value for society through technology transfer. This paper proposes a Random Forest (RF) approach to predict the licensing potential of Brazilian academic patents in the biotechnology and pharmaceutical technology fields by mining data contained in licensed and unlicensed patent documents. First, two databases were built: the first with the licensed patents from 8 universities that licensed the most patents in the pharmaceutical and biotechnology technological fields, and the second with the remaining characteristics of licensed and unlicensed patents. Next, three machine learning algorithms were assessed on the data, with the Random Forest algorithm achieving the best relative performance. Therefore, an RF model was developed to evaluate the licensing potential of these patents. At the end, the model performance was evaluated with metrics of accuracy, precision and recall. The result is an RF model for predicting the licensing potential of patents by determining the most influential variables.

Keywords: random forest; data mining; academic patent licensing; biotechnology; pharmaceutical.

Data de aprovação: 25/11/2022 (data de apresentação do TCC)

¹ Graduando do curso de Tecnologia em Análise e Desenvolvimento de Sistemas. Instituto Federal do Piauí - Floriano. E-mail: caflo.2020114tads24@aluno.ifpi.edu.br.

² Professor do Eixo Gestão e Negócios do Instituto Federal do Piauí - Floriano. E-mail: rafaelangelo@ifpi.edu.br

SUMÁRIO

1 INTRODUÇÃO	5
2 REVISÃO TEÓRICA	7
2.1 INOVAÇÃO ABERTA NAS UNIVERSIDADES E A INDÚSTRIA BIOFARMACÊUTICA	7
2.2 MINERAÇÃO DE DADOS E ANÁLISE DE PATENTES	7
2.3 APRENDIZADO DE MÁQUINA	8
2.4 TRABALHOS RELACIONADOS	9
3 METODOLOGIA	11
3.1 PERCURSO METODOLÓGICO	12
3.2 RANDOM FOREST	13
4 RESULTADOS	14
4.1 ANÁLISE DESCRITIVA	14
4.2 AVALIAÇÃO DO MODELO PARA PREVER LICENCIAMENTO.	15
4.2.1 Escolha do melhor algoritmo de classificação	15
4.2.2 Importância relativa das variáveis no modelo final	16
4.2.3 Avaliação da performance do modelo	17
5 DISCUSSÃO DE RESULTADOS	18
6 CONCLUSÃO	22

1 INTRODUÇÃO

A compreensão de quais aspectos norteiam a determinação do potencial de licenciamento de uma patente é uma questão central para as universidades e Institutos de ciência e tecnologia com capacidade de inovação (ZHYLINSKA; SITNICKI; VIKULOVA, 2019), e os profissionais responsáveis pela gestão da inovação nos Núcleos de Transferência de Tecnologia das Instituições de Ciência Tecnologia e Inovação (ICT) têm a tarefa de avaliar e identificar como alavancar ou aprimorar as tecnologias desenvolvidas, mantendo a capacidade de inovação organizacional (BURGELMAN; CHRISTENSEN; WHEELWRIGHT, 2013).

Nesse sentido, para que estas instituições possam efetivamente criar valor para a sociedade por meio da transferência de tecnologia (SOARES; TORKOMIAN; NAGANO, 2020) com parcerias mais efetivas (DE MOORTELE; CHRISPEELS, 2018), a produção tecnológica precisa ser avaliada quanto a sua qualidade técnica, comercial e legal.

O licenciamento de patentes é o método mais popular de transação de patentes devido à sua flexibilidade (WANG; LIN; LO, 2012) com valor do mercado global estimado em US\$ 200 bilhões em 2011 (ALVAREZ; LÓPEZ, 2015). Nesse sentido, é importante identificar características de patentes relacionadas a sua qualidade técnica, comercial e legal que mais afetam o potencial de licenciamento das patentes.

Considerando que os campos tecnológicos biofarmacêuticos (biotecnologia e farmacêutico) representam 35,64% dos licenciamentos de patentes acadêmicas brasileiras (LEITE *et al.*, 2022) e que a maior parte das patentes desenvolvidas nas ICT públicas no Brasil são desses dois campos, surge a seguinte questão de pesquisa: *Quais fatores mais influenciam no potencial de licenciamento de patentes acadêmicas brasileiras nos campos tecnológicos biotecnologia e farmacêutica?*

Nesse contexto, este estudo tem por objetivo identificar quais características são influentes para o licenciamento de patentes acadêmicas brasileiras nos campos tecnológicos biotecnologia e farmacêutica a partir da mineração de dados contidos em documentos de patentes licenciadas e não licenciadas.

Existem muitos trabalhos ao redor do mundo com propostas de aplicação de algoritmos de aprendizado de máquina em patentes em diversas aplicações (ABBAS; ZHANG; KHAN, 2014; CHI; WANG, 2022; DE CLERCQ *et al.*, 2019; SHEREMETYEVA, 2003) como classificação automática dos documentos, análise de reivindicações com

processamento de linguagem natural, previsão de tecnologias emergentes (CHI; WANG, 2022; DE CLERCQ *et al.*, 2019; SHEREMETYEVA, 2003).

Porém, no contexto do Brasil não há estudos com aplicação de aprendizado de máquina em dados de licenciamentos, pois essas informações são difíceis de reunir (LEITE *et al.*, 2022). Desta forma, a aplicação de um algoritmo para previsão de licenciamentos se torna um desafio ainda maior devido a dificuldade de encontrar e reunir grandes quantidades de dados para o treinamento do modelo. Assim, sob o ponto de vista teórico esse estudo pode preencher lacunas científicas respondendo questões escassas no contexto da América Latina e Brasil quando se trata de fatores que influenciam o potencial de licenciamentos de patentes biofarmacêuticas.

Considerando a grande representatividade do campo biofarmacêutico (biotecnologia e farmacêutico) nos licenciamentos realizados pela academia e que a maior parte das patentes desenvolvidas nas ICT públicas brasileiras são desses dois campos, este estudo contribui para melhores decisões dos gestores de propriedade intelectual dessas ICT quando esses entendem quais ações estratégicas podem aumentar o desempenho em licenciamentos das suas patentes.

Além disso, ações estratégicas que aumentem licenciamentos de patentes significam aumentar futuras receitas públicas e assim reduzir a dependência dos recursos federais tão escassos nos últimos anos, o que demonstra a relevância social deste estudo.

Além da introdução, esta pesquisa contém outras 5 seções. Em seguida a revisão teórica apresenta uma síntese da literatura acerca de inovação aberta, mineração de dados e análise de patentes, aprendizado de máquina e trabalhos relacionados. Depois, um passo descreve-se um passo a passo da abordagem do problema na metodologia, seguida dos resultados alcançados. Por fim, é feita a discussão e a conclusão dos resultados.

