

Dados no apoio à gestão pública estadual: conhecendo o “Programa Inteligência Artificial do Paraná – PIÁ”

RESUMO

Marcelo Maia

Marcos Maia

Denise Fukumi Tsunoda

Cicero Aparecido Bezerra

O Estado do Paraná criou uma iniciativa para uso de inteligência artificial denominado Paraná Inteligência Artificial (PIÁ), responsável pela centralização dos serviços prestados pelo governo. A consulta de informação pelos usuários, torna a captura desses dados uma possibilidade da utilização dos mesmos para possível análise na gestão pública. Com este intuito, procedeu-se mineração de dados na base do programa, a qual fora disponibilizada pela Celepar - Tecnologia da Informação e Comunicação do Paraná. Os dados se referem ao ano de 2019, ano em que a plataforma unificando mais de 400 serviços teve seu lançamento no final do mês de junho. Utilizaram-se os *softwares Microsoft Excel®* para a preparação dos dados e o Weka para a mineração via algoritmo APRIORI. A pesquisa teve início com a parceria entre as instituições Companhia de Informação do Paraná e a Universidade Federal do Paraná, para a constituição da base de dados para a análise do sistema. A partir da disponibilização ocorre a realização das etapas de análise estatísticas e mineração de dados, utilizando o algoritmo de associação (APRIORI). As análises revelaram que no ano de 2019, o período da tarde (12h às 18h exclusive) apresenta maior demanda, exigindo que os serviços devem estar disponíveis independente do “tipo de exibição” (*online e/ou presencial, por telefone, somente online, somente presencial*), passando aos gestores que este horário deve apresentar funcionários disponíveis para atender a população. Pesquisas futuras podem aumentar o escopo da pesquisa, mais períodos de análise, comparando investimentos aos órgãos com a demanda do serviço.

PALAVRAS-CHAVE: Algoritmo APRIORI. Site do Governo do Paraná. Gestão Pública Estadual.

INTRODUÇÃO

Dados do Instituto Brasileiro de Geografia e Estatística (IBGE), publicados em 2019 (IBGE, resolução nº.3, 2019) apontam o estado do Paraná com 11.433.957 habitantes, “[...] crescimento de 0,75% percentual [...] um aumento de 85.020 habitantes em relação aos dados de 2018, [...] o quinto estado mais populoso do país, atrás de São Paulo, Minas Gerais, Rio de Janeiro e Bahia, [...] a capital do Paraná, Curitiba, a cidade mais populosa do estado (1.9333.105)” (PAVANELI, 2019, n.p). O aumento demonstra a abrangência que o governo deve apresentar acesso a serviços, que “[...] o uso das tecnologias amplia o entendimento” (HOJDA; FARINIUK; SIMÃO; 2019, p.80, tradução nossa) e auxilia no processo de ferramentas de gestão com o recurso das Tecnologias de Informação (TI).

O Programa Inteligência Artificial (PIÁ), de acordo com o gestor em atuação do governo do estado do Paraná, no fórum internacional de inovação “ThinkinG”, “debate as práticas de referência internacional e dissemina o uso de inteligência e de inovação na gestão pública” (LAYNES, 2019, n.p), na cidade de Foz de Iguaçu, nos dias 27 e 28 de junho de 2019. No lançamento do programa, o governador em exercício destaca que o “Paraná passa a partir de hoje ser o primeiro estado do Brasil com inteligência artificial a serviço do cidadão, coisa que apenas a iniciativa privada tinha” (LAYNES, 2019, n.p). O programa “é o primeiro [...] de inteligência artificial do Brasil focado na prestação de serviços à população. A plataforma e o aplicativo reúnem mais de 380 serviços do Governo em um só lugar e funcionam como canais de diálogo com o cidadão para atender as demandas e reclamações” (LAYNES, 2019, n.p).

Neste contexto, o PIÁ, completa um ano de execução. O diretor de desenvolvimento de sistemas da Companhia de Tecnologia da Informação e Comunicação do Paraná (Celepar) Daniel Scalet, relata que o “[...] sistema de inteligência artificial demorou três meses para ser construído e foi criado para ser simples, intuitivo e para que seja enxergado rapidamente pela população como uma plataforma auxiliar (LAYNES, 2019, n.p). Desde modo, os dados foram solicitados pelos pesquisadores por meio da ouvidoria da Companhia de Tecnologia da Informação do Paraná (Celepar), sociedade de economia mista do governo do Estado do Paraná. A companhia “desenvolve soluções tecnológicas para facilitar a vida do cidadão”, sendo “empresa responsável pela estrutura tecnológica que mantém os dados de sistemas que atendem as secretarias e órgãos estaduais” (CELEPAR, 2019, n.p). Desde fevereiro de 2020 ocorreu a solicitação por parte dos pesquisadores junto a Celepar, para a disponibilização da base de dados do ano de 2019. Os dados foram anonimizados antes de serem concedidos, ou seja, não permite por meios técnicos a reconstrução do caminho para identificação da pessoa titular.

