RMP
Revista dos Mestrados Profissionais
UFPE / CCSA – MGP
https://periodicos.ufpe.br/revistas/rmp
ISSN – 2317 - 0115
V. xx - N. x (xxxx)
Recife – PE

Análise Preditiva de Pendências em Chamados de Suporte Técnico no Judiciário Brasileiro: Um Estudo de Caso

Lenylda Maria de Souza Albuquerque
Universidade de Pernambuco
lmsa@ecomp.poli.br
https://orcid.org/0009-0007-7119-8455

Jonas da Silva Aquino
Universidade de Pernambuco
jsa2@ecomp.poli.br
https://orcid.org/0009-0006-2363-4198

Francisco Marcelo de Barros Maciel
Universidade de Pernambuco
fmbm@ecomp.poli.br
https://orcid.org/0009-0006-9884-6961

Paulino Calei Alves Domingos
Universidade de Pernambuco
pcad@ecomp.poli.br
https://orcid.org/0009-0003-8889-4343

Tiago de Sousa Ribeiro
Universidade Federal de Pernambuco
Tiago.ribeiro@ufpe.br
https://orcid.org/0000-0002-7317-893X

Resumo – A gestão de chamados de suporte é crítica para a continuidade, especialmente no Poder Judiciário. Este artigo propõe uma abordagem preditiva para identificar e modelar padrões de pendências no atendimento. Foram analisados 13.102 registros categóricos do sistema Assyst de um órgão do Judiciário brasileiro. Após pré-processamento, aplicou-se K-Modes para agrupar solicitações com perfis similares. Emergiram cinco clusters, associados a causas como ausência de informações essenciais, falta de retorno do usuário e dependências técnicas. A consistência foi validada por Análise de Correspondência Múltipla. Os achados subsidiam ações proativas, melhor alocação de recursos e redução de interrupções em serviços de TI no setor público.

Palavras-chave: Mineração de Dados; Suporte Técnico; Pendência de Atendimento; K-Modes; Análise Preditiva

Predictive Analysis of Pending Issues in Technical Support Tickets in the Brazilian Judiciary: A Case Study

Abstract – *Managing technical support tickets is critical to service continuity, especially in high-criticality environments such as the Judiciary. This paper proposes a predictive approach to identify and model patterns of pending issues in ticket handling. We analyzed 13.102 categorical records from the Assyst system of a Brazilian judicial body. After data preprocessing, the K-Modes algorithm was applied to cluster requests with similar pending*

profiles. Five distinct clusters emerged, associated with causes such as missing essential information, lack of user feedback, and technical dependencies. Cluster consistency was validated through Multiple Correspondence Analysis. The findings support proactive actions, improved resource allocation, and reduced disruptions in public-sector IT services.

Keywords: *Data Mining; Technical Support; Pending Issues; K-Modes; Predictive Analytics*

Data da Submissão: Deixar em branco - **Data de aceitação:** Deixar em branco

1. Introdução

A eficiência na prestação de serviços e na entrega de produtos é um pilar fundamental para a sustentabilidade e reputação de qualquer organização contemporânea. Nesse cenário, o suporte técnico desempenha um papel estratégico ao atuar como o principal elo entre a instituição e seus usuários, sendo crucial para a manutenção da qualidade do serviço e a satisfação do cliente (Fantinatti, 2008), (Cavalcanti, 2012). Contudo, mesmo em estruturas que adotam as melhores práticas de gestão da qualidade, falhas operacionais e interrupções no fluxo de atendimento são inevitáveis, gerando gargalos que podem comprometer a eficiência e acarretar custos significativos (Amorim, 2020).

Particularmente em instituições do setor público, como o Poder Judiciário brasileiro, a operacionalidade dos sistemas de Tecnologia da Informação (TI) é vital para a continuidade de suas atividades. Os chamados de suporte técnico, nesse contexto, não apenas solucionam problemas pontuais de hardware, *software* e sistemas, mas garantem a fluidez dos processos judiciais e administrativos, impactando diretamente a efetividade dos serviços prestados à sociedade.

Um dos principais desafios enfrentados por essas equipes de assistência técnica reside na gestão do tempo em que os chamados permanecem pendentes de atendimento. Essas interrupções podem ser causadas por múltiplos fatores, como a ausência de informações essenciais por parte do solicitante, a dificuldade de contato ou a necessidade de investigações técnicas complexas. A reiteração dessas pendências não só aumenta os custos operacionais diretos e indiretos, como a alocação prolongada de mão de obra e o consumo de recursos, mas também prejudica a imagem da instituição e a percepção de qualidade do serviço de suporte (Amorim, 2020).

Os chamados de assistência técnica do órgão em estudo correspondem a solicitações relacionadas à infraestrutura de tecnologia da informação, abrangendo desde problemas com hardware e *software* até pedidos de instalação e sustentação de sistemas. O fluxo de atendimento ocorre em três níveis: o Primeiro Nível, executado por funcionários terceirizados responsáveis por solucionar demandas simples ou encaminhá-las aos níveis superiores; o Segundo Nível, composto por servidores da instituição com especialização em sistemas administrativos e judiciais; e o Terceiro Nível, voltado ao tratamento de problemas que

exigem alterações no código-fonte dos sistemas. Após o encerramento do chamado, é enviada uma pesquisa de satisfação ao usuário.

Caso a solução seja considerada insatisfatória, o chamado é reaberto. Durante esse processo, ele pode permanecer em estado de "pendente de atendimento" por diferentes razões, como: ausência de resposta do solicitante a uma requisição de informação adicional, falha no contato com o usuário ou a necessidade de investigação técnica mais aprofundada por parte da equipe, entre outras. Essas pendências podem ocorrer repetidamente em um mesmo chamado, o que acarreta custos adicionais, compromete a eficiência operacional da equipe e prejudica a imagem do serviço de suporte.

Diante desse cenário, a mineração de dados (MD) emerge como uma abordagem promissora para desvendar padrões ocultos e extrair conhecimento acionável de vastos volumes de dados históricos (Rezende, 2005). Nesse sentido, a questão central que orienta esta pesquisa é: quais padrões podem ser extraídos dos dados históricos de chamados de suporte técnico para identificar os fatores que contribuem para interrupções no atendimento?

Este estudo tem como objetivo geral investigar a incidência e as causas que influenciam as pendências no atendimento de chamados de assistência técnica, utilizando técnicas de mineração de dados. Para tanto, os objetivos específicos são: (*a*) estudar as características dos chamados abertos no período de 01 de janeiro a 31 de dezembro de 2024, com foco nas variáveis que influenciam as pendências; (b) investigar e classificar as causas dessas pendências em categorias para uma análise detalhada; (c) modelar as principais variáveis que contribuem para a ocorrência de pendências nos chamados de suporte técnico; e (d) avaliar a frequência e os padrões de interrupção no atendimento, bem como suas implicações operacionais para o desempenho do suporte técnico.

A identificação precoce desses fatores permite uma gestão mais eficiente dos recursos humanos e financeiros. Cada um dos 13.102 registros analisados correspondeu a uma pendência que gerou custos diretos (alocação de mão de obra) e indiretos (eletricidade, telecomunicações e custo de oportunidade), além de impactar negativamente a percepção de qualidade dos serviços prestados. A aplicação de técnicas preditivas favorece decisões baseadas em evidências, promovendo alocação estratégica de recursos e redução de custos operacionais no âmbito do Poder Judiciário.

Diferentemente de abordagens existentes que focam predominantemente na abertura ou no desfecho dos chamados, a exemplo de estudos que otimizam custos ou analisam a reincidência de novos atendimentos (Mayer et al., 2014; Gomes et al., 2014) ou que preveem o não comparecimento em agendamentos (Adeodato et al., 2008), esta pesquisa se aprofunda na análise das fases intermediárias do atendimento e nas causas das pendências, proporcionando uma compreensão mais granular dos gargalos operacionais. Ao oferecer *insights* baseados em evidências diretamente de um sistema de gestão técnica em uso (Assyst) e validar os agrupamentos por meio de técnicas como a Análise de Correspondência Múltipla (ACM), este trabalho contribui para a tomada de decisões estratégicas, a otimização da alocação de recursos e a implementação de planos de ação direcionados, elevando a eficiência e a qualidade dos serviços de suporte técnico no Poder Judiciário.

Para a exposição do trabalho desenvolvido, este artigo foi estruturado em oito seções. Além desta introdução, o texto organiza-se da seguinte forma: a Seção 2 apresenta os conceitos fundamentais sobre chamados de assistência técnica, mineração de dados e algoritmos de agrupamento; a Seção 3 discute os trabalhos relacionados e o diferencial desta pesquisa; a Seção 4 detalha a metodologia adotada, incluindo a descrição da base de dados,

o pré-processamento e a metodologia experimental; a Seção 5 expõe a análise e a discussão dos resultados; a Seção 6 aborda as limitações do estudo; a Seção 7 apresenta sugestões para trabalhos futuros; e a Seção 8 traz as considerações finais.