2 REVISÃO TEÓRICA

2.1 INOVAÇÃO ABERTA NAS UNIVERSIDADES E A INDÚSTRIA BIOFARMACÊUTICA

Inovação aberta é o processo de acelerar a inovação usando conhecimento interno e externo da organização. Nesse processo, as organizações podem trazer tecnologias externas para dentro da empresa e incorporá-las no processo de inovação (fora para dentro) ou a empresa compartilha e permite a incorporação por terceiros de suas ideias e tecnologias subutilizadas (dentro para fora) (CHESBROUGH, 2003). Um exemplo do processo de dentro

para fora é o licenciamento de tecnologias desenvolvidas em universidades que se tornam assim provedoras de soluções para indústria (ZIEGLER *et al.*, 2013).

Considerando os licenciamentos já realizados pelas universidades públicas brasileiras, os campos tecnológicos biofarmacêuticos (biotecnologia e farmacêutico) representam 35,64% dos licenciamentos de patentes acadêmicas brasileiras. Dentro do campo tecnológico Farmacêutico, as Universidade de São Paulo, Universidade de Campinas (Unicamp) e Universidade Federal de Minas Gerais (UFMG) representam 67,7% dos licenciamentos. Dentro do campo tecnológico Biotecnologia, UFMG e Unicamp juntas têm 41% dos licenciamentos nesse mercado. Essa distribuição nos licenciamentos demonstrando *expertise* e liderança de algumas instituições nesse mercado (LEITE *et al.*, 2022).

Considerando agora a perspectiva da inovação de fora para dentro, a indústria biofarmacêutica é um ecossistema de inovação aberta, e a estratégia de licenciamento tornou-se um modelo de negócios central para empresas de biotecnologia (HOFMAN; NIKLASSON, 2016). A indústria biofarmacêutica tem uma propensão relativamente alta para licenciamento (ANAND; KHANNA, 2003) e as patentes são meios de apropriação de rendas de uma tecnologia. Na indústria de biotecnologia, o licenciamento é uma das estratégias mais importantes para superar a crise econômica.

2.2 MINERAÇÃO DE DADOS E ANÁLISE DE PATENTES

Com o avanço da tecnologia e da internet a necessidade de armazenamento de dados *online* cresceu nos últimos anos. O grande crescimento no volume de dados gerou uma necessidade urgente de novas técnicas e ferramentas para transformar dados em informações e conhecimento (CHEN; HAN; YU, 1996).

A mineração de texto, compreendida como um campo interdisciplinar entre a mineração de dados, linguística, estatística e ciência da computação, tornou-se uma disciplina amplamente aplicada, utilizando métodos estatísticos e de aprendizado de máquina (FEINERER; HORNIK; MEYER, 2008).

Os documentos de patentes contêm importantes resultados de pesquisas. No entanto, eles são longos e ricos em terminologia técnica, o que exige muito esforço humano para as análises (TSENG; LIN; LIN, 2007). Nesse contexto, técnicas de análise de patentes que envolve uma série de etapas incluindo extração de patentes de base de dados, extração de informação de patentes (mineração de texto e *web scrapping*) e análise da informação extraída para inferências lógicas (ABBAS; ZHANG; KHAN, 2014) são empregadas para extração de informações em velocidade e quantidade além da humana.

A geração de grandes base de dados de patentes também permitiu a aplicação de algoritmos de aprendizado de máquina em dados de patentes, seja para classificação automática dos documentos (DE CLERCQ *et al.*, 2019), análise de reivindicações com processamento de linguagem natural (SHEREMETYEVA, 2003) ou previsão de tecnologias emergentes (CHI; WANG, 2022).

2.3 APRENDIZADO DE MÁQUINA

Aprendizado de máquina refere-se a habilidade de um sistema de adquirir e integrar conhecimento através de observações em larga escala, aperfeiçoando-se e aprendendo novos conhecimentos ao invés de ser programado com aquele conhecimento (WOOLF, 2010).

Essa tecnologia vem sendo aplicada em diversos contextos como previsão de câncer em pacientes (SHAIKH; RAO, 2022), previsão de preços de ações (VIJH *et al.*, 2020) e até previsão e análise da propagação de epidemias de COVID-19 (LI *et al.*, 2020).

Em matéria de aprendizado de máquina supervisionada, os problemas são abordados com a aplicação de algoritmos como Regressão Linear, Regressão Logística, Árvore de Decisão, *Random Forest*, Redes Neurais Artificiais entre outros .

A Árvore de Decisão é uma abordagem de Aprendizado de Máquina Supervisionado para resolver problemas de classificação e regressão continuamente dividindo dados com base em um determinado parâmetro. As decisões são as folhas e os dados são divididos em nós. Dentro da árvore de classificação a variável de decisão é categórica (Sim/Não) (RAY, 2019).

Uma das desvantagens da Árvore de Decisão é a instabilidade, pequena variação nos dados de treino podem resultar em árvores completamente distintas. Isso pode ser evitado ao treinarmos várias árvores distintas e agregar suas predições. Nesse sentido, surge a *Random Forest* que é uma combinação de árvores de decisão as quais votam pela classe mais popular, abordagem que permite um aumento significativo na acurácia das classificações em comparação com o algoritmo de Árvore de Decisão (BREIMAN, 2001).

2.4 TRABALHOS RELACIONADOS

Wang, Lin e Lo (2012) discutem fatores influentes para a comercialização de patentes acadêmicas: política e meio ambiente, mercado, organizações acadêmicas, inventores, propriedade intelectual e tecnologia. Sendo que o mercado é o fator de maior influência para a comercialização de patentes. Kapoor *et al.* (2012) mediram a qualidade de patentes do setor de energia eólica usando três indicadores de valor diferentes: (i) a contagem de citações para a

frente, (ii) O tamanho da família de patentes e (iii) a Classificação Internacional de Patentes (IPC).

Usando vários exemplos de acordos de licenciamento de tecnologia da indústria biofarmacêutica, Frattini, Bianchi e Franzó (2019) desenvolveram uma estrutura baseada em pesquisa de negociação, que identifica os fatores que afetam o preço de uma licença de tecnologia e os mecanismos pelos quais eles exercem sua influência (SONG; LEKER, 2019; THOMAS; CHUGAN, 2020). Outros trabalhos têm sido conduzidos usando grande bases de dados de licenciamentos na indústria biofarmacêutica, seja para diferenciar estratégias distintas de inovação (SONG; LEKER, 2019) ou para analisar determinantes no nível da empresa que afetariam o acordo de licenciamento (THOMAS; CHUGAN, 2020).

