

José Wendel dos Santos

eng.wendel@live.com

[Orcid: 0000-0002-9138-0437](https://orcid.org/0000-0002-9138-0437)

Doutorando em Ciência da Propriedade Intelectual (PPGPI), Universidade Federal de Sergipe, Aracaju, Sergipe, Brasil.

Leonôra Virginia de Jesus Dias

leonoradidas13@gmail.com

[Orcid: 0009-0004-7790-3846](https://orcid.org/0009-0004-7790-3846)

Mestra em Ciência da Propriedade Intelectual (PPGPI), Universidade Federal de Sergipe, Aracaju, Sergipe, Brasil.

José Barreto Netto

jbarretonetto@hotmail.com

[Orcid: 0000-0001-9387-0846](https://orcid.org/0000-0001-9387-0846)

Doutorando em Ciência da Propriedade Intelectual (PPGPI), Universidade Federal de Sergipe, Aracaju, Sergipe, Brasil.

Cristiane Monteiro de Farias

Rezende

chrismont@gmail.com

[Orcid: 0000-0002-7629-133X](https://orcid.org/0000-0002-7629-133X)

Doutoranda em Ciência da Propriedade Intelectual (PPGPI), Universidade Federal de Sergipe, Aracaju, Sergipe, Brasil.

Marcelo de Paula

marcelop@ufob.edu.br

[Orcid: 0000-0003-0234-7270](https://orcid.org/0000-0003-0234-7270)

Doutor em Estatística (PIPGES), Universidade Federal do Oeste da Bahia (UFOB), Barreiras, Bahia, Brasil.

Cristiane Toniolo Dias

cristianetdias@utfpr.edu.br

[Orcid: 0000-0002-7204-8930](https://orcid.org/0000-0002-7204-8930)

Doutora em Ciência da Propriedade Intelectual (PPGPI), Universidade Tecnológica Federal do Paraná (UTFPR), Curitiba, Paraná, Brasil.

Uso de transformadores generativos pré-treinados para agilizar o exame de pedidos de patente

RESUMO

A avaliação de patentes garante a legitimidade das invenções, mas os examinadores enfrentam o desafio de conciliar agilidade com análises precisas. Este estudo explora o potencial dos transformadores generativos pré-treinados (GPTs) para automatizar etapas do processo de exame de patentes. O GPT-4 foi treinado com um corpus de patentes, aprendendo a identificar padrões complexos e extrair informações relevantes. Após um refinamento com dados específicos sobre os requisitos de patenteabilidade, o modelo foi avaliado em novas tarefas de classificação de pedidos de patente. Os resultados indicaram uma alta precisão do GPT-4, embora a necessidade de reduzir a taxa de falsos negativos seja evidente. Acredita-se que a otimização dos hiperparâmetros e a expansão do conjunto de dados de treinamento possam aprimorar a capacidade do modelo de capturar nuances semânticas e conceituais, tornando-o uma ferramenta ainda mais eficaz na classificação precisa de pedidos de patente.

PALAVRAS-CHAVE: Aprendizado de máquina. Inteligência artificial. Patentes. Inovação.

INTRODUÇÃO

O sistema de patentes atua como um catalisador da inovação, incentivando a criação de novas tecnologias que impulsionam a competitividade das organizações nos diversos setores da economia. No entanto, os escritórios de patentes ao redor do mundo têm enfrentado críticas devido à morosidade e à onerosidade do processo de exame de pedidos de patentes, com prazos que se estendem por anos (Yang; Zhang; Liu, 2024). Essa lentidão gera uma série de efeitos negativos tanto para os inventores quanto para a sociedade em geral – insegurança jurídica, perda de competitividade e retardo do progresso tecnológico e econômico dos países. Entre os principais fatores que contribuem para essa situação estão a escassez de examinadores qualificados e a crescente complexidade das invenções submetidas à análise, resultando em um aumento significativo no volume de informações contidas em cada pedido (Chi; Wang, 2022).