2. Conceitos

Os Chamados de Assistência Técnica são solicitações formais feitas por usuários para relatar problemas ou necessidades relacionadas a produtos ou sistemas tecnológicos. Esses chamados são essenciais para garantir que as demandas sejam registradas e tratadas de forma eficiente, proporcionando uma resposta adequada e oportuna às necessidades dos usuários.

No contexto do Poder Judiciário, esses chamados são cruciais para manter a operacionalidade dos sistemas de tecnologia da informação, garantindo que os servidores e usuários possam realizar suas atividades sem interrupções significativas. A eficiência na gestão desses chamados impacta diretamente a qualidade do serviço prestado e a imagem da instituição.

2.1 Mineração de Dados

Por ser considerada multidisciplinar, a MD se utiliza da Estatística, do Aprendizado de Máquina e de Bancos de Dados. Sob a perspectiva da Estatística, a MD é a análise de grandes volumes de dados para identificar padrões ocultos e resumir informações de forma útil e compreensível (Han; Kamber, 2012).

Já na perspectiva de Aprendizado de Máquina, a MD é uma etapa crucial do processo KDD (*Knowledge Discovery in Databases*), ou Descoberta de Conhecimento em Bases de Dados, onde algoritmos são aplicados para construir modelos que aprendem a partir dos dados, permitindo a identificação de padrões e a realização de previsões sob restrições computacionais (Fayyad et al., 1999). O KDD é a área que investiga a extração de conhecimento a partir de grandes volumes de dados experimentais ou observacionais. Esse processo envolve várias etapas, desde a seleção e preparação dos dados até a interpretação dos resultados, sendo a mineração de dados uma delas, responsável por identificar padrões ocultos e relevantes nas informações (Fayyad et al., 1999).

Por fim, na perspectiva de Bancos de Dados, MD combina técnicas de Inteligência Artificial, Estatística e Banco de Dados para extrair informações valiosas de grandes conjuntos de dados (Han; Kamber, 2012).

Nesse cenário, a Mineração de Dados apresenta um potencial transformador para otimizar a eficiência dos processos de assistência técnica, especialmente no ambiente de dados descritivos e categóricos presente nos sistemas de gestão de chamados.

2.2 Algoritmos de Agrupamento

O agrupamento, ou *clustering*, é uma tarefa fundamental da MD que visa particionar um conjunto de objetos em grupos, ou clusters, de forma que objetos no mesmo grupo sejam mais similares entre si do que com aqueles em outros grupos (Han; Kamber, 2012). A escolha do algoritmo de agrupamento é crucial e depende da natureza dos dados.

O algoritmo K-Means é amplamente utilizado para agrupar dados numéricos, baseando sua lógica na minimização da soma dos quadrados das distâncias (geralmente euclidianas) entre cada ponto e o centroide do cluster ao qual pertence (Jain et al., 1999). No

entanto, sua aplicação é inadequada para dados categóricos, pois a noção de média e distância euclidiana não se aplica a variáveis qualitativas.

Para conjuntos de dados predominantemente categóricos, como os registros de chamados de suporte técnico em questão, o algoritmo K-Modes é a escolha metodológica mais apropriada (De Vos, 2015). Desenvolvido por Huang (1998), o K-Modes Clustering é uma extensão do K-Means que supera a restrição de dados numéricos. Em vez de médias, o K-Modes define os centróides dos clusters com base na moda (o valor mais frequente) das variáveis categóricas dentro de cada grupo. A dissimilaridade entre os pontos de dados é medida pela distância de Hamming, que contabiliza o número de atributos em que dois objetos diferem. Essa abordagem baseada em correspondência categórica permite a formação de clusters que refletem a similaridade de perfis, sendo altamente coerente com a análise de dados textuais e descritivos presentes em chamados (Ahmad; Khan, 2013).

Dessa forma, o algoritmo K-Modes Clustering, possibilita a classificação de dados categóricos por meio de combinações simples de medidas de similaridade entre categorias. Esse método mantém a estrutura do K-Means, porém remove a restrição de que os dados precisam ser numéricos, tornando-o adequado para conjuntos de dados compostos por variáveis qualitativas.

3. Trabalhos Relacionados

O posicionamento de uma pesquisa no cenário científico exige uma revisão criteriosa dos trabalhos já desenvolvidos na área. Essa etapa é fundamental, sobretudo, para identificar as lacunas que a presente investigação busca preencher, consolidando o embasamento teórico e evidenciando a originalidade e a contribuição do estudo em questão (Webster; Watson, 2002). Com o propósito de fundamentar teoricamente a proposta e situá-la, foram analisados trabalhos correlatos que aplicam técnicas de Mineração de Dados em contextos de atendimento ao público.

Uma das abordagens encontradas na literatura é a de Mayer *et al*. (2014), que se dedicaram à redução de custos nos canais de atendimento de uma empresa de energia (Ampla Energia). Este estudo empregou métodos estatísticos e técnicas de Mineração de Dados sobre um volume substancial de 5 milhões de registros de atendimento, abrangendo 1.7 milhões de clientes ao longo de 12 meses. A principal contribuição desses autores foi a identificação de padrões e tendências que possibilitaram a otimização da gestão de custos nos diversos pontos de contato com o cliente. Embora valioso para a otimização macroeconômica de serviços, esse trabalho não se aprofundou na granularidade das causas de interrupção ou pendências em chamados de suporte técnico em andamento, focando mais na gestão estratégica de custos gerais de atendimento.

Em outra linha de pesquisa, Gomes *et al*. (2014) investigaram a identificação de atendimentos sucessivos a um mesmo usuário no Serviço de Atendimento Móvel de Urgência (SAMU) de um município na Região Metropolitana de Curitiba, PR. Utilizando regras de associação e análise de janela de tempo sobre uma base de 5.839 eventos, o estudo conseguiu descobrir padrões de reincidência nos atendimentos, o que contribuiu para uma melhor compreensão da demanda do serviço. Este trabalho demonstra a eficácia da mineração de dados em revelar comportamentos repetitivos dos usuários. Contudo, seu foco recai sobre a reincidência de novos atendimentos e não sobre as causas intrínsecas ou a dinâmica das pendências que ocorrem dentro de um único processo de suporte.

Adicionalmente, o trabalho de Adeodato *et al*. (2008) focou na previsão do não comparecimento (no-show) em atendimentos agendados por *call-center* em uma empresa de assistência médica. Para isso, aplicaram técnicas de Mineração de Dados com uma rede neural Multi-Layer Perceptron (MLP), treinada com 30.000 consultas e testada com 10.000. A pesquisa culminou no desenvolvimento de um sistema de apoio à decisão capaz de realizar o reagendamento em tempo real de consultas de alto risco. Este é um exemplo pertinente da aplicação preditiva da mineração de dados para otimizar a alocação de recursos. No entanto, o problema central difere do presente estudo, pois aborda a previsão de um comportamento do usuário (a não adesão ao agendamento) em vez de analisar as causas operacionais das interrupções no fluxo de um chamado já iniciado.

O estudo desenvolvido por McConnell *et al*. (2021) aborda a aplicação de aprendizado de máquina na previsão do resultado de moções em processos civis norte-americanos. O objetivo central consistiu em avaliar a eficácia de modelos como XGBoost e Naive Bayes na antecipação de decisões judiciais, destacando também o papel da explicabilidade dos algoritmos. Para isso, recorreram a dados de casos civis com engenharia de recursos baseada em texto jurídico (TF-IDF). Os autores salientam que, embora as previsões tenham sido promissoras, ainda permanecem limitações quanto à transparência e contexto dos casos, sendo recomendada a evolução de métodos interpretáveis e a ampliação dos estudos empíricos para uma integração ética da IA no direito, com ênfase nas pesquisas sobre viés e comparação com a previsão humana. Este trabalho, embora utilize técnicas de IA no contexto jurídico, concentra-se na predição de desfechos processuais a partir de informações textuais legais, diferindo substancialmente do objetivo de analisar a eficiência operacional de chamados de suporte técnico em TI. A granularidade da análise está no conteúdo jurídico e em decisões, não nos processos internos de gestão de infraestrutura.

Na investigação realizada por Bianchini *et al*. (2023), são discutidos os desafios da análise automatizada de processos judiciais italianos após a digitalização. O trabalho teve como meta compreender de que forma técnicas modernas de IA, como mineração de processos e redes LSTM+CNN, podem ser aplicadas ao diagnóstico e automatização das tarefas judiciais. A metodologia incluiu a análise de registros digitais e a reengenharia de processos mediante pipelines de dados. Os resultados evidenciam dificuldades de adaptação dos tribunais à complexidade informacional, mas também o potencial da IA (Inteligência Artificial) para aumentar automação e transparência. Entre as recomendações, destacam-se a integração de sistemas, a formação de operadores e o desenvolvimento de metodologias adequadas à realidade do judiciário italiano. Similarmente ao trabalho anterior, Bianchini *et al*. (2023), focam na automação e análise de fluxos de trabalho judiciais, abordando desafios da digitalização e o potencial da IA para reengenharia de processos judiciais. Contudo, seu escopo ainda difere da investigação detalhada das causas de pendências em chamados de suporte técnico de TI, que é o foco da presente pesquisa.