No contexto de análise da qualidade das patentes usando abordagens de aprendizado de máquina, Chi e Wang (2022) criaram um modelo de redes neurais de aprendizado profundo para prever a probabilidade de aprovação de pedidos de patentes. Já Lai e Che (2009) propuseram um modelo preditivo do valor jurídico/legal monetário das patentes para apoiar sentenças de indenização ou para negociar a taxa de liquidação em litígios de processos de violação de patentes. Han e Sohn (2015) usaram mineração de texto para identificar variáveis/indicadores importantes relacionadas ao valor da patente, sendo este valor medido pelo tempo de vida da patente usando uma análise de sobrevivência.

Por fim, Kim *et al.* (2019) propuseram um modelo de avaliação de tecnologias de origem universitária e concentram sua atenção nos indicadores relativos ao campo e escopo tecnológico (classes IPC), novidade e originalidade (qualidade das citações usadas na patente e Índice *Herfindah*³), proteção e cobertura de mercado (reivindicações e família da patente), esforços e capacidades de desenvolvimento (colaboração e número de inventores; quantidade e qualidade de invenções do titular no campo da invenção) e prévio desempenho acadêmico (quantidade e qualidade da produção acadêmica do inventor principal).

Uma leitura geral dos trabalhos sobre características de patentes e qualidade ou valores destas, percebe-se que variáveis preditoras do potencial de licenciamento das patentes têm relação com a qualidade técnica, mercadológica e legal destas. Assim, uma série de indicadores podem ser úteis para medir uma ou mais dessas qualidades, por isso alguns estudos dão ênfase a um ou outro aspecto da qualidade das patentes e não necessariamente a todos. Com base nisso, o quadro 1 apresenta as variáveis que serão testadas no modelo de aprendizado de máquina neste estudo.

³ Índice Herfindah é uma medida da dimensão das **empresas** relativamente à sua **indústria** e é um indicador do grau de concorrência entre elas

Quadro 1 - Variáveis da pesquisa

Variável	Conceito operacional	Fonte
Titularidade	Quantidade de instituições que participaram ativamente do desenvolvimento da invenção.	(LEE, 2008)
Citações recebidas (<i>forward citations</i>)	Quando a patente é citada por outras patentes posteriores	(FISCHER; LEIDINGER, 2014/4).
Escopo da patente	Quantidade de aplicações da patente. E medido pela quantidade de classificações IPC (<i>International Patent Classification</i>)	(SQUICCIARINII; DERNISI; CRISCUOLO, 2013)
Tamanho da Família.	Quantidade de territórios onde a patente já foi depositada.	(EPO, 2020; SQUICCIARINII; DERNISI; CRISCUOLO, 2013)
Vida Útil	Refere-se ao tempo para que a patente expire.	(LEE <i>et al.</i> , 2018)
Reivindicações	Refere-se à extensão de proteção das patentes, ou seja, definem e delimitam os direitos da invenção.	(TANTIYASWASDIKUL, 2014; WEILING; HAO, 2018; WU; WELCH; HUANG, 2015)
Reivindicações dependentes	São associadas a uma reivindicação independente fazendo à esta última uma referência explícita. Esta categoria de reivindicação faz o papel de complemento da reivindicação independente.	(BARNEY, 2004)
Inventores	Indica a quantidade de recursos humanos/conhecimento envolvidos no processo.	(TAHMOORESNEJAD; BEAUDRY, 2019)
Produtividade do Inventor Principal	Número de patentes que o inventor principal já depositou.	(KIM <i>et al.</i> , 2019)
PCT ou Depósito Fora	Quando a patente tem um depósito via Tratado de Cooperação em matéria de Patentes ou pelo menos um em um país diferente do original.	(BARNEY, 2004)
<i>No Patent Citation</i>	Refere-se a quantidade de documentos, que não são patentes, citados no texto. Indicam qualidade significativamente superior às patentes que não citam a literatura científica Branstetter (2005).	(VELAYOS-ORTEGA; LÓPEZ-CARREÑO, 2021)
Citações de patentes (<i>backward citations</i>).	Refere-se a quantidade de patentes citadas no documento	(HWANG; KIM; JEONG, 2021)

Força Tecnológica	Refere-se a soma de citações recebidas de todas as patentes coletadas de um campo tecnológico em uma universidade.	(KIM <i>et al.</i> , 2019)
-------------------	--	----------------------------

3 MÉTODO

Quanto ao objetivo é um estudo exploratório, pois verifica a influência das características no licenciamento. Quanto a abordagem é um estudo quantitativo, pois usa técnicas de aprendizado de máquina para captar as características de uma população ao fazer inferências a partir das características da amostra. Quanto ao método de coleta de dados é um estudo de monitoramento, pois registram-se as informações disponíveis através de observações de patentes em banco de dados. É um estudo *ex post facto*, pois o pesquisador não manipula as variáveis mas apenas observa o que já aconteceu, ou seja, os licenciamentos realizados (COOPER; SCHINDLER, 2013).

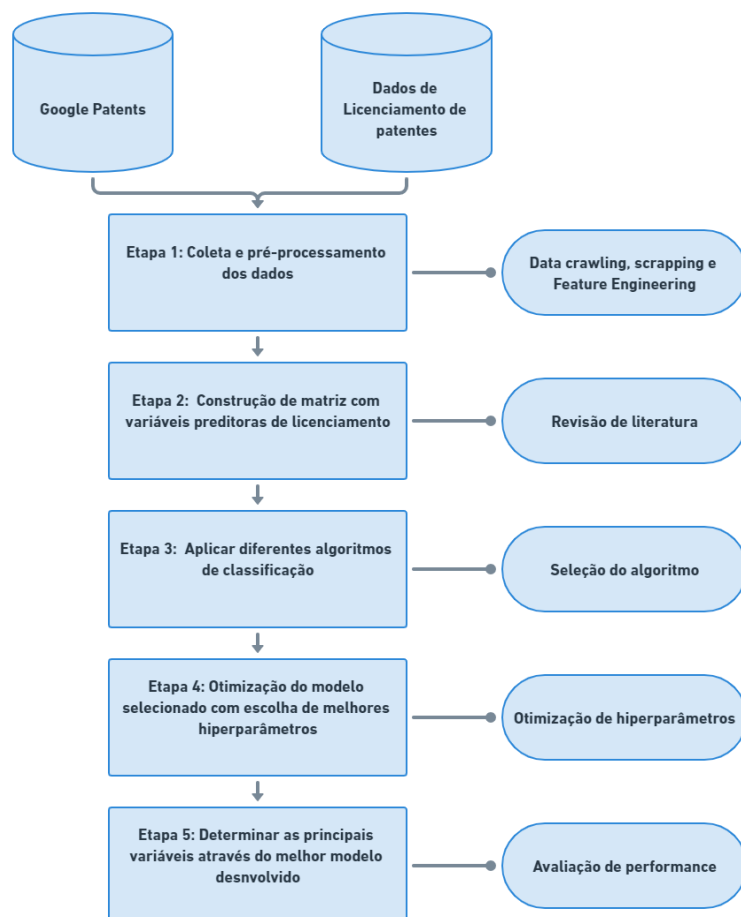
O método de amostragem é do tipo não probabilística e intencional, pois é escolhido de acordo com as características específicas das universidades licenciadoras, isto é, aquelas que têm mais licenciamentos nos campos tecnológicos objeto de estudo. É também um estudo longitudinal, pois coletou dados de licenciamento de patentes de ICT no Brasil de 1998 a 2022.