A crescente demanda por agilidade e precisão na tramitação dos pedidos de patente tem impulsionado a busca por soluções inovadoras. Recentemente, técnicas de aprendizagem profunda, como os transformadores generativos pré-treinados (GPTs), têm mostrado grande potencial para tarefas de processamento de linguagem natural (PLN). Os GPTs são uma classe de modelos de Inteligência Artificial (IA) que permitem a geração e análise de textos complexos com base em padrões aprendidos. Esses modelos são pré-treinados em um grande volume de dados e podem ser ajustados para executar tarefas semelhantes às dos seres humanos (Deeb *et al.*, 2024; Nguyen-Mau *et al.*, 2024). Estudos recentes já demonstraram o potencial dos modelos GPT na triagem automatizada de revisões sistemáticas de literatura (Guo *et al.*, 2024; Gargari *et al.*, 2024; Landschaft *et al.*, 2024). Portanto, é razoável esperar que os GPTs também possam ser treinados para identificar nuances semânticas utilizadas em pedidos de patentes.

Diante desse cenário promissor, este estudo teve como objetivo analisar o potencial da aplicação de GPTs na triagem inicial do processo de exame de pedidos de patente. A justificativa para a condução deste estudo se fundamenta nos benefícios que a implementação dessa tecnologia pode trazer aos escritórios de patentes. O uso de GPTs para tarefas repetitivas e de baixo risco permite que os examinadores se dediquem a tarefas mais complexas. Essa automatização pode ainda reduzir custos do processo, minimizar falhas e melhorar a qualidade das decisões dos examinadores, aumentando a confiabilidade dos escritórios de patentes.

METODOLOGIA

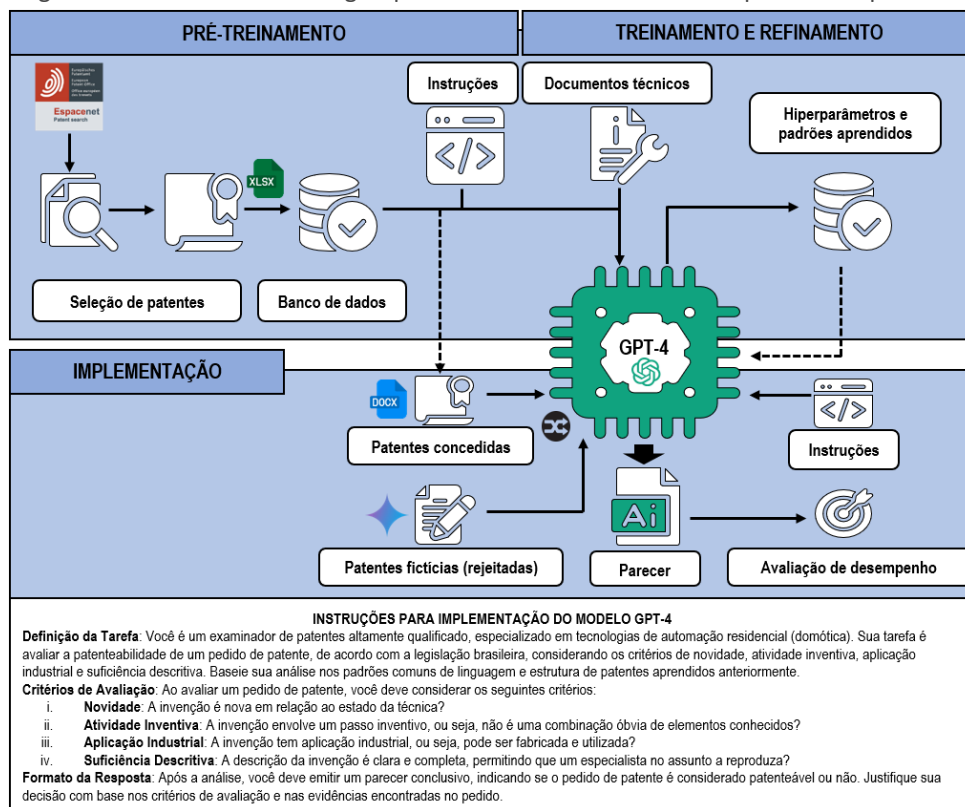
Este estudo propôs uma abordagem exploratória para a avaliação preliminar de pedidos de patente, utilizando o modelo de linguagem GPT-4 (OpenAI®) com o módulo Scholar AI® integrado (Figura 1). Esse módulo foi empregado para complementar o pré-treinamento do GPT-4, proporcionando acesso a um corpus relevante de literatura científica. Durante a fase de treinamento, o GPT-4 foi alimentado com um conjunto de patentes concedidas por examinadores humanos, relacionadas à tecnologia de automação residencial (domótica), extraídas do Espacenet® em 1º de agosto de 2024. Dentre os 164.042 documentos encontrados e ordenados por relevância pela própria base de dados, os primeiros 500 tiveram seus status legais analisados manualmente. Após a exclusão dos documentos não