Com os dados disponibilizados, este estudo tem por objetivo verificar o que os dados do Programa Inteligência Artificial do Paraná – PIÁ podem proporcionar para tomada de decisão na administração pública, descrevendo o sistema, suas características, analisando os dados.

Este estudo se justifica pelo fato de “[...] o governo [...] aproximar-se dos cidadãos [...] disponibilizados na internet como recurso estratégico para facilitar o acesso a informação e serviços, agilizar processos” (DUARTE; et al, 2020, p.54), sendo, “[...] disseminação e acesso às informações, [...] um meio de acesso integrado que oferece aos visitantes um ponto único de contato para

fornecimento de informações” (SIMÃO; RODRIGUES, 2005, p.81). Justifica-se também pelas “[...] aplicações usuais de inteligência artificial, como modelagem de usuário e interfaces inteligentes e auto adaptativas, o intuito [...] conhecer o usuário do sistema, individualmente, na tentativa de detectar suas necessidades e desejos (SILVEIRA, 1996, p.16).

No entanto, no presente estudo serão abordados: a importância da gestão pública, descrevendo a notoriedade perante a União, Estados e Municípios; visão geral da inteligência artificial, conceito e técnica; histórico do Piauí, constituição, realização e funcionalidades; métodos e materiais utilizados, o meio utilizado para obtenção dos dados; resultados e discussões, tratamento e métodos utilizados; considerações finais, apontamento em relação ao objeto estudado e referências.

A IMPORTÂNCIA DA GESTÃO PÚBLICA

A Constituição da República Federativa do Brasil (1988), “[...] formada pela União indissolúvel dos Estados e Municípios e do Distrito federal [...]” (BRASIL, 1988, Art. 1), que compete aos Estados “[...] conservar o patrimônio público; [...] cuidar da saúde e assistência pública; [...] proporcionar aos meios de acesso à cultura, à educação e à ciência; [...] promover programas de construção de moradias [...] saneamento básico; [...] implantar política de educação para segurança no trânsito[...]” (BRASIL, 1988, Art. 1), que “[...] precisa ser avaliada quanto à adequação entre as decisões tomadas pelos gestores públicos” (Ribeiro, et al., 2020, p.52). Assim, Oliveira, (2019, p.107) descreve que “a gestão pública [...] sobre o papel do gestor priorizando, no exercício de suas funções, a transparência e a ética”.

Nesse contexto, “a administração pública compreende, [...] o conjunto de órgãos que compõe um governo para planejamento, organização e coordenação de ações, projetos, serviços [...] e se materializa pelo conjunto das funções necessárias para execução dos serviços públicos em geral” (OLIVEIRA, 2019, p.107), sendo “[...] atividades prestacionais de obrigação do Estado para satisfação dos interesses sociais” (BARRETO, 2014, p.18).

Assim, de acordo com Vaz (2013, n.p), “a gestão pública passa por um processo de globalização [...] alavancado pelo desenvolvimento das tecnologias de informação e comunicação [...] para que apoie o salto de qualidade e produtividade, necessário na área de prestação de serviços públicos à população”, a fim de “promover a competitividade do país, num mundo cada vez mais integrado” (VAZ, 2013, n.p).

A integração ocorre desde o “início dos anos 1990, o governo tenta acompanhar as tendências internacionais de gestão pública para aproximar-se dos cidadãos [...] disponibilizados na internet [...] para facilitar o acesso a informações e serviços, agilizar processos” (DUARTE; et al, 2020, p.54). Desta forma, o governo do estado do Paraná, por intermédio da tecnologia da informação, desenvolve o sistema Paraná Inteligência Artificial (PIÁ), que busca criar mecanismos para que os computadores assemelhem ao ser humano, tanto na forma de pensar como de agir de maneira inteligente, utilizando a inteligência artificial.