A amostra selecionada justifica-se por dois motivos. Primeiro os campos tecnológicos biofarmacêuticos representam a maior porcentagem de licenciamentos de patentes (35,64%) das ICT públicas no Brasil. Segundo, a escolha das 8 universidades para análise deve-se ao critério de maior produtividade de licenciamentos nesses dois campos tecnológicos. As universidades líderes nos últimos 24 anos foram Universidade Federal de Minas Gerais (UFMG), Universidade de Campinas (UNICAMP), Universidade de São Paulo (USP), Universidade Federal de São Carlos (UFScar), Universidade de Brasília (UnB), Universidade Federal de Viçosa (UFV), Universidade Federal do Paraná (UFPR), Universidade Federal do Rio Grande do Sul (UFRGS), respectivamente (LEITE *et al.*, 2022).

3.1 PERCURSO METODOLÓGICO

A figura 1 descreve o percurso metodológico.

Figura 1 - Percurso metodológico



Fonte: Elaborada pelos autores (2022)

Conforme a figura 1, na etapa 1 mineramos informações sobre 99 licenciamentos de patentes junto aos Núcleos de Inovação Tecnológica das ICT e nos diários oficiais através de buscas manuais. Em seguida, utilizamos a plataforma Google Patents como base para buscar (*crawling*) e coletar (*scrapping*) características dessas patentes licenciadas e das não licenciadas pelas ICT selecionadas. Ainda nesta etapa ocorrem as atividades de limpeza, filtragem e organização dos dados para aplicação das próximas etapas.

Na etapa 2, uma síntese da literatura relacionada a indicadores de qualidade/valor de patentes foi conduzida para elaborar uma matriz de variáveis preditoras do potencial de licenciamento das patentes analisadas. Estruturamos nossa matriz de características com 13 variáveis: Titularidade, Citações recebidas, Escopo da patente, Família da patente, Vida útil da patente, Reivindicações, Reivindicações independentes, Número de inventores, Produtividade

do inventor principal, PCT ou Depósito Fora, *No Patent Citations*, Citações, Força Tecnológica. Essas variáveis analisadas têm relação com fatores técnicos, mercadológicos e legais das patentes, conforme Quadro 1.

Na etapa 3, foram aplicados três algoritmos de classificação: *Random Forest*, Redes Neurais e Máquina de Vetores de Suporte e verificou-se que o modelo de *Random Forest* alcançava desempenho melhor que as outras duas. Esses três modelos vêm sendo largamente aplicados em problemas não lineares (NIKOU; MANSOUR FAR; BAGHERZADEH, 2019). Em seguida, na etapa 4, aplicou-se a otimização de hiperparâmetros do algoritmo de *Random Forest* para classificação de patentes licenciadas e não licenciadas.

Por fim, na etapa 5 analisou-se o resultado do desempenho do modelo gerado na etapa anterior e determinamos os atributos mais relevantes para prever a probabilidade de licenciamento de uma patente biofarmacêutica.

3.2 RANDOM FOREST

Breiman (2001) propôs o algoritmo de *Random Forest* que consiste em um conjunto N de árvores de decisão que votam em uma classe para realizar a predição. Por ser um algoritmo de alta performance (RAY, 2019) vem sendo utilizado em diversas aplicações na literatura, desde previsão de preços de ações (VIJH *et al.*, 2020), previsão de câncer (SHAIKH; RAO, 2022) e mensuração do valor econômico de patentes (KIM *et al.*, 2019). Algumas das principais vantagens da *Random Forest* é a habilidade de inferir relações complexas e não lineares entre variáveis (KIM *et al.*, 2019).

Esse algoritmo foi implementado em *Python* com a biblioteca *scikit-learn*. A base de dados foi balanceada e, em seguida, dividida em 75% dos dados para treinamento e 25% dos dados para teste do modelo. Devido ao desbalanceamento dos dados selecionamos a técnica de sub amostragem randômica visando balancear a base reduzindo a classe majoritária (patentes não licenciadas).

Para a implementação de um modelo em *Random Forest* com a biblioteca *scikit-learn*, os principais hiperparâmetros a serem ajustados são o número de árvores a serem utilizadas (*na estimators*) e o cálculo de importância dos atributos (*criterion*).

Visando a escolha de melhores hiperparâmetros para o modelo foi utilizada a classe *GridSearch* da biblioteca em conjunto com validações manuais para revelar combinações de hiperparâmetros que alcançassem melhor acurácia.

Para avaliação da performance do modelo utilizamos métricas de *accuracy*, *precision* e *recall*.

4 RESULTADOS

4.1 ANÁLISE DESCRITIVA

Os dados são obtidos a partir de informações disponíveis em patentes licenciadas e não licenciadas de 8 (oito) instituições acadêmicas brasileiras líderes em licenciamentos nos campos tecnológicos de biotecnologia e farmacêutica. As instituições são Universidade Federal de Minas Gerais (UFMG), Universidade de Campinas (UNICAMP), Universidade de São Paulo (USP), Universidade Federal de São Carlos (UFScar), Universidade de Brasília (UnB), Universidade Federal de Viçosa (UFV), Universidade Federal do Paraná (UFPR), e Universidade Federal do Rio Grande do Sul (UFRGS).

Tabela 1 - Quantitativo de patentes licenciadas e não licenciadas especificamente nos campos tecnológicos biotecnologia e farmacêutico juntos.

Universidade	Licenciadas	Não licenciadas	Total
Universidade Federal de Minas Gerais (UFMG)	23	536	559
Universidade de Campinas (Unicamp)	22	360	382
Universidade de São Paulo (USP)	18	499	517
Universidade Federal de São Carlos (UFScar)	11	37	48
Universidade de Brasília (UnB)	8	73	81
Universidade Federal de Viçosa (UFV)	7	104	111
Universidade Federal do Paraná (UFPR)	5	129	134
Universidade Federal do Rio Grande do Sul (UFRGS)	5	246	251
Totais	99	1984	2083

Fonte: Elaborada pelos autores.