A partir da revisão desses trabalhos, observa-se uma tendência consolidada na aplicação da Mineração de Dados e Inteligência Artificial para a otimização de processos de atendimento ao público e, mais recentemente, no âmbito jurídico. Os estudos de Mayer *et al*. (2014), Gomes *et al*. (2014) e Adeodato *et al*. (2008) exemplificam abordagens valiosas, variando desde a otimização de custos e a identificação de reincidências de usuários até a previsão de não comparecimento em serviços agendados, empregando metodologias diversas. Por sua vez, McConnell *et al*. (2021) e Bianchini *et al*. (2023) demonstram a aplicação de IA em cenários jurídicos, focando em predição de resultados e automação de processos, respectivamente.

Contudo, um aspecto comum aos primeiros três estudos é o enfoque em pontos de contato específicos do ciclo de atendimento (custos gerais, abertura de novos chamados, agendamentos) ou no comportamento externo do usuário (reincidência, no-show), sem aprofundar na dinâmica interna das interrupções durante o processo de atendimento de um único chamado. Os trabalhos mais recentes no contexto judiciário, embora pertinentes à aplicação da IA neste domínio, não investigam as especificidades operacionais dos chamados de suporte técnico e suas pendências, mas sim aspectos de predição de decisões ou automação de processos jurídicos.

Nesse contexto, embora os trabalhos supracitados sejam valiosos para o campo da otimização de serviços e da aplicação da Mineração de Dados e IA, a presente pesquisa distingue-se significativamente por seu foco e granularidade, especialmente ao considerar a totalidade dos estudos revisados.

Em contraste com as abordagens focadas em otimização de custos (Mayer et al., 2014), análise de reincidência em novos atendimentos (Gomes et al., 2014) ou previsão de não comparecimento em agendamentos (Adeodato et al., 2008), este estudo concentra-se especificamente na identificação, categorização e modelagem das causas intrínsecas das pendências em chamados de assistência técnica em andamento.

A investigação aprofundada das variáveis operacionais e descritivas diretamente extraídas de um sistema de gestão técnica (Assyst) permite uma compreensão detalhada dos fatores que comprometem a resolução eficiente dos chamados, atuando nas fases intermediárias do atendimento, e não apenas na sua abertura ou desfecho.

Adicionalmente, em comparação com os estudos de McConnell *et al*. (2021) e Bianchini *et al*. (2023), que também utilizam IA no Poder Judiciário, o presente trabalho se diferencia por sua aplicação específica na gestão de TI. Enquanto McConnell *et al*. (2021), focam na predição de resultados de moções legais e Bianchini *et al*. abordam a automação e reengenharia de processos judiciais de forma mais ampla, esta pesquisa se aprofunda na micrologística e eficiência do suporte técnico, um pilar fundamental para a operacionalidade de qualquer instituição, incluindo o Judiciário. O foco aqui é otimizar a infraestrutura de TI e a prestação de serviços de suporte, elementos cruciais que impactam indiretamente, mas de forma vital, a capacidade do Judiciário de funcionar eficazmente.

Essa perspectiva oferece um diferencial metodológico ao explorar a dinâmica interna do suporte técnico no setor público, gerando insights práticos para a melhoria contínua dos serviços e a elaboração de estratégias proativas de mitigação de gargalos operacionais.

A utilização do algoritmo K-Modes, particularmente adequado para dados categóricos como os encontrados em registros de chamados de suporte técnico, e a validação por Análise de Correspondência Múltipla, reforçam a robustez da metodologia na extração de padrões significativos a partir de informações descritivas. Assim, a pesquisa se posiciona como uma contribuição original e estratégica no campo da mineração de dados aplicada à gestão de chamados técnicos, com alta aplicabilidade para o Poder Judiciário brasileiro, preenchendo uma lacuna na literatura ao focar na identificação de padrões de pendência que impactam diretamente a satisfação do usuário e a eficiência operacional em um ambiente de alta criticidade.

4. Metodologia

Esta seção detalha os procedimentos metodológicos adotados para a consecução dos objetivos desta pesquisa. Apresenta-se a análise dos *stakeholders* envolvidos, a descrição da base de dados utilizada, as etapas de pré-processamento dos dados e a metodologia experimental aplicada para a análise preditiva das pendências em chamados de suporte técnico.

4.1 *Stakeholders* Envolvidos

A identificação e a análise dos stakeholders envolvidos em um projeto de pesquisa aplicada são essenciais para o seu planejamento, execução e para o sucesso na validação e implementação dos resultados (Freeman, 1984). Esta etapa permite mapear os interesses, o grau de influência e o nível de envolvimento de cada parte, fundamentando a gestão estratégica da comunicação e das expectativas ao longo do desenvolvimento da pesquisa.

Os "participantes" primários da análise são os 13.102 registros de chamados de suporte técnico. Os *stakeholders* (Analista Judiciário de TI e usuários finais) atuaram como validadores e fontes de contextualização do domínio, e não como participantes na coleta de dados primários do fenômeno estudado.

O Poder Judiciário Federal (TRT) atua como o stakeholder patrocinador institucional deste projeto. Sua importância reside na provisão do contexto real de aplicação, da infraestrutura e dos dados necessários para a investigação.

O analista judiciário do setor de Tecnologia da Informação do órgão patrocinador foi identificado como um *stakeholder* central, possuindo alto interesse e alta disponibilidade. Seu papel é crucial, sendo responsável direto pelas solicitações de requisitos, pela validação dos entregáveis da pesquisa e, sobretudo, pela tomada de decisões estratégicas que orientam o estudo. Sua atuação como facilitador e decisor assegura a relevância e a aderência dos resultados às necessidades operacionais da instituição.

A própria instituição (TRT) e seus usuários finais (servidores do setor) constituem outro grupo de *stakeholders* igualmente relevante. Embora apresentem um interesse médio e, em geral, baixa disponibilidade devido às suas atividades laborais primárias, seu envolvimento é de natureza informativa, sendo os principais fornecedores de *feedback* sobre a aplicabilidade e o impacto prático das soluções propostas. Sua perspectiva é fundamental para garantir que as melhorias desenvolvidas sejam de fato úteis e bem recebidas.

No âmbito acadêmico, o grupo de autores desse estudo representa o principal executor da pesquisa. Caracterizados por alto interesse e alta disponibilidade, são os responsáveis diretos pela condução da investigação, coleta e análise de dados, e pela elaboração dos resultados. Além de serem responsáveis pela metodologia e validação científica dos procedimentos, assegurando o rigor e a qualidade acadêmica do trabalho.

A análise dessas inter-relações entre os stakeholders permitiu um planejamento estratégico do projeto, visando a maximização do engajamento das partes interessadas e a garantia de que os resultados da pesquisa sejam alinhados às expectativas e passíveis de aplicação prática no ambiente do Poder Judiciário.

4.2 Interação e Validação dos Achados com os *Stakeholders*-Chave do TRT

A interação e comunicação com os *stakeholders* do TRT foram planejadas e executadas de forma estratégica para assegurar a relevância do estudo e a validação prática

dos resultados.

O Analista Judiciário do setor de Tecnologia da Informação, identificado como o *stakeholder* central para decisões e validações, desempenhou um papel crucial em todas as fases da pesquisa. Desde o início, houve uma colaboração estreita para a solicitação de requisitos e o consenso na seleção de atributos durante a etapa de redução de dados, garantindo que as variáveis analisadas fossem significativas para a realidade operacional da instituição. Este analista foi o principal ponto de contato para a apresentação e discussão dos resultados preliminares e finais.

A validação dos achados com este stakeholder se deu por meio de sessões de feedback e discussões aprofundadas por meio de reuniões semanais, apresentações de resultados parciais e finais, além de registros em atas de reunião contemplando propostas de melhoria, resultados e próximos passos (quando cabíveis). Os cinco perfis de clusters identificados pelo algoritmo K-Modes, as principais causas de pendência e seus respectivos tempos médios de resolução foram apresentados detalhadamente.

A validação por especialistas de domínio (*Face Validity*) foi fundamental nesse processo em que o analista confirmou que os agrupamentos e as interpretações das causas de pendência correspondiam a situações e desafios já percebidos no dia a dia do suporte técnico. Este reconhecimento por parte de um especialista interno garantiu que os resultados não eram meros artefatos estatísticos, mas sim representações válidas e acionáveis da realidade operacional do TRT.