correspondentes a patentes concedidas, 200 patentes foram selecionadas aleatoriamente, exportadas e traduzidas automaticamente para o português, compondo assim a amostra deste estudo. Dessas, 150 patentes foram utilizadas para treinar o GPT-4 a identificar padrões semânticos característicos de patentes de domótica, enquanto as 50 restantes foram reservadas para a fase de implementação.

Na fase de refinamento, foram aplicadas técnicas de aprendizado de máquina supervisionado para otimizar os hiperparâmetros do GPT-4. Ao longo de cinco dias consecutivos, o modelo foi exposto a um conjunto diversificado de dados, incluindo relatórios técnicos, cartilhas, normas e legislações, com o objetivo de aprimorar sua capacidade de identificar características de patenteabilidade e se adaptar ao contexto legal brasileiro.

Na fase de implementação, o GPT-4 foi configurado para executar o exame das patentes, classificando-as em duas categorias: patenteável ou não patenteável. A classificação foi baseada em critérios utilizados por examinadores humanos, como novidade, atividade inventiva, aplicação industrial e suficiência descritiva. Para avaliar a capacidade do modelo em identificar violações aos princípios fundamentais do direito patentário, foram geradas 50 patentes fictícias relacionadas à domótica, utilizando a IA Gemini®. Essas patentes continham falhas intencionais em relação aos critérios de patenteabilidade, como a descrição de ideias abstratas ou a ausência de um invento concreto. Assim, um arquivo contendo 100 documentos de patente, composto por 50 patentes reais e 50 patentes fictícias, foi carregado na interface do sistema.

Figura 1 – Estrutura metodológica para uso do GPT-4 no exame de pedidos de patentes



Fonte: Elaborada pelos autores (2024).

O corpus textual, constituído por 200 patentes extraídas do Espacenet®, foi submetido a uma análise lexical utilizando o Iramuteq® (versão 0.7 alfa 2). Essa ferramenta, que integra recursos do R® e do Python®, possibilitou a identificação de diferenças lexicais entre os textos, evidenciando a heterogeneidade da amostra. Em seguida, procedeu-se à implementação e à avaliação do desempenho do GPT-4 com as 100 patentes (50 reais e 50 fictícias), utilizando métricas de classificação binária obtidas a partir da matriz de confusão. A precisão (*P*) mede a proporção de patentes corretamente classificadas como patenteáveis, enquanto a revocação (*R*) quantifica a proporção de patentes realmente patenteáveis que foram corretamente identificadas pelo GPT-4. O *F1-score*, por sua vez, representa a média harmônica de *P* e *R*, oferecendo uma medida balanceada do desempenho geral do modelo. Mais detalhes sobre o cálculo dessas métricas podem ser encontrados em Kulkarni, Chong e Batarseh (2020).

RESULTADOS E DISCUSSÃO

Dentre as duzentas patentes de automação residencial analisadas para treinar e implementar o GPT-4 no exame de patenteabilidade, 75% foram concedidas nos últimos dois anos. A maior parte dessas patentes é de titularidade de grandes corporações chinesas (41,3%) e americanas (36,6%). Essas patentes descrevem sistemas que permitem o controle remoto de dispositivos domésticos por meio de redes interconectadas. Apesar do grande potencial, a complexidade de instalação, frequentemente associada à necessidade de configuração de múltiplos dispositivos e protocolos de comunicação, ainda é um obstáculo para a adoção em larga escala. No entanto, as patentes analisadas demonstram um crescente interesse em soluções inovadoras, como interfaces intuitivas, integração com assistentes virtuais e aplicativos móveis, visando simplificar a instalação e expandir cada vez mais as aplicações para áreas como o cuidado com idosos e a acessibilidade para pessoas com necessidades especiais.