A Tabela 1 apresenta os quantitativos de patentes licenciadas e não licenciadas de cada instituição no campo tecnológico em análise. Percebe-se que a UFMG é a que mais licenciou nesses campos tecnológicos e dessa lista de 8 universidades, 2 são de São Paulo e são estaduais. Juntas, UFMG, Unicamp e USP representam mais de 50% dos licenciamentos desta base. Apesar de possuir o quarto maior número de patentes, a UFRGS tem baixos resultados em licenciamento.

4.2 AVALIAÇÃO DO MODELO PARA PREVER LICENCIAMENTO.

Neste tópico é apresentado um comparativo de desempenho entre três modelos de aprendizado de máquina para os dados analisados através das métricas de *accuracy*, *precision* e *recall*. Em seguida é apresentado o desempenho e a importância relativa de cada variável para o modelo final.

4.2.1 Escolha do melhor algoritmo de classificação

A tabela 2 mostra um comparativo simplificado entre os modelos de classificação para os dados analisados. O resultado mostra que o algoritmo de *Random Forest* alcançou o melhor desempenho considerando as métricas de *accuracy*, *precision* e *recall*.

Tabela 2 - Comparação dos modelos

Modelo	Categoria	Acurácia	<i>Precision</i>	<i>Recall</i>
Redes Neurais	0 (não licenciada)	0.73	0.70	0.73
	1 (licenciada)		0.77	0.74
Máquinas de Vetores de Suporte	0 (não licenciada)	0.73	0.78	0.69
	1 (licenciada)		0.69	0.78
Random Forest	0 (não licenciada)	0.78	0.83	0.73
	1 (licenciada)		0.73	0.83

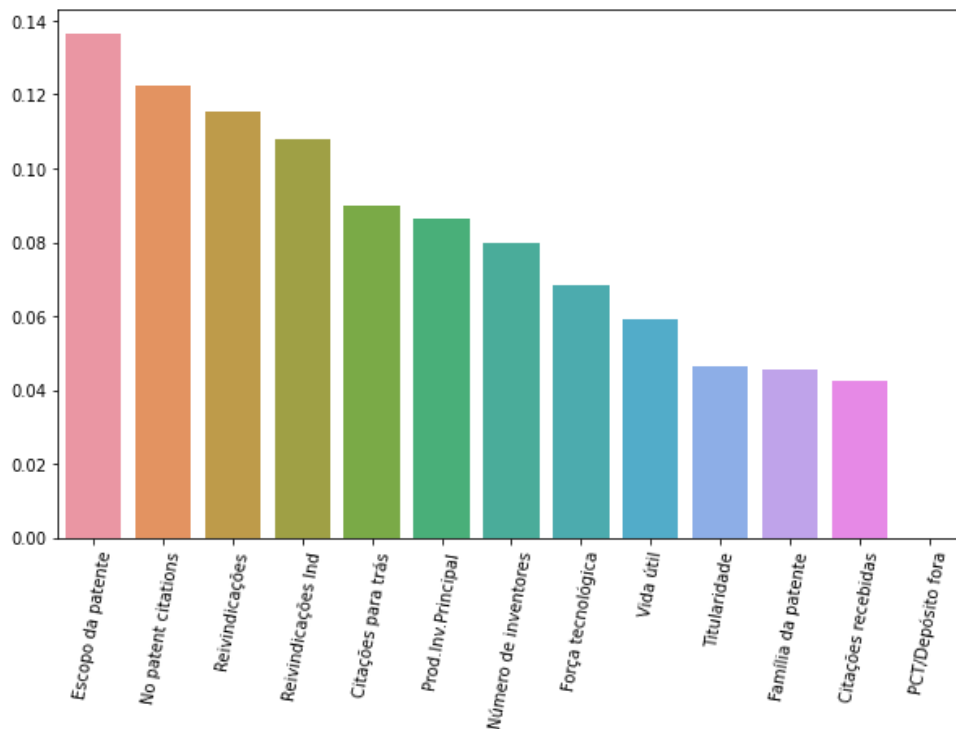
Fonte: Elaborada pelos autores.

Observando a tabela 2 nota-se que o modelo em *Random Forest* obteve a maior *accuracy* (78%) com os hiperparâmetros de *n_estimators*=100, *criterion*=entropy, e *random_state*=0. Contudo essa métrica pode ainda ser muito ingênua para determinar o melhor modelo, por isso é importante observarmos as métricas *precision* e *recall* para compreendermos o quão bem cada modelo está performando para cada classe. Desse modo observa-se que, de maneira geral, as pontuações da *Random Forest* (*Precision* classe 0: 83%; *Precision* classe 1 73%; *Recall* classe 0 73%; *Recall* classe 1 83%) são maiores que as dos outros modelos com exceção da *precision* para a classe “1”, em que Redes Neurais estão 4% acima (77%).

4.2.2 Importância relativa das variáveis no modelo final

Em seguida, a figura 2 exibe a importância relativa de cada variável para o modelo. Este resultado é gerado a partir da diminuição total das impurezas do nó da divisão da variável, calculada em média para todas as árvores.

Figura 2 - Importância relativa das variáveis



Fonte: Elaborada pelos autores.

Pode-se observar na figura 2 que as quatro variáveis mais influentes são o escopo da patente, *no patent citations*, reivindicações e reivindicações independentes, respectivamente nessa ordem. Já as variáveis menos influentes são titularidade, família de patente, citações recebidas e PCT/Depósito fora.

4.2.3 Avaliação da performance do modelo

A matriz de confusão mostra uma relação entre valores previstos (pelo modelo) e os valores reais (o que os valores de fato eram) para cada classe.

Tabela 3 - Resultados da avaliação de performance

(a) Matriz de confusão

		Previsto	
		Não Licenciada	Licenciada
Real	Não Licenciada	20	6
	Licenciada	3	20

(b) Avaliação da performance das métricas

	Não Licenciada	Licenciada
<i>Precision</i>	87%	77%
<i>Recall</i>	77%	87%
<i>OOB Error Rate</i>	31%	
<i>Accuracy</i>	82%	

Fonte: Elaborada pelos autores.

Conforme a tabela 3 , o modelo apresentou uma *accuracy* de 82%. Isso significa que o modelo acertou 82% das previsões nos dados de teste. O *precision* de 87% para 'não licenciadas' significa que quando o modelo previu 'não licenciada' estava certo em 87% dos casos, enquanto o *precision* de 77% para as 'licenciadas' significa que quando o modelo previu 'não licenciada' estava certo em 77% dos casos. Enquanto no *recall* as 'não licenciadas' alcançaram 87%, o que significa que quando a classe de fato era 'não licenciada' o modelo acertou 87% e as licenciadas 77%, o que significa que quando a classe de fato era 'licenciada' o modelo acertou 77% dos casos.