Adicionalmente, os usuários finais (servidores do setor), embora com menor disponibilidade, foram engajados de forma a fornecer *feedback* sobre a percepção e o impacto das pendências em seu trabalho. Essa entrada foi essencial para corroborar as implicações práticas dos resultados e para entender as nuances da satisfação do usuário, especialmente em *clusters* como o 5 (problemas de comunicação e mensageria) onde a insatisfação era elevada. Suas perspectivas ajudaram a contextualizar a gravidade de certas pendências e a validar a relevância das propostas de melhoria.

Em suma, a comunicação foi estabelecida em múltiplos níveis: com a gerência de TI para validação técnica e estratégica, e com os usuários finais para validação da experiência. Essa abordagem colaborativa e a validação contínua com os stakeholders-chave do TRT foram essenciais para garantir que os resultados do estudo não apenas tivessem rigor científico, mas também alta aplicabilidade e aceitação no ambiente real do Poder Judiciário, pavimentando o caminho para a implementação de melhorias baseadas em evidências.

4.3 Descrição da Base de Dados

A base de dados para este estudo foi obtida a partir do Sistema Assyst, uma plataforma de gestão de serviços de TI utilizada no órgão patrocinador da pesquisa. Como principal "instrumento de coleta" o Sistema Assyst forneceu os dados brutos estruturados dos chamados. Dessa forma, não foram utilizados questionários ou formulários criados especificamente para a coleta inicial de dados.

Este sistema forneceu um conjunto robusto de 13.102 registros de ocorrências em chamados de assistência técnica, abrangendo o período de 01 de janeiro de 2024 a 31 de dezembro de 2024. A seleção desse volume de dados visa garantir a representatividade e a abrangência temporal necessárias para a identificação de padrões sazonais e comportamentais.

O dicionário de dados final, compreendeu os seguintes campos principais, categorizados por tipo e tamanho, com valores permitidos específicos:

- TIPO_CHAMADO: Caractere (10), descrevendo o tipo de solicitação (e.g., Incidente, Mudança, Tarefa);
- UNIDADE_USUARIO_SOLICITANTE: Caractere (20), identificando a origem do solicitante (e.g., Usuário interno, Advogados, Aposentado);
- GRUPO_SERVICO: Caractere (50), categorizando o serviço (e.g., Serviços de rede, Sistemas Judiciais);
- SERVICO: Caractere (40), especificando o serviço dentro do grupo (e.g., Active Directory, PJE);
- CATEGORIA: Caractere (10), fornecendo uma categoria mais específica do chamado (e.g., Erro/Falha, Desenvolvimento);
- CAUSA_PENDENTE: Caractere (30), descrevendo o motivo da pendência na resolução do chamado;
- DT_ABERTURA: Data (8), indicando a data de abertura do chamado;
- TEMPO_RESOLUCAO: Numérico (10), registrando o tempo em minutos para a resolução do chamado;
- TEMPO_RESOLUCAO_HORAS: Numérico (10), registrando o tempo em horas para a resolução do chamado.

Esta riqueza de atributos permite uma análise multifacetada dos fatores que influenciam as pendências no atendimento.

4.4 Análise Descritiva dos Dados

A etapa de análise descritiva teve como objetivo primordial compreender as características gerais da base de dados, identificando padrões, distribuições estatísticas e possíveis inconsistências antes da aplicação das técnicas de mineração. Com base nos 13.102 chamados de assistência técnica do ano de 2024, observou-se que os chamados do tipo "Incidente" configuram os principais objetos de análise. Esta priorização deve-se à sua maior suscetibilidade a reaberturas, caso a resolução inicial não seja satisfatória, tornando-os críticos para a eficiência do suporte.

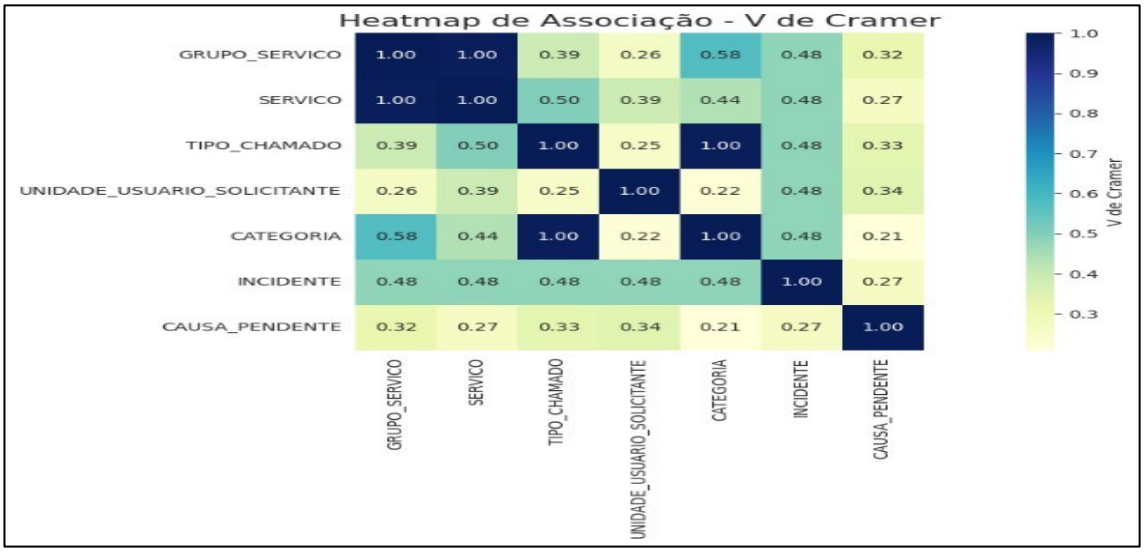
A investigação inicial também incluiu a exploração de associações entre as *features* (variáveis) do conjunto de dados. A Figura 1, apresentou uma relevante associação positiva entre diversas variáveis (e.g., “GRUPO_SERVICO” e “CATEGORIA” (0.58), e entre “TIPO_CHAMADO” e “CATEGORIA” (0.48)), sugerindo interdependência estrutural nas regras operacionais do serviço, o que justifica uma investigação mais aprofundada para entender as relações subjacentes.

As associações sistematicamente baixas da variável “CAUSA_PENDENTE” com as demais suportam a tese deste estudo de que causas de pendências são multidimensionais e pouco previsíveis apenas pelas variáveis analisadas inicialmente. Isso justifica a necessidade de avançar para modelagens preditivas e agrupamentos (como o K-Modes), em linha com o objetivo do estudo de obter insights granulares para mitigação de gargalos.

Essa descoberta sugere que características dos chamados, já citados no parágrafo anterior, estão interligadas, o que pode ser fundamental para a modelagem preditiva, permitindo a identificação de quais tipos específicos de fechamento estão mais fortemente associados a quais causas de pendência. Assim, a análise integrada dessas variáveis suporta a antecipação e o gerenciamento proativo de situações problemáticas no fluxo do atendimento técnico.

A figura 1 é uma representação visual da matriz de associação entre as variáveis categóricas e numéricas pré-selecionadas do conjunto de dados de chamados de suporte técnico.

Figura 1 - *Heatmap* de Associação.



Essa ferramenta metodológica desempenha um papel fundamental na fase exploratória, permitindo a identificação de relações e padrões subjacentes antes da aplicação de técnicas de mineração de dados mais complexas, como o agrupamento. O objetivo primário de um *heatmap* de associação, na identificação de relações, é mapear visualmente a força e, em alguns casos, a direção das relações entre pares de variáveis.

Para variáveis categóricas, tais como: TIPO_CHAMADO, CAUSA_PENDENTE, TIPO_FECHAMENTO, GRUPO_SERVICO, SERVICO, CATEGORIA, UNIDADE_USUARIO_SOLICITANTE, conforme Seção 4.3 e 4.5, a "associação" é tipicamente medida por estatísticas como o V de Cramer, o coeficiente de contingência ou o Qui-quadrado, que quantificam a dependência entre variáveis nominais. Para variáveis numéricas (TEMPO_RESOLUCAO, TEMPO_RESOLUCAO_HORAS) ou entre numéricas e categóricas (após a codificação das categóricas), podem ser utilizados coeficientes de correlação (Miot, 2018).

A identificação de associações significativas é essencial para a modelagem preditiva e para o agrupamento. Variáveis fortemente associadas podem indicar redundância ou, mais pertinentemente, que uma variável é um forte preditor ou está fortemente ligada à ocorrência de outra, como um tipo de fechamento específico com uma causa de pendência.