VISÃO GERAL DA INTELIGÊNCIA ARTIFICIAL

A Inteligência Artificial (IA) “[...] vêm conduzindo mudanças [...] em nossa sociedade. [...] compreende diferentes áreas [...] possibilita construir soluções que permitam lidar com uma ampla variedade de problemas, trazendo melhorias socioeconômicas significativas para a sociedade” (CÓBE, 2020, p.39). A IA “[...] é um ramo da ciência da computação que usando algoritmos [...] é capaz de reconhecer um problema, ou uma tarefa a ser realizada, analisar dados e tomar decisões, simulando a capacidade humana” (LOBO, 2019, p.3)

Nesse contexto, “as técnicas de IA, associadas à abundante quantidade de dados digitais e ao onipresente poder de processamento paralelo entregue [...] deverão, [...] suprir a alta demanda pública por serviços [...]” (CÓBE, 2020, p.39), demonstrando que “[...] servem como um parâmetro de monitoramento” (CÓBE, 2020, p.46). Nesse contexto, a Companhia de Informação do Paraná (Celepar), desenvolveu o PIÁ.

PROGRAMA INTELIGÊNCIA ARTIFICIAL DO PARANÁ

No segundo semestre de 2016 ocorre a criação da plataforma do governo digital, desenvolvida pela Celepar, que conforme diálogo com a gerente de inovação corporativa “[...] em 2017 começou a oferecer alguns serviços ao cidadão, firmar parcerias com os gestores dos serviços nos diversos órgãos do Estado, alinhamentos com a legislação, definição de arquitetura, com o nome Governo Digital do Paraná”, em que o lançamento oficialmente, ocorreu em 2018, pela governadora em exercício.

No primeiro semestre de 2019, “[...] incluímos na plataforma a solução de *chatbot* com inteligência artificial” (Gerente de Inovação Corporativa – Celepar), “[...] iniciativas com o uso de inteligência, inovação e tecnologia para descomplicar a vida do cidadão paraense. Dentre os produtos está o Paraná Inteligência Artificial “(AGÊNCIA, 2019), “[...] mais conhecido como Piá (nome que faz referência ao termo paranaense para designar meninos)” (DUARTE, et al, 2020, p.58).

Para Duarte (2020, p.58) “[...] é o portal do governo de estado”, lançado na quinta feira (27 de junho de 2019), em Foz do Iguaçu, durante o fórum internacional de inovação *ThinkinG* (debate as práticas de referência internacional e dissemina o uso de inteligência e de inovação na gestão pública), pelo governador em exercício (Carlos Massa Ratinho Junior).

O governador enfatizou que “O PIÁ é um passo importante para o Paraná se tornar o Estado mais inovador do Brasil”, enfatizando que é “[...] inspirado em iniciativas exitosas do setor privado e [...] muitos países conseguiram prestar melhores serviços à população usando esta tecnologia [...] todas as informações que o cidadão paranaense precisa estão disponíveis com muito mais facilidade” (AGÊNCIA, 2019, n.p).

No fórum o presidente da Celepar, Leandro Moura, relata

[...] agregamos valor no uso da inteligência, trabalhamos com a inclusão, registramos em nossas soluções as etapas da vida do cidadão, pois através de tudo que criamos estamos presentes no dia

a dia da população”, “[...] trazendo mais qualidade de vida às pessoas ao eliminarmos filas em processos antes presenciais e agora digitais, devolvendo ao cidadão algo imensurável em seu cotidiano, o tempo, além de enaltecer o ex-presidente (Allan Costa) (AGÊNCIA, 2019, n.p.)

O PIÁ, “[...] congrega, informações de documentos emitidos pelo Governo Federal e pelos municípios. A ideia é ter o Estado no celular, tablet computador [...] reúne programa que o governo precisa prestar de maneira rápida, fácil” (AGÊNCIA, 2019, n.p)

Para a utilização, é necessário “entrar no site www.pia.pr.gov.br e fazer perguntas [...] caso o usuário informe o seu CPF, o nível de customização aumenta, a plataforma responderá o cidadão e, conseqüentemente, ganhará mais recursos a cada pergunta feita pela sociedade (AGÊNCIA, 2019, n.p). “O cadastro [...] é uma forma inovadora de acessar serviços público e informações oficiais pela internet [...] o acesso pode ser feito em qualquer navegador ou pelo aplicativo disponível para iOS e Android” (PIÁ, 2020, n.p). O cadastro pode ser realizado pela Carteira Nacional de Habilitação (CNH) ou carteira de identidade (RG), emitida no Paraná, informar um e-mail e um telefone celular, para receber o SMS (mensagem eletrônica) com o código de segurança.

Uma vez cadastrado, a

[...] plataforma vai avisar sobre vencimento das faturas de água e luz, dos pontos na carteira de motorista, possibilita cadastrar o animal de estimação, agendar atendimento em órgãos públicos, ver e resgatar créditos do Nota Paraná, boletim escolar dos filhos matriculados [...] dentre outros diversos serviços (registro de acidentes sem vítimas pelas rodovias estaduais” (AGÊNCIA, 2019, n.p)”.