5 DISCUSSÃO DE RESULTADOS

Embora seja reconhecido na literatura o impacto das variáveis mercadológicas, legais e técnicas no licenciamento de patentes, muito pouco se sabe sobre o impacto dessas características nas patentes acadêmicas brasileiras no campo de biotecnologia e farmacêutica. Assim, à luz da necessidade de entender essa questão, exploramos como a titularidade, citações recebidas, escopo da patente, família da patente, vida útil da patente, reivindicações, reivindicações independentes, número de inventores, produtividade do inventor principal,

PCT ou depósito fora, *no patent citations*, citações de patentes, força tecnológica influenciam na chance de licenciamento dessas patentes. Para explorar essa questão utilizamos a abordagem de aprendizado de máquina com o algoritmo de *Random Forest* em um grupo de 99 patentes licenciadas e 1984 patentes não licenciadas das ICT públicas brasileiras com melhor desempenho em licenciamentos nesses campos tecnológicos.

Constatamos que as variáveis mais influentes para prever um licenciamento acadêmico nos campos tecnológicos analisados são o escopo da patente, *no patent citations*, reivindicações e reivindicações independentes, respectivamente nessa ordem. De fato, quanto maior o escopo da patente, ou seja, maiores possibilidades de aplicação de uma patente, mais valiosa ela pode ser, especialmente quando possíveis substitutos na mesma classe de produto são disponíveis (BARNEY, 2004; LERNER, 1994), sendo provado em estudo robusto especialmente dentro do campo tecnológico de biotecnologia e farmacêutica (SQUICCIARINII; DERNISI; CRISCUOLO, 2013). Como esperávamos, um quantitativo de "*no patent citation*" - quantidade de citações a documentos que não são patentes - indica qualidade significativamente superior às patentes que não citam a literatura científica (BRANSTETTER, 2005).

Especificamente quanto ao escopo da patente e *no patent citation* nossos resultados estão de acordo com Kim *et al.* (2019) quando também determina que essas duas variáveis estão entre as 5 mais influentes na predição do valor de tecnologias originadas da academia.

Além disso, esses resultados estão em harmonia com teorias que indicam que o número de "*no patent citation*" determina a atualidade da bibliografia utilizada e, portanto, a tecnologia a ser patenteada é considerada mais inovadora, no entanto, quando as patentes citam apenas outras patentes anteriores, elas trazem apenas melhorias para tecnologias patentadas (VELAYOS-ORTEGA; LÓPEZ-CARREÑO, 2021).

Quanto à importância relativa das reivindicações - extensão de proteção das patentes, ou seja, definem e delimitam os direitos da invenção, conforme art. 41 da Lei de propriedade industrial (BRASIL, 1996) - confirma-se que, realmente quanto mais reivindicações independentes mais valiosa é a patente (KWON; RYU; PARK, 2013), sendo que patentes com reivindicações mais curtas e bem elaboradas (palavras técnicas) são mais valiosas, pois tem mais poder legal (BARNEY; BARNEY, 2003). Além disso, essa relação é significativamente positiva quando se trata do valor da patente universitária (TANTIYASWASDIKUL, 2014). Não é à toa que variável têm relação direta com a estabilidade dos direitos da patente, ou seja, o quanto a patente tem potencial para manter-se válida quando questionada sobre sua novidade.

Especificamente quando as reivindicações, nossos resultados estão em harmonia com a literatura científica e tecnológica que têm mostrado estreita relação positiva com o valor da patente (BARNEY, 2004; SQUICCIARINI; DERNISI; CRISCUOLO, 2013; TANTIYASWASDIKUL, 2014) pois mede sua força legal (LIU; CHANG, 2022; WANG; HSIEH, 2015/3; WILLIAMS, 2006).

6 CONCLUSÃO

Este trabalho propôs uma abordagem de aprendizado de máquina para identificar quais características são influentes para o licenciamento de patentes acadêmicas brasileiras nos campos tecnológicos biotecnologia e farmacêutica. Para isso, primeiro foi necessário a coleta de dados em fontes diversas para alimentar o modelo final. Chegamos a um modelo em *Random Forest* com *accuracy* de 82% e observamos que as variáveis com mais importância para o modelo são escopo da patente, no patent citation e o número de reivindicações.

Esperamos contribuir com melhor compreensão de quais aspectos norteiam a determinação do potencial de licenciamento de uma patente. Além disso, esse estudo tenta preencher lacunas científicas respondendo questões escassas no contexto da América Latina e Brasil quando se trata de fatores que influenciam o potencial de licenciamentos de patentes acadêmicas.

Esse trabalho também contribui para melhores decisões dos gestores de propriedade intelectual dessas ICT com suporte teórico para empreender ações estratégicas que podem aumentar o desempenho em licenciamentos das suas patentes. Por fim, pode-se aumentar receitas públicas com novas estratégias, reduzindo a dependência de recursos federais.

O trabalho enfrenta algumas limitações devido ao tamanho reduzido da base de dados e a dificuldade de acesso a dados de licenciamento no Brasil. Além disso, nossas variáveis preditoras foram obtidas apenas de dados contidos em documentos de patentes.

Sugerimos para trabalhos futuros que se utilize mais dados de licenciamento de patentes para melhor aprendizado do modelo. Também é possível explorar novas variáveis e verificar sua importância.

REFERÊNCIAS

ABBAS, A.; ZHANG, L.; KHAN, S. U. **A literature review on the state-of-the-art in patent analysis**. 2014.

ALVAREZ, R.; LÓPEZ, R. A. Foreign Technology Acquisition and Changes in the Real Exchange Rate. **The World Economy**, v. 38, n. 4, p. 613–628, 31 abr. 2015. Disponível em: <<http://doi.wiley.com/10.1111/twec.12253>>.

ANAND, B. N.; KHANNA, T. The Structure of Licensing Contracts. **The Journal of industrial economics**, v. 48, n. 1, p. 103–135, 27 mar. 2003. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-6451.00114>>.

BARNEY, J. **Method and system for valuing intangible assets**. 20040010393:A1. Concessão: 15 Jan. 2004.

BRANSTETTER, L. Exploring the Link Between Academic Science and Industrial Innovation. **Annals of economics and statistics**, n. 79-80, p. 119–142, 2005. Disponível em: <<https://ideas.repec.org/a/adr/anecst/y2005i79-80p119-142.html>>. Acesso em: 2 maio. 2022.

BRASIL. **Lei 9.279, de 14 de Maio de 1996. Regula direitos e obrigações relativos à propriedade industrial**. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/L9279.htm>. Acesso em: 13 ago. 2016.