4.5 Pré-processamento dos Dados

O pré-processamento de dados é uma etapa crítica na Mineração de Dados, garantindo a qualidade, acurácia e eficiência dos processos analíticos subsequentes (Hand; Mannila; Smyth, 2001). Para este estudo, o processo foi sistematicamente dividido em quatro macroetapas detalhadas a seguir

4.5.1 Integração de Dados

- Corrigir Inconsistências Semânticas: Através de uma declaração DML SELECT, aproximadamente 400 versões textuais do campo CAUSA_PENDENTE foram padronizadas para descrições uniformes, aumentando a consistência dos dados categóricos;
- Mesclar Dados: Realizou-se uma série de junções internas (inner joins) para consolidar informações de diferentes tabelas do sistema Assyst. As junções incluíram: sa.assyst usr com TECNICO_ACT pelo atributo assyst_usr_id; incident com act_rec pelo atributo incident_id; inc_cat com inc pelos atributos inc_cat_id e item_id; e product com it pelo atributo product_id.

4.5.2 Redução de Dados

- Selecionar Atributos: Após consenso com o stakeholder, um conjunto reduzido e relevante de atributos foi selecionado para a análise. Os atributos mantidos foram: TIPO_CHAMADO, UNIDADE_USUARIO_SOLICITANTE, GRUPO_SERVIÇO, SERVIÇO, CATEGORIA, CAUSA_PENDENTE, TIPO_FECHAMENTO, DT_ABERTURA, TEMPO_RESOLUCAO e TEMPO_RESOLUCAO_HORAS.
- Selecionar Registros: Foram mantidos apenas os registros cuja data de abertura (DT_ABERTURA) correspondesse ao ano de 2024, utilizando um método de *random sampling* (amostragem aleatória) sem reposição para garantir a imparcialidade da amostra.

4.5.3 Limpeza de Dados

- Corrigir Valores Inconsistentes: Esta etapa focou na retificação do formato do atributo TEMPO_RESOLUCAO. Os valores originais, exportados do banco de dados em um formato incompatível com as ferramentas de análise, foram corrigidos por meio de um script desenvolvido em Python, assegurando a integridade dos demais atributos.

4.5.4 Transformação de Dados

- Normalizar os Dados: Aplicou-se a normalização para ajustar a escala dos atributos numéricos, garantindo que nenhuma variável dominasse indevidamente o processo de análise devido à sua magnitude;
- Codificar Dados Categóricos: Para preparar os atributos categóricos para os algoritmos de aprendizado de máquina, foi empregada a técnica de *One-Hot Encoder*, transformando cada categoria em uma nova variável binária, evitando vieses de ordenamento.

Os dados brutos utilizados são de natureza confidencial do órgão do Poder Judiciário. Entretanto, todo o conjunto de dados foi integralmente anonimizado e pré-processado para

remover qualquer informação sensível ou que pudesse identificar indivíduos ou unidades específicas. Para garantir a reprodutibilidade e a transparência, os autores se colocam à disposição na possibilidade de disponibilizar uma versão do dataset processado e anonimizado em um repositório de dados ou mediante solicitação formal e aprovação do órgão, conforme a política de dados abertos e privacidade. Assim como, em relação ao algoritmo de agrupamento.

K-Modes e as demais etapas de processamento e análise foram implementados utilizando a linguagem Python e suas bibliotecas padrão. O código-fonte dos scripts relevantes para o agrupamento (K-Modes) e a validação será disponibilizado mediante solicitação formal, para facilitar a verificação e replicação por outros pesquisadores.

4.6 Metodologia Experimental

A metodologia experimental foi desenhada para investigar os padrões de pendência em chamados de assistência técnica, com foco na aplicação de técnicas de agrupamento e validação.

4.6.1 Algoritmo de Agrupamento K-Modes

Dada a natureza predominantemente categórica dos atributos selecionados (TIPO_CHAMADO, UNIDADE_USUARIO_SOLICITANTE, GRUPO_SERVICO, CATEGORIA, CAUSA_PENDENTE), o algoritmo K-Means, que se baseia em médias e distâncias euclidianas, mostrou-se inadequado para esta análise (Jain et al., 1999). Optou-se, portanto, pelo K-Modes, uma extensão robusta do K-Means desenvolvida especificamente para dados categóricos (Huang, 1998; De Vos, 2015; Ahmad; Khan, 2013).

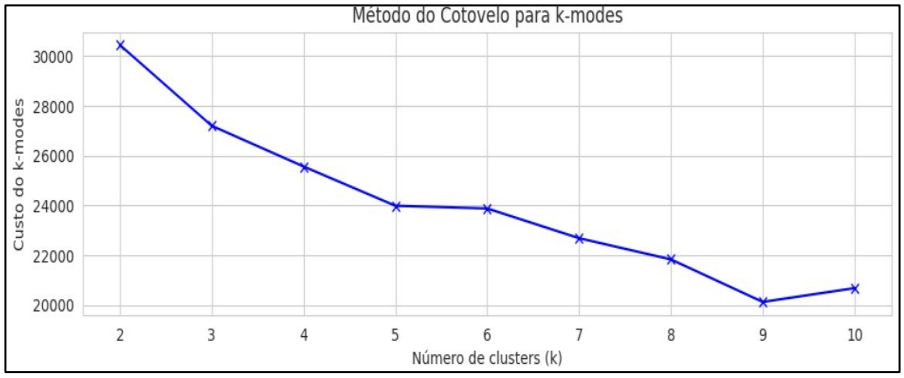
O K-Modes define os centróides de cada grupo com base na moda (o valor mais frequente) dos atributos categóricos dentro de cada cluster, em vez da média. A medida de dissimilaridade utilizada foi a distância de Hamming, que quantifica o número de atributos em que dois pontos de dados diferem. Essa abordagem permitiu agrupar chamados por similaridade de perfis, revelando padrões coerentes com as características qualitativas dos dados.

4.6.2 Definição do Número de *Clusters*

A determinação do número ideal de *clusters* (o hiperparâmetro k) foi realizada por meio do Método do Cotovelo (Elbow Method) (Jain et al., 1999). Esta técnica envolve a plotagem do custo (ou soma dos desvios para a moda) em função do número de clusters e a identificação do ponto de inflexão na curva, que representa o balanço ótimo entre a redução do custo e a complexidade do modelo.

Nos dados analisados, o ponto de inflexão ocorreu em $k \approx 5$, sugerindo que cinco grupos distintos proporcionavam a melhor segmentação sem excesso de granularidade, conforme ilustrado na figura 2. Outros valores para k foram empiricamente testados, mas não apresentaram resultados superiores.

Figura 2 - Método do Cotovelo (*Elbow Method*).



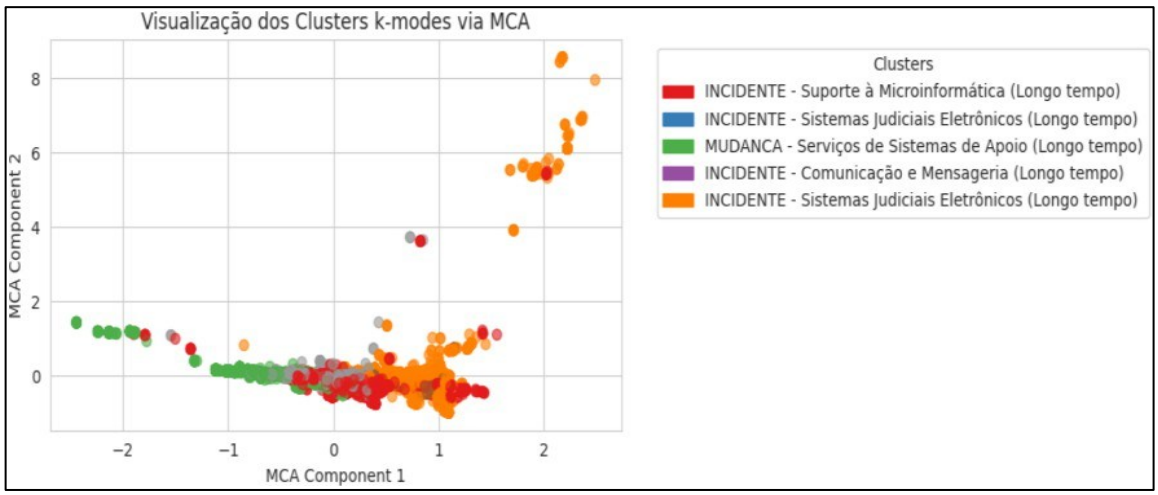
4.6.3 Visualização e Validação dos *Clusters* via Análise de Correspondência Múltipla

Para visualizar os *clusters* resultantes da aplicação do K-Modes e confirmar sua coerência, utilizou-se a ACM, uma técnica estatística robusta para dados categóricos (Abdi; Williams, 2010; Greenacre; Blasius, 2017). A ACM projeta as variáveis categóricas e os clusters em um espaço contínuo bidimensional, permitindo uma representação visual clara das relações entre eles (conforme apresentado na Figura 3, K-modes).