O cadastro não é obrigatório, “[...] mas com ele você recebe informações personalizadas, tem acesso mais rápido ao que procura e um maior número de serviços quem podem ser feitos pela internet” (PIÁ, 2020, n.p).

Nesta plataforma, independente de efetuar o cadastro, estão presentes diversos serviços ofertados pelo governo. Estes os serviços, estão separados em partes maiores como: “justiça, família e trabalho”, “segurança”, “trânsito”, “pets”, “agricultura e abastecimento”, “transporte”, “educação”, “esporte”, “meio ambiente”, “cultura”, e cada uma dessas áreas há a separação por empresa (Cohapar, Copel, Compagas, Detran, Sanepar, entre outras). Caso efetue o cadastro, o acesso a informações que apresentem o seu cadastro é facilitado (por apresentar assuntos já pesquisados). Assim, o Paraná é “O primeiro estado do Brasil a facilitar a vida do cidadão” (AGÊNCIA, 2019, n.p), reunindo atualmente (junho de 2020) mais de 4007 serviços (3578 serviços federais, 424 serviços estaduais e cinco serviços municipais).

METODOLOGIA E MATERIAIS UTILIZADOS

A pesquisa é de abordagem quantitativa, classificada como exploratória e descritiva. Os dados foram solicitados pelos pesquisadores por meios formais via

ouvidoria, em que ocorre a realização da parceria entre a Universidade Federal do Paraná, por intermédio do Programa de Pós-Graduação em Gestão da Informação, e a Celepar, quanto a solicitação e preenchimento de termo de anuência, no qual foi requerido a coleta da informação e o termo de compromisso pesquisador.

Para a disponibilização dos metadados em planilha eletrônica, foram necessários pouco mais de dois meses na tramitação dos documentos e liberação dos dados. O quadro 1 apresenta os dados tal como foram disponibilizados pela companhia.

Quadro 1 – Dados disponibilizados pela Companhia de Tecnologia da Informação do Paraná do Programa Inteligência Artificial do Paraná do ano de 2019

| Serviços (1) | Tipo Exec. Serviço (2) | Mês (3) | Data de Publicação (4) | Total (5) | Hora (6) | | | |
|---|-------------------------------|-----------|------------------------|-----------|----------|-------|-----|-------|
| | | | | | 00:00 | 01:00 | ... | 23:00 |
| Acessar a Agência de Notícias do Paraná | Somente <i>online</i> | Janeiro | 15/01/2019 | 328 | 4 | 6 | ... | 10 |
| Acionar o Disque Idoso Paraná | <i>Online</i> e/ou presencial | Fevereiro | 03/10/2018 | 135 | 2 | 4 | ... | 5 |
| Receber orientações sobre direitos autorais | Por Telefone | Março | 06/11/2017 | 45 | 2 | 0 | ... | 1 |
| Agendar visita no Porto de Paranaguá | Diversas formas | Agosto | 25/07/2019 | 16 | 0 | 0 | ... | 0 |

Fonte: Elaborado pelos autores (2020)

(1) Descrição dos serviços (mais de 400).

(2) Exibição das diversas opções de acesso (diversas formas, *online* e/ou presencial, por telefone, somente *online*, somente presencial).

(3) Período (limitado mês a mês) de janeiro a dezembro.

(4) A partir de qual dia o serviço está disponível ao cidadão.

(5) Quantidade total de acessos do serviço.

(6) Separado por hora (compreendendo de 00:00 a 23:00 – 00:00; 01:00; 02:00; 03:00; 04:00; 05:00; 06:00; 07:00; 08:00; 09:00; 10:00; 11:00; 12:00; 13:00; 14:00; 15:00; 16:00; 17:00; 18:00; 19:00; 20:00; 21:00; 22:00; 23:00) quantidade de acessos.

Após o recebimento destes dados, foram incluídas as categorias e subcategorias relacionadas aos serviços prestados pelo programa e constantes nos dados originais. Os serviços foram consultados no site programa Piá e foram responsáveis pela geração das categorias e subcategorias, as quais foram inseridas

na base original de dados. O quadro 2 apresenta os dados inseridos pelos pesquisadores.

Quadro 2 – Inserção das categorias e subcategorias do Programa Inteligência Artificial do Paraná do ano de 2019

| Grandes áreas | Novos menus |
|--------------------------------|---|
| Administração | agência reguladora, documentos, governo digital, legislação estadual, protocolo, servidor público e transparência |
| Agropecuária