A combinação de fatores demográficos, sociais e econômicos indica uma tendência de crescimento no desenvolvimento e proteção de tecnologias para automação residencial. Esse fenômeno, observado também em outros setores tecnológicos, tem intensificado a demanda por proteção da propriedade industrial, como evidenciado pelos 164.042 pedidos de patentes identificados neste estudo. Por outro lado, a crescente complexidade técnica desses pedidos exige dos examinadores uma alta capacidade analítica para avaliar a novidade, o passo inventivo e a aplicação industrial das invenções. Essa crescente complexidade, conforme apontado por Golodne e Silva (2022), tem gerado uma demanda cada vez maior, associada ao aumento do número de páginas e à necessidade de cumprir prazos cada vez mais curtos.

A análise lexical dos documentos de patentes utilizados neste estudo foi realizada em três fases distintas: treinamento, implementação com patentes reais e implementação com patentes fictícias (Tabela 1). Na fase de treinamento, foram analisados 150 textos, totalizando 34.768 segmentos e 43.714 palavras ativas (123.634 ocorrências nos textos). A densidade lexical, um indicador da diversidade vocabular de um texto, foi de 0,56. Foram identificadas 4 classes semânticas distintas, sendo a classe "Controle" a mais representativa (50,89% dos segmentos; $\chi^2_{máx} = 147,24$; $p < 0,001$). Na fase de implementação com patentes reais, a análise de 50 textos resultou em 5.189 segmentos e 8.033 palavras ativas (18.587

ocorrências), com uma densidade lexical de 0,59. A classe semântica predominante foi "Interface" (52,2%; $\chi^2_{máx} = 165,29$; $p < 0,001$). Por fim, na fase de implementação com patentes fictícias, a análise de 50 textos apresentou 2.503 segmentos, 1.412 palavras ativas (9.125 ocorrências) e uma densidade lexical de 0,27. Apenas 3 classes semânticas foram identificadas, sendo "Segurança" a mais representativa (56,7%; $\chi^2_{máx} = 92,46$; $p < 0,05$). Esses achados indicam diferenças significativas na extensão, complexidade linguística e no vocabulário utilizado nos textos. A redução drástica da densidade lexical das patentes fictícias pode ser atribuída à natureza simulada, que tendem a ter um vocabulário mais simples, repetitivo e com menos termos técnicos.

Tabela 1 – Estrutura metodológica para uso do GPT-4 no exame de pedidos de patentes

Fase do experimento	Treinamento	Implementação (Patentes reais)	Implementação (Patentes fictícias)
Nº de textos	150	50	50
Segmentos de texto	34.768 (68,7)	5.189 (45,02)	2.503 (53,08)
Nº de palavras	77.634	13.587	5.259
Nº palavras ativas	43.714	8033	1412
Nº de ocorrências	123.634	18.587	9.125
Densidade lexical	0,56	0,59	0,27
Classes semânticas	6	4	3

Fonte: Elaborada pelos autores (2024).

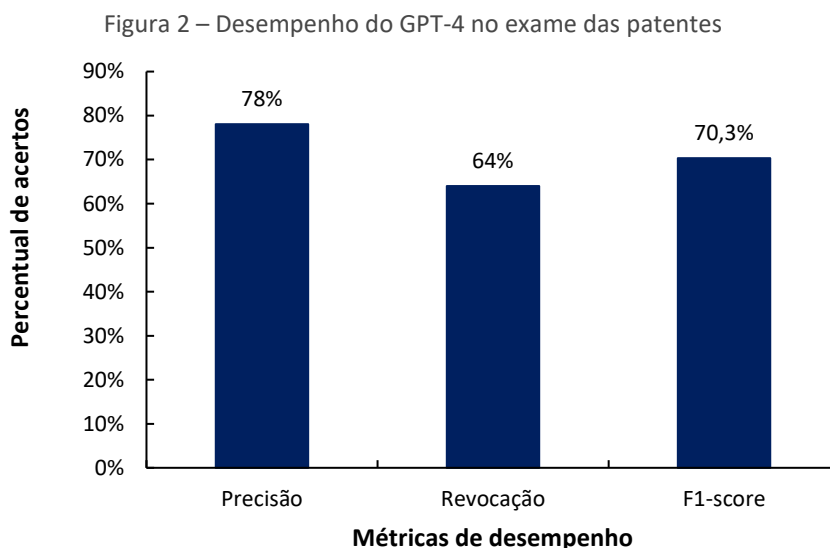
Após a confirmação da heterogeneidade lexical do corpus de patentes e a implementação do GPT-4, procedeu-se à avaliação de seu desempenho na classificação das 100 patentes. A matriz de confusão indica que o modelo identificou corretamente 32 patentes reais como patenteáveis, demonstrando uma boa capacidade de reconhecer características distintivas de invenções passíveis de proteção legal. No entanto, classificou incorretamente 18 patentes reais como não patenteáveis. O modelo identificou corretamente 41 patentes fictícias como não patenteáveis, evidenciando uma boa capacidade de rejeitar invenções que não se enquadram nos requisitos legais. Por outro lado, classificou incorretamente 9 patentes fictícias como patenteáveis. Essa taxa de falsos positivos indica que o GPT-4 pode ser suscetível à superestimação, classificando algumas invenções como patenteáveis mesmo quando não atendem aos critérios de patenteabilidade (Tabela 2).