Os *clusters* definidos pelo K-Modes, quando projetados pela ACM, apresentaram boa separação e coerência com os perfis esperados dos chamados, conforme se vê na figura 3:

- *Cluster* Vermelho (Incidentes de Suporte à Microinformática): Apresentam-se isolados e concentrados, indicando características muito específicas e homogêneas
- *Clusters* Azul e Laranja (Incidentes em Sistemas Judiciais Eletrônicos): Estão próximos, mas com diferenciação, sugerindo perfis semelhantes, mas com nuances;
- *Cluster* Verde (Mudanças em Sistemas de Apoio): Caracterizados por uma maior dispersão e um perfil distinto, indicando mais variabilidade interna nas mudanças em sistemas de apoio;
- *Cluster* Roxo (Problemas de Comunicação e Mensageria): É o menor e mais isolado, com um perfil de pendência muito particular.

Figura 3 - K-Modes.



A observação de alguma sobreposição parcial de pontos sugeriu a existência de perfis limítrofes entre clusters, indicando a complexidade inerente aos dados. A ACM não só confirmou a validade dos agrupamentos, mas também facilitou a interpretação dos eixos: o primeiro componente se mostrou eficaz para distinguir os perfis principais, enquanto o segundo capturou variações secundárias. Esta análise orienta a formulação de ações específicas por grupo, como melhorias direcionadas ou ajustes operacionais.

A figura 3 confirma visualmente a estrutura dos agrupamentos, reforça a correlação entre as características dos chamados e as causas de pendência, e é crucial para a i*nterpretabilidade* dos resultados e para a formulação de ações estratégicas direcionadas para cada tipo de chamado no Poder Judiciário.

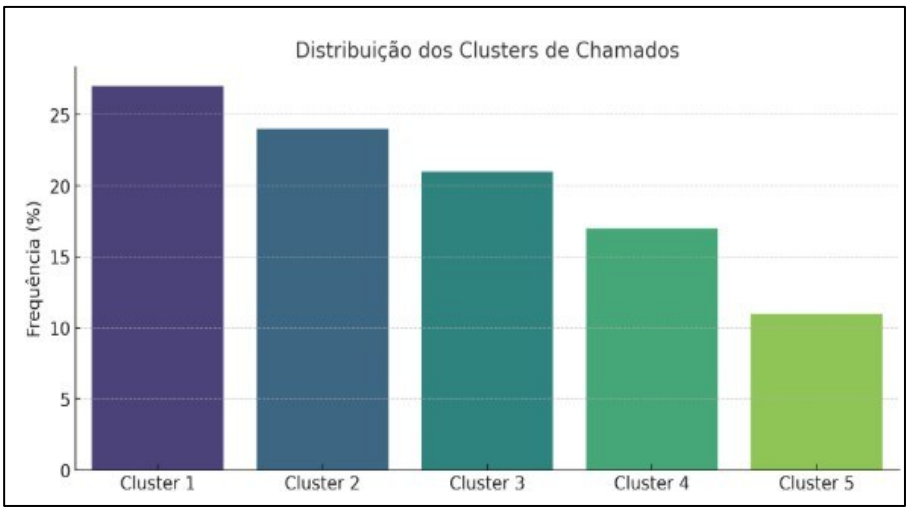
5. Análise e Discussão dos Resultados

Esta seção apresenta e discute os resultados obtidos a partir da aplicação das técnicas de Mineração de Dados, em conformidade com a metodologia delineada na Seção 4. Os achados foram analisados criticamente, buscando responder à questão central de pesquisa e avaliar as implicações teóricas e práticas para a gestão de chamados de suporte técnico no Poder Judiciário. A integração entre os resultados empíricos e a discussão conceitual permitiu uma compreensão aprofundada dos padrões de pendência e suas causas subjacentes.

A aplicação do algoritmo de agrupamento K-Modes, com o hiperparâmetro k definido como 5 (cinco) por meio do Método do Cotovelo (Figura 2), foi bem-sucedida na identificação de cinco clusters distintos de chamados.

Cada um desses agrupamentos revelou características semelhantes em relação às causas de pendência e variáveis operacionais, demonstrando que as interrupções no atendimento não ocorrem de maneira aleatória, mas seguem padrões operacionais específicos que podem ser categorizados e analisados (Huang, 1998; De Vos, 2015; Ahmad; Khan, 2013). A figura 4, representa a proporção de chamados em cada *cluster*.

Figura 4 - Distribuição dos *Clusters* de Chamados.



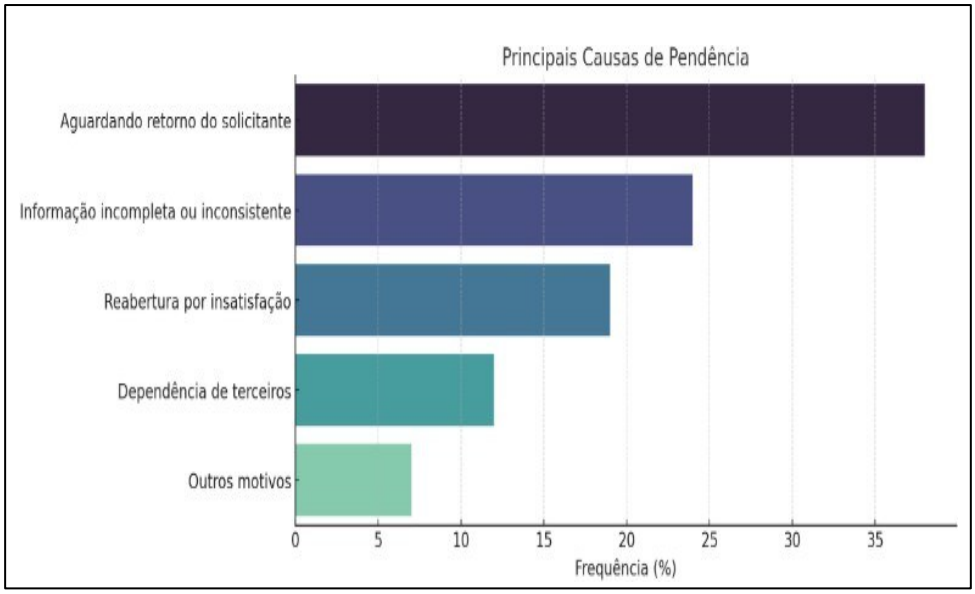
A análise detalhada de cada *cluster* permitiu a identificação de perfis operacionais distintos e suas respectivas implicações:

- *Cluster* 1 (27% dos chamados): Este grupo foi predominantemente composto por incidentes de suporte à microinformática, caracterizados por uma baixa complexidade técnica. A principal causa de pendência identificada foi o aguardo de retorno do solicitante, o que resultou em um tempo médio de resolução de aproximadamente 22 horas. Esta causa foi reforçada pela validação do Analista Judiciário de TI, que confirmou a alta incidência e o impacto operacional dessa causa específica na rotina do suporte. Este achado é particularmente relevante, pois aponta para a necessidade de estratégias mais eficazes de comunicação e engajamento do usuário, bem como para a implementação de protocolos claros para o encerramento de chamados que dependem de informações externas. A demora no retorno do solicitante, embora pareça um fator externo à equipe técnica, impacta diretamente a eficiência operacional e o cumprimento dos níveis de serviço;
- *Cluster* 2 (24% dos chamados): Este agrupamento compreendeu chamados relacionados a sistemas judiciais, com foco em problemas de autenticação e integração. As pendências neste cluster foram frequentemente associadas à intervenção de equipes externas ou de terceiro nível de suporte, culminando em um tempo médio de resolução de 48 horas. A maior complexidade e a dependência de múltiplos agentes ou especialistas para a solução indicam a importância de políticas robustas de escalonamento e de acordos de nível de serviço (SLA) bem definidos com equipes de suporte especializadas, tanto internas quanto externas;
- *Cluster* 3 (21% dos chamados): Este *cluster* concentrou incidentes administrativos em sistemas de apoio. A causa de pendência mais comum foi a falta de informações no momento da abertura do chamado, o que gerou reiterações de contato e um consequente aumento no ciclo de atendimento, com um tempo médio de resolução de 35 horas. Este resultado sublinha a necessidade de aprimorar os mecanismos de registro inicial dos chamados, possivelmente através de formulários mais estruturados ou sistemas de validação de informações que minimizem a entrada de dados incompletos. A educação dos usuários sobre a importância de fornecer detalhes precisos desde o início também se mostra fundamental para mitigar este tipo de pendência;
- *Cluster* 4 (17% dos chamados): Este grupo englobou casos que envolviam modificações ou atualizações em sistemas corporativos, exigindo intervenções coordenadas entre diferentes áreas. Notavelmente, este cluster apresentou o maior tempo médio de resolução (72 horas) e o maior percentual de reabertura (acima de 30%). A complexidade intrínseca de modificações sistêmicas, que frequentemente envolvem múltiplas equipes e testes rigorosos, justifica o tempo de resolução. Contudo, a alta taxa de reabertura sugere que as soluções implementadas podem não estar sendo totalmente eficazes ou que a comunicação pós-intervenção é inadequada, levando à insatisfação do usuário. Isso indica a necessidade de processos de validação de entrega mais rigorosos e de uma comunicação mais transparente com o usuário sobre o escopo e as limitações das modificações;
- *Cluster* 5 (11% dos chamados): Este *cluster* foi caracterizado por problemas de comunicação e mensageria institucional, com pendências frequentemente ligadas a configurações de usuário ou dependências de recursos de rede. Este grupo apresentou o menor índice de satisfação nas pesquisas pós-atendimento. A baixa satisfação, apesar de um tempo de resolução não explicitamente o maior, pode ser atribuída à frustração gerada por problemas que afetam a comunicação diária do usuário, ou à natureza das soluções,

que podem ser percebidas como temporárias ou paliativas. Este resultado sugere a necessidade de uma revisão das soluções padrão para esses tipos de problemas e uma atenção especial à experiência do usuário em serviços de comunicação.