BREIMAN, L. Random Forests. **Machine learning**, v. 45, n. 1, p. 5–32, 1 out. 2001. Disponível em: <<https://doi.org/10.1023/A:1010933404324>>.

BURGELMAN, R. A.; CHRISTENSEN, C. M. WHEELWRIGHT, S. C. **Gestão Estratégica da Tecnologia e da Inovação: Conceitos e Soluções**. [s.l.] AMGH Editora, 2013. 638 p.

CHEN, M.-S.; HAN, J.; YU, P. S. Data mining: an overview from a database perspective. **IEEE transactions on knowledge and data engineering**, v. 8, n. 6, p. 866–883, dez. 1996. Disponível em: <<http://dx.doi.org/10.1109/69.553155>>.

CHESBROUGH, H. The Logic of Open Innovation: Managing Intellectual Property. **California management review**, v. 45, n. 3, p. 33–58, 1 abr. 2003. Disponível em: <<https://doi.org/10.1177/000812560304500301>>.

CHI, Y.-C.; WANG, H.-C. Establish a patent risk prediction model for emerging technologies using deep learning and data augmentation. **Advanced Engineering Informatics**, v. 52, p. 101509, 1 abr. 2022. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1474034621002585>>.

COOPER, D. R.; SCHINDLER, P. S. **Business research methods**. New York: McGraw-Hill/Irwin, 2013. v. Twelfth edition

DE CLERCQ, D.; DIOP, N. F.; JAIN, D.; TAN, B.; WEN, Z. **Multi-label classification and interactive NLP-based visualization of electric vehicle patent data**. 2019.

DE MOORTEL, K.; CRISPEELS, T. International university-university technology transfer: Strategic management framework. **Technological forecasting and social change**, 18 maio 2018. Disponível em:

<<http://www.sciencedirect.com/science/article/pii/S0040162518307418>>.

EPO. DOCDB Simple Patent Family. 2020. Disponível em:

<<https://www.epo.org/searching-for-patents/helpful-resources/first-time-here/patent-families/docdb.html>>. Acesso em: 2 maio. 2020.

FEINERER, I.; HORNIK, K.; MEYER, D. Text Mining Infrastructure in R. **Journal of statistical software**, v. 25, p. 1–54, 31 mar. 2008. Disponível em:

<<https://www.jstatsoft.org/article/view/v025i05>>. Acesso em: 7 jul. 2022.

FISCHER, T.; LEIDINGER, J. Testing patent value indicators on directly observed patent value—An empirical analysis of Ocean Tomo patent auctions. **Research policy**, v. 43, n. 3, p. 519–529, 2014/4. Disponível em:

<<http://www.sciencedirect.com/science/article/pii/S0048733313001297>>.

FRATTINI, F.; BIANCHI, M.; FRANZÓ, S. Identifying Factors Influencing the Price of Technology Licenses: A Framework Grounded in Negotiation Research. **International Journal of Innovation Management**, v. 23, n. 05, p. 1950039, 1 jun. 2019. Disponível em:

<<https://doi.org/10.1142/S1363919619500397>>.

HAN, E. J.; SOHN, S. Y. Patent valuation based on text mining and survival analysis. **The Journal of technology transfer**, v. 40, n. 5, p. 821–839, 2015. Disponível em:

<<http://dx.doi.org/10.1007/s10961-014-9367-6>>.

HOFMAN, J.; NIKLASSON, A. **Success Factors in Product Licensing in the Pharmaceuticals Industry Identification and evaluation of factors influencing likelihood and financial value of a licensing deal Master's Thesis in the Master's Programme Entrepreneurship and Business Design**. Disponível em:

<<https://odr.chalmers.se/bitstream/20.500.12380/244578/1/244578.pdf>>. Acesso em: 27 jun. 2022.

HWANG, J.-T.; KIM, B.-K.; JEONG, E.-S. Patent Value and Survival of Patents. **Journal of Open Innovation: Technology, Market, and Complexity**, v. 7, n. 2, p. 119, 26 abr. 2021.

Disponível em: <<https://www.mdpi.com/2199-8531/7/2/119>>. Acesso em: 10 jul. 2022.

KAPOOR, R.; KARVONEN, M.; LEHTOVAARA, M.; KÄSSI, T. Patent value indicators: Case of emerging wind energy markets. Em: 2012 Proceedings of PICMET '12: Technology Management for Emerging Technologies, 2012, [...]. 2012. p. 1042–1048.

KIM, Y.; AHN, J. M.; KWON, O.; LEE, C. Valuation of University-Originated Technologies: A Predictive Analytics Approach. **IEEE Transactions on Engineering Management**, p. 1–13, 2019. Disponível em: <<http://dx.doi.org/10.1109/TEM.2019.2938182>>.

LAI, Y.-H.; CHE, H.-C. Integrated evaluator extracted from infringement lawsuits using extension neural network accommodated to patent assessment. **International Journal of Computer Applications in Technology**, v. 35, n. 2-4, p. 84–96, 2009. Disponível em:

<<https://www.inderscienceonline.com/doi/abs/10.1504/IJCAT.2009.026585>>.

LEE, J. H.; SUNG, T.-E.; KIM, E.; SHIN, K. Evaluating determinant priority of license fee in biotech industry. **Journal of Open Innovation: Technology, Market, and Complexity**, v. 4, n. 3, 2018. Disponível em:

<<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85058292158&doi=10.3390%2fjoit>>.

mc4030030&partnerID=40&md5=2db734aad7b367bd6223c38601b6d164>.

LEE, Y.-G. Patent Licensability and Life: A Study of U.S. Patents Registered by South Korean Public Research Institutes. **Scientometrics**, v. 75, n. 3, p. 463–471, jun. 2008. Disponível em: <<http://link.springer.com/10.1007/s11192-007-1879-5>>.

LEITE, R. A. S.; REIS, I. B.; DE SOUSA BARROS LIMA, S.; REIS, A. A. M. Perfil de Licenciamentos de Patentes Acadêmicas no Brasil. (IFPI) **In: Semana Nacional de Ciência e Tecnologia**, 2022, Florianópolis. [...]. Florianópolis: Instituto Federal do Piauí, 2022.

LEITE, R. A. S.; REIS, I. B.; DE SOUSA BARROS LIMA, S.; REIS, A. A. M. Perfil de Licenciamentos de Patentes Acadêmicas no Brasil. Em: SNCT 2022, 2022, Florianópolis. [...]. Florianópolis: Instituto Federal do Piauí, 2022.

LERNER, J. The Importance of Patent Scope: An Empirical Analysis. **The Rand Journal of economics**, v. 25, n. 2, p. 319–333, 1994. Disponível em: <<http://www.jstor.org/stable/2555833>>.