Tabela 2 – Matriz de confusão da previsão do GPT-4 no exame das patentes (N = 100)

Real	Previsto pelo GPT-4	
	Patenteável	Não patenteável
Patenteável	32	9
Não patenteável	18	41

Fonte: Elaborada pelos autores (2024).

A Figura 2 apresenta a comparação das métricas de desempenho do GPT-4 na análise das patentes. O modelo demonstrou um desempenho satisfatório na classificação de patentes como patenteáveis, com uma precisão de 78% e uma revocação de 64%. O *F1-score*, que busca equilibrar essas duas métricas, apresentou um valor de 70,3%.



Fonte: Elaborada pelos autores (2024).

Esses resultados indicam um viés do GPT-4 em favor da precisão, reforçando a ideia de que o modelo é mais conservador em suas classificações positivas. Essa característica pode ser desejável em cenários onde o custo de classificar erroneamente uma patente não patenteável como patenteável (falso positivo) é alto, pois evita sobrecarregar os examinadores com um grande número de pedidos que podem ser posteriormente invalidados (nulidade da patente). No entanto, é fundamental considerar o impacto de classificar erroneamente uma patente patenteável como não patenteável (falso negativo), pois isso pode resultar em prejuízos financeiros significativos e na perda de potenciais inovações no mercado.

CONCLUSÕES

A superação dos desafios relacionados ao processo de concessão de patentes no Brasil e no mundo exige um esforço conjunto de diversos atores, incluindo o governo, a iniciativa privada e a academia. A implementação de medidas para agilizar e modernizar esse processo é fundamental para estimular a inovação, fortalecer o setor produtivo e promover o desenvolvimento econômico dos países. Neste estudo, exploramos um conjunto de dados patentários associados à automação doméstica para demonstrar a eficácia da aplicação do GPT-4 nesse processo. O modelo apresentou um desempenho razoável na classificação das patentes, com uma precisão acima de 70%. No entanto, a revocabilidade mais baixa que a precisão indica que o GPT-4 ainda requer aprimoramentos para ser utilizado de forma autônoma.

Os resultados deste estudo evidenciam a necessidade de pesquisas adicionais para explorar o potencial do GPT-4 na área de propriedade intelectual. Apesar de

ser um modelo de linguagem avançado, as limitações na classificação de algumas patentes podem ser atribuídas, em parte, às complexidades inerentes à compreensão de conceitos técnicos e à realização de raciocínios jurídicos. A ausência de um conhecimento especializado e aprofundado sobre a tecnologia em questão pode ter levado a classificações errôneas. O domínio da automação doméstica, caracterizado por sua constante inovação tecnológica, em conjunto com a complexidade e subjetividade da legislação de patentes brasileira, pode ter contribuído para a dificuldade na tarefa de exame de pedidos de patentes internacionais (traduzidas do inglês). Um ajuste fino nos hiperparâmetros do modelo e a coleta de dados de treinamento adicionais podem ajudar a melhorar seu desempenho ao longo do tempo.