Em uma visão consolidada, as causas de pendência mais frequentes em toda a base de dados foram: aguardando retorno do solicitante (38%), informação incompleta ou inconsistente (24%), reabertura por insatisfação (19%) e dependência de terceiros (12%), conforme detalhado na figura 5 (Pendências mais frequentes). Esses dados reforçam que uma parcela significativa das interrupções é influenciada por fatores relacionados à qualidade da informação inicial e ao engajamento do usuário, bem como por dependências externas ou internas que requerem coordenação.

Figura 5 - Pendências mais frequentes.



A ACM, utilizada para a visualização dos clusters, evidenciou uma separação clara entre os grupos e reforçou a correlação entre as variáveis categóricas "tipo de fechamento" e "causa da pendência" (Abdi; Williams, 2010; Greenacre; Blasius, 2017). A Figura 3 (K-modes) demonstra graficamente essa separação e coerência, validando a estrutura de agrupamento identificada pelo K-Modes.

Essa correlação é um achado metodologicamente significativo, pois sugere que a natureza do fechamento de um chamado não é independente da causa de sua pendência, mas sim influenciada por ela, abrindo caminho para a criação de modelos preditivos mais precisos.

Os achados gerais desta pesquisa indicam a existência de padrões recorrentes e previsíveis nas pendências de atendimento. A capacidade de identificar esses padrões permite uma gestão mais proativa e direcionada, em contraste com abordagens reativas. Isso possibilita o desenvolvimento de estratégias de intervenção específicas para cada perfil de chamado. Tais como, chamados do Cluster 3 (incidentes administrativos), a implementação de campos de preenchimento obrigatório e validação de informações no momento da abertura pode reduzir drasticamente as pendências por informações insuficientes. Para o Cluster 1

(microinformática), campanhas de comunicação direcionadas aos usuários para incentivar o retorno rápido de informações podem ser eficazes. Já para os Clusters 2 e 4, a otimização dos processos de escalonamento e a melhoria na coordenação com equipes externas ou de terceiro nível são cruciais.

Em comparação com os trabalhos relacionados analisados na Seção 3, este estudo se diferencia por aprofundar a análise nas fases intermediárias do atendimento, em vez de focar apenas na abertura ou no desfecho final do chamado (Webster; Watson, 2002). Essa abordagem granular permite uma compreensão mais acurada dos gargalos operacionais específicos que levam às pendências, oferecendo insights diretos para planos de ação institucionalmente relevantes.

A identificação de padrões de pendência por cluster específico, com causas e tempos de resolução distintos, fornece subsídios para uma alocação mais eficiente de recursos humanos e tecnológicos, além de orientar programas de treinamento e políticas de comunicação.

Em suma, os resultados obtidos com a aplicação do algoritmo K-Modes, combinados com a validação pela Análise de Correspondência Múltipla, fornecem evidências robustas da complexidade dos chamados de assistência técnica no Poder Judiciário. A formação de clusters bem definidos revela que as pendências seguem padrões operacionais específicos, relacionados à natureza do serviço, ao nível técnico demandado e à qualidade da informação fornecida. A forte correlação entre o tipo de fechamento e a causa da pendência sugere a viabilidade de modelos preditivos que alertem a equipe técnica sobre chamados com alta probabilidade de interrupções, permitindo intervenções proativas e, consequentemente, uma significativa melhoria na governança da Tecnologia da Informação e Comunicação (TIC) e na percepção de qualidade dos serviços prestados.

6. Limitações da Pesquisa

Este estudo, especialmente, por sua natureza exploratória e focada em um determinado contexto e recorte temporal, possui inerentes limitações que podem moldar o escopo de suas conclusões e a generalização possível de seus resultados.

O reconhecimento e a explicitação dessas fronteiras são cruciais para a integridade científica do estudo, fornecendo uma compreensão clara dos parâmetros dentro dos quais os achados devem ser interpretados e servindo como um catalisador para futuras investigações.

Neste contexto, identificam-se as limitações deste estudo, tais como:

- Restrição Temporal dos Dados: A análise realizada neste estudo baseou-se exclusivamente em um conjunto de dados abrangendo o período de 01 de janeiro de 2024 a 31 de dezembro de 2024, conforme detalhado na Seção 4.3 - Descrição da Base de Dados. Embora este recorte tenha permitido uma investigação aprofundada dos padrões de pendência durante um ano fiscal completo, é importante notar que as dinâmicas operacionais, as tecnologias empregadas e os comportamentos dos usuários podem evoluir ao longo do tempo. Consequentemente, as conclusões aqui apresentadas são mais representativas do período analisado e podem não capturar variações ou tendências significativas que se manifestem em outros períodos temporais;
- Estudo de Caso em um Único Contexto Institucional: Os dados utilizados nesta pesquisa

foram coletados de um único órgão do Poder Judiciário brasileiro. Embora esta abordagem de estudo de caso tenha permitido uma análise granular e a validação dos achados por especialistas do domínio dentro de um ambiente real e de alta criticidade, ela impõe uma limitação à generalização direta dos resultados para outras instituições. Diferenças regionais, culturais, processuais e de infraestrutura tecnológica entre distintas entidades do Poder Judiciário ou outros setores da administração pública podem influenciar os padrões de pendência. Contudo, os insights metodológicos e os tipos de padrões identificados oferecem uma base valiosa para futuras replicações e comparações.

É fundamental destacar que estas limitações não diminuem a validade interna e a relevância dos achados no contexto específico do estudo, mas sim balizam sua interpretação e o direcionamento de pesquisas subsequentes. As perspectivas para mitigar e explorar essas limitações são abordadas detalhadamente na Seção 7 - Trabalhos Futuros, onde propomos a ampliação e replicação deste estudo em outros órgãos, bem como a integração de análises em tempo real para validar a generalização dos padrões identificados e o aprimoramento contínuo das abordagens.

7. Trabalhos Futuros

A presente pesquisa lança as bases para avanços significativos na gestão de serviços de TI no setor público, mas também abre diversos caminhos para aprofundamentos metodológicos e aplicações práticas. Neste sentido, as seguintes direções são propostas para trabalhos futuros:

- Integração de Modelos Preditivos Avançados para Previsão em Tempo Real – ir além da mera "integração" e focar na construção de modelos preditivos robustos e capazes de operar em tempo real. Neste contexto, investigar e comparar a performance de algoritmos de aprendizado de máquina supervisionado, como *Random Forest*, *Gradient Boosting Machines* (e.g., *XGBoost*, *LightGBM*) e Redes Neurais Artificiais (RNAs), para prever a probabilidade de um chamado sofrer pendência ou ser reaberto com base em suas características iniciais. O objetivo seria fornecer alertas proativos às equipes de suporte, permitindo intervenções preventivas antes que a pendência se estabeleça ou se agrave. Isso representa uma transição da análise descritiva para a capacidade prescritiva;
- Desenvolvimento de *Dashboards* Operacionais Interativos com Capacidade Preditiva – enfatizar a usabilidade e a capacidade de suporte à decisão desses dashboards. Assim, propor a criação de painéis analíticos interativos, integrados diretamente ao Sistema Assyst ou a sistemas correlatos, que não apenas monitorem indicadores de desempenho em tempo real (e.g., causas de pendência, tempos médios de resolução, índice de reabertura), mas também visualizem as previsões dos modelos preditivos. Esses dashboards deveriam oferecer insights acionáveis para gestores e equipes, permitindo ajustes dinâmicos na alocação de recursos e na priorização de chamados;
- Ampliação e Generalização da Base de Dados – replicar o estudo em outros órgãos do Poder Judiciário Federal e/ou Estadual. Essa ampliação permitiria não apenas validar o grau geral dos padrões de pendência identificados, mas também realizar análises comparativas (*benchmarking*) entre diferentes unidades. Isso revelaria boas práticas institucionais e variações regionais ou processuais, enriquecendo a compreensão do fenômeno e potencializando a adoção de soluções padronizadas onde aplicável;
- Incorporação de Dados Qualitativos e Análise de Sentimento para Compreensão Holística – complementar a análise quantitativa com dados qualitativos. Isso incluiria a realização

de entrevistas com usuários finais e equipes de suporte, aplicação de questionários abertos, e, de forma mais avançada, a aplicação de técnicas de análise de sentimento sobre o conteúdo textual das pesquisas de satisfação e dos registros de comunicação nos chamados. Essa abordagem multimétodo proporcionaria uma visão mais holística da experiência do usuário e das percepções da equipe, revelando nuances e fatores emocionais que dados estruturados não capturam, aprofundando a compreensão das causas de insatisfação e pendência.