LI, L.; YANG, Z.; DANG, Z.; MENG, C.; HUANG, J.; MENG, H.; WANG, D.; CHEN, G.; ZHANG, J.; PENG, H.; SHAO, Y. Propagation Analysis and Prediction of the COVID-19. **Infectious Disease Modelling**, v. 5, p. 282–292, 31 mar. 2020. Disponível em: <<http://dx.doi.org/10.1016/j.idm.2020.03.002>>.

LIU, K.; CHANG, C. Number of Claims for Evaluation of Individual Patent Based on LiDAR Technology. 2022. Disponível em: <<http://www.jmest.org/wp-content/uploads/JMESTN42354006.pdf>>.

NIKOU, M.; MANSOURFAR, G.; BAGHERZADEH, J. Stock Price Prediction Using DEEP Learning Algorithm and Its Comparison with Machine Learning Algorithms. **Intelligent Systems in Accounting Finance & Management**, v. 26, n. 4, p. 164–174, out. 2019. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/isaf.1459>>.

RAY, S. A Quick Review of Machine Learning Algorithms. Em: 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), 2019, [...]. 2019. p. 35–39.

SHAIKH, F. J.; RAO, D. S. Prediction of Cancer Disease using Machine learning Approach. **Materials Today: Proceedings**, v. 50, p. 40–47, 1 jan. 2022. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2214785321027206>>.

SHEREMETYEVA, S. Natural Language Analysis of Patent Claims. Em: Proceedings of the ACL-2003 Workshop on Patent Corpus Processing, 2003, , [...]. ,: Association for Computational Linguistics, 2003. p. 66–73.

SOARES, T. J.; TORKOMIAN, A. L. V.; NAGANO, M. S. University regulations, regional development and technology transfer: The case of Brazil. **Technological forecasting and social change**, v. 158, p. 120129, 1 set. 2020. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0040162520309550>>.

SONG, C. H.; LEKER, J. Differentiation of innovation strategies based on pharmaceutical licensing agreements: Insight from Korean pharmaceutical firms. **Technology Analysis & Strategic Management**, v. 31, n. 2, p. 169–185, 1 fev. 2019. Disponível em:

<<https://doi.org/10.1080/09537325.2018.1490711>>.

SQUICCIARINII, M.; DERNISI, H.; CRISCUOLO, C. Measuring Patent Quality: Indicators of Technological and Economic Value. **OECD Science, Technology and Industry Working Papers**, p. 69, 2013. . Acesso em: 29 jun. 2020.

TAHMOORESNEJAD, L.; BEAUDRY, C. Capturing the economic value of triadic patents. **Scientometrics**, v. 118, n. 1, p. 127–157, 1 jan. 2019. Disponível em: <<https://doi.org/10.1007/s11192-018-2959-4>>.

TANTIYASWASDIKUL, K. Determinants of Patent Value in US and Japanese University Patents. **International Journal of Technical Research and Applications**, v. 2, n. Special 4, p. 8–12, 2014. Disponível em: <<http://www.ijtra.com/special-issue-view/determinants-of-patent-value-in-us-and-japanese-university-patents.pdf>>.

THOMAS, B.; CHUGAN, P. K. **Determinants of Licensing Deals in Pharmaceutical Industry**. 1 maio 2020. . Disponível em: <<https://papers.ssrn.com/abstract=3678644>>. Acesso em: 27 jun. 2022.

TSENG, Y.-H.; LIN, C.-J.; LIN, Y.-I. Text mining techniques for patent analysis. **Information processing & management**, v. 43, n. 5, p. 1216–1247, 1 set. 2007. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0306457306002020>>.

VELAYOS-ORTEGA, G.; LÓPEZ-CARREÑO, R. Non-Patent Literature. **Encyclopedia**, v. 1, n. 1, p. 198–205, 12 fev. 2021. Disponível em: <<https://www.mdpi.com/996998>>. Acesso em: 10 jul. 2022.

VIJH, M.; CHANDOLA, D.; TIKKIWAL, V. A.; KUMAR, A. **Stock Closing Price Prediction using Machine Learning Techniques****Procedia Computer Science**2020. Disponível em: <<http://dx.doi.org/10.1016/j.procs.2020.03.326>>.

WANG, B.; HSIEH, C.-H. Measuring the value of patents with fuzzy multiple criteria decision making: insight into the practices of the Industrial Technology Research Institute. **Technological forecasting and social change**, v. 92, p. 263–275, 2015/3. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0040162514002856>>.

WANG, M.; LIN, J.; LO, H. Influential factors of the commercialization of academic patents: The Taiwan experience. Em: 2012 Proceedings of PICMET '12: Technology Management for Emerging Technologies, 2012, [...]. 2012. p. 1830–1836.

WEILING, S.; HAO, Y. X. Z. The Patent Portfolios Value Evaluation Based on the TOPSIS-Entropy Method. Em: Proceedings of PICMET '18: Technology Management for Interconnected World, 2018, [...]. 2018. p. 1–14.

WILLIAMS, A. **Bias compensated method and system for patent evaluation**. 20060036453:A1. Concessão: 16 fev. 2006.

WOOLF, B. P. **Building Intelligent Interactive Tutors: Student-centered Strategies for Revolutionizing E-learning**. [s.l.] Morgan Kaufmann, 2010. 480 p.

WU, Y.; WELCH, E. W.; HUANG, W. L. Commercialization of university inventions:

Individual and institutional factors affecting licensing of university patents. **Technovation**, 2015. Disponível em:
<https://www.sciencedirect.com/science/article/pii/S0166497214001266?casa_token=7AxdMU-9F6QAAAAA:RAgqfkXGRzzygc1gOiqYaZcxTcPcwpPlXpAeM_vzb9NxQsAdvaKaFz1axRv7A10Yn_p4zK82mA>.

ZHYLINSKA, O.; SITNICKI, M.; VIKULOVA, A. Systematic Assessment Of The Innovative Potential Of A Research University. **Baltic Journal of Economic Studies**, v. 5, n. 2, p. 38–44, 13 maio de 2019. Disponível em:
<<http://www.baltijapublishing.lv/index.php/issue/article/view/633>>. Acesso em: 20 maio. 2019.

ZIEGLER, N.; RUETHER, F.; BADER, M. A.; GASSMANN, O. Creating value through external intellectual property commercialization: A desorptive capacity view. **The Journal of technology transfer**, v. 38, n. 6, p. 930–949, 2013. Disponível em:
<<https://www.scopus.com/inward/record.uri?eid=2-s2.0-84886726986&doi=10.1007%2fs10961-013-9305-z&partnerID=40&md5=c3d072139b56a7f0fb6a133f79765f3f>>.