De maneira geral, este estudo pioneiro forneceu subsídios para o desenvolvimento de ferramentas inovadoras que auxiliem na modernização do sistema de propriedade industrial e na promoção da inovação no país.

Using generative pre-trained transformer to streamline patent application examining

ABSTRACT

Patent evaluation ensures the legitimacy of inventions, but examiners face the challenge of balancing speed with accurate analysis. This study explores the potential of generative pre-trained transformers (GPTs) to automate steps in the patent examination process. GPT-4 was trained on a corpus of patents, learning to identify complex patterns and extract relevant information. After refinement with specific data on patentability requirements, the model was evaluated on new patent application classification tasks. The results indicated high accuracy for GPT-4, although there is a clear need to reduce the false negative rate. Optimizing the hyperparameters and expanding the training dataset is believed to improve the model's ability to capture semantic and conceptual nuances, making it an even more effective tool for accurately classifying patent applications.

KEYWORDS: Machine learning. Artificial intelligence. Patents. Innovation.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

REFERÊNCIAS

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, e101509, 2022.

DEEB, M. *et al.* **“The emerging role of generative artificial intelligence in transplant medicine”**. American Journal of Transplantation, In Press, 2024.

GARGARI, O. K. *et al.* **“Enhancing title and abstract screening for systematic reviews with GPT-3.5 turbo”**. BMJ Evidence-Based Medicine, v. 19, n. 1, p. 69-70, 2024.

GOLODNE, D. M.; SILVA, A. S. **“Managing patent examiners’ workload by using patent application volume data”**. World Patent Information, v. 68, e102095, 2022.

GUO, E. *et al.* **“Automated Paper Screening for Clinical Reviews Using Large Language Models: Data Analysis Study”**. Journal of Medical Internet Research, v. 26, n. 1, e48996, 2024.

INPI. **“Patentes”**, 2020. Disponível em: <https://www.gov.br/inpi/pt-br/>. Acesso em: 15 jul. 2024.

KULKARNI, A.; CHONG, D.; BATARSEH, F. A. **“Foundations of data imbalance and solutions for a data democracy”**. Data Democracy, p. 83–106, 2020.

LANDSCHAFT, A. *et al.* **“Implementation and evaluation of an additional GPT-4-based reviewer in PRISMA-based medical systematic literature reviews”**. International Journal of Medical Informatics, v. 189, n. 1, e 105531, 2024.

MDIC. **“Usuário pode reduzir em até 3 anos prazo para concessão de patente no Brasil”**. Disponível em: <https://www.gov.br/mdic/pt-br/assuntos/noticias/2024/marco/usuario-pode-reduzir-em-ate-3-anos-prazo-para-concessao-de-patente-no-brasil>. Acesso em: 23 jul. 2024.

NGUYEN-MAU, T.; LE, A. C.; PHAM, D. H; HUYNH, V. N. **“An information fusion-based approach to context-based fine-tuning of GPT models”**. Information Fusion, v. 104, e102202, 2024.

YANG, H.; ZHANG, Y.; LIU, L. **“Patent priority examination under mission orientation: can speed and quality coexist?”** Technology in Society, v. 76, e102452, 2024.

. José Wendel dos Santos, José Barreto., Leonôra Virginia de Jesus Dias, Cristiane Monteiro de Farias Rezende, Mário Jorge Campos dos Santos

Recebido: 2024/09/17

Aprovado: . 2024/09/17

DOI: 103895/recit. V15 n37-19116

Como citar: SANTOS, J. W.; NETTO, J. B NETTO, SANTOS, I. N; DIAS, L. D. V. J.; REZENDE, C. M. F., SANTOS, M. J.C... Uso de transformadores generativos pré-treinados para agilizar o exame de pedidos de patente R. Eletr. Cient. Inov. Tecnol, Medianeira, v. 15. n. 37, p. 30 - 40, mai/ago, 2024. Disponível em: <<https://periodicos.utfpr.edu.br/recit>>. Acesso em: XXX.

Correspondência:

Leonôra Virginia de Jesus Dias Correio
Universidade Federal de Sergipe, São Cristóvão, Sergipe.

Direito autoral: Este artigo está licenciado sob os termos da Licença creativecommons.org/licenses/by-nc/4.0 Internacional.