Essas propostas visam não apenas a continuidade científica da investigação, mas também a aplicação concreta dos resultados em prol da modernização dos serviços públicos, alinhando-se às diretrizes de transformação digital e eficiência administrativa no setor público brasileiro.

8. Conclusão

Este estudo teve como objetivo investigar as causas e padrões de pendência no atendimento de chamados de suporte técnico, utilizando técnicas de mineração de dados em uma base de 13.102 registros do Sistema Assyst, provenientes de um órgão do Poder Judiciário Brasileiro. A pesquisa abordou a questão central de como padrões podem ser extraídos de dados históricos para identificar fatores que contribuem para interrupções no atendimento, oferecendo uma resposta robusta através de uma análise exploratória e preditiva.

Os resultados obtidos demonstram de forma inequívoca que as pendências não constituem eventos aleatórios, mas sim refletem características estruturais e operacionais subjacentes aos chamados. A aplicação do algoritmo K-Modes revelou a existência de cinco perfis distintos de chamados, cada um com padrões específicos de pendência, tempos de resolução e causas recorrentes, permitindo uma compreensão aprofundada da heterogeneidade dos desafios enfrentados pelo suporte técnico. A relevância e a representatividade desses agrupamentos foram corroboradas por meio da validação com os stakeholders-chave do TRT, em especial o Analista Judiciário do setor de TI, que reconheceu os padrões identificados como condizentes com a realidade operacional da instituição (conforme detalhado na Seção 4.2).

As principais causas de interrupção identificadas: aguardo de retorno do solicitante, ausência de informações essenciais, dificuldades técnicas não resolvidas internamente e reaberturas por insatisfação do usuário, reafirmam a complexidade da gestão de serviços de TI, que transcende a mera capacidade técnica, envolvendo aspectos de comunicação, qualidade da informação de entrada e processos de coordenação. A forte correlação entre o tipo de fechamento do chamado e a causa da pendência, validada pela Análise de Correspondência Múltipla, evidencia a propriedade preditiva dessas características.

Portanto, conclui-se que a análise preditiva, aplicada de forma criteriosa a dados históricos de suporte técnico, possui um potencial transformador para a gestão pública. Ela não apenas ratifica a hipótese de que é possível prever a propensão de um chamado a sofrer pendências, como também oferece subsídios valiosos e interpretáveis para o redesenho de fluxos operacionais, o aprimoramento de estratégias de atendimento e a alocação mais eficiente de recursos. A validação prática desses achados pelos especialistas do domínio (conforme Seção 4.2) reforça o caráter acionável das conclusões, garantindo que as propostas de melhoria sejam pertinentes e passíveis de implementação no contexto do Poder Judiciário.

Este estudo contribui significativamente para a governança de TIC no Judiciário, alinhando-se às diretrizes de transformação digital e promovendo a melhoria contínua da eficiência e da qualidade dos serviços prestados.

Referências

ABDI, H.; WILLIAMS, L. J. Multiple correspondence analysis. **Wiley Interdisciplinary Reviews: Computational Statistics**, v. 2, n. 3, p. 306-310, 2010.

ADEODATO, P. J. L.; ARNAUD, A. L.; BRAZ, V. M.; Vasconcelos, G. C. **Previsão de No-Show no Agendamento de Serviços Médicos Baseada em Mineração de Dados**. Centro de Informática, Universidade Federal de Pernambuco (UFPE), Recife, Pernambuco, Brasil, 2008. Disponível em: https://run.unl.pt/bitstream/10362/114092/1/TGI0388.pdf.

AHMAD, A.; KHAN, S. S. A survey of k-modes clustering algorithms. **International Journal of Computer Applications,** v. 73, n. 4, 2013.

AMORIM, L. R. **Proposta De Melhoria Dos Procedimentos Construtivos De Uma Construtora Com Vistas à Minimização Do Aparecimento De Manifestações Patológicas**. Trabalho de Conclusão de Curso - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2020.

BIANCHINI, D.; BONO, C.; CAMPI, A.; CAPPIELLO, C.; CERI, S.; DE LUZI, F. MECELLA, M.; PERNICI, B.; PLEBANI, P. Challenges in AI-supported process analysis in the Italian judicial system: what after digitalization? **ACM Digital Library**, 2023. Disponível em: https://dl.acm.org/doi/10.1145/3630025.

CAVALCANTI, G. C. B. **Procedimento de assistência técnica para empresas construtoras de edificações residenciais.** Dissertação de Mestrado - Instituto de Pesquisas Tecnológicas do Estado de São Paulo, São Paulo. Novembro, 2012.

DE VOS, N. J. **Kmodes categorical clustering library**. 2015. Disponível em: https://github.com/nicodv/kmodes. Acesso em: 17/05/25.

FANTINATTI, P. A. P. **Ações de Gestão do Conhecimento na Construção Civil: evidências a partir da assistência técnica de uma construtora**. Dissertação (Mestrado em Engenharia Civil) – Escola de Engenharia, Universidade Estadual de Campinas, Campinas, 2008. Disponível em: https://repositorio.unicamp.br/acervo/detalhe/421706

FAYYAD, U; PIATETSKY-SHAPIRO, G; SMYTH, P. From Data Mining to Knowledge Discovery in Databases. **ACM SIGKDD**. June 1999. Disponível em: https://dl.acm.org/doi/pdf/10.1145/846170.846181.

FREEMAN, R. E. **Strategic Management: A Stakeholder Approach**. Pitman, 1984.
GOMES, D. C.; CARVALHO, D. R.; CUBAS, M. R.; SHMEIL, M. A. H. Mineração de

Dados no Serviço de Atendimento de Urgências. **Journal Health Inform**, 2014. Outubro-Dezembro. Disponível em: https://jhi.sbis.org.br/index.php/jhi-sbis/article/view/302/215.

GREENACRE, M. J.; BLASIUS, J. **Correspondence Analysis in Practice**. 3rd ed. Chapman and Hall/CRC, 2017.

HAN, J.; KAMBER, M.; Pei, J. **Data Mining: Concepts and Techniques**. 3rd ed. Morgan Kaufmann Publishers, 2012.

HAND, D; MANNILA, H; SMYTH, P. **Principles of Data Mining**. MIT Press, 2001. Disponível em: DOI: 10.1016/S0933-3657(02)00058-1.

HUANG, Z. Extensions to the k-means algorithm for clustering large data sets with categorical values. **Data Mining and Knowledge Discovery**, 2(3), 283–304, 1998. Disponível em: https://doi.org/10.1023/A:1009769707641. Acesso em: 17/05/25.

JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data clustering: a review. **ACM computing surveys** (CSUR), v. 31, n. 3, p. 264-323, 1999.

MAYER, V. F.; Marquesa, O. R. B.; Plastino, A. Uso de Múltiplos Canais de Atendimento: Estudo Sobre o Comportamento dos Consumidores da Ampla Energia S.A. Utilizando Técnicas de Mineração de Dados. **Sistemas & Gestão** 9, pp 340-354, 2014. Disponível em:

https://web.archive.org/web/20160110092922id_/http://www.revistasg.uff.br:80/index.php/sg/article/viewFile/V9N3A10/SGV9N3A10. Acesso em: 29/03/2025.

McCONNELL, D. J.; ZHU, J.; PANDYA, S.; AGUIAR, D. Case-level prediction of motion outcomes in civil litigation. **ACM Digital Library**, 2021. Disponível em: https://dl.acm.org/doi/10.1145/3462757.3466101.

MIOT, Hélio Amante. Análise de correlação em estudos clínicos e experimentais. **Jornal Vascular Brasileiro**, São Paulo, v. 17, n. 4, p. 275-279, 2018. Disponível em: https://www.scielo.br/j/jvb/a/YwjG3GsXpBFrZLQhFQG45Rb/.

REZENDE, Solange Oliveira. (Org.). **Sistemas inteligentes: fundamentos e aplicações**. Barueri, SP: Manole, 2005.

WEBSTER, J.; WATSON, R. T. Analyzing the past to prepare for the future: Writing a literature review. **MIS Quarterly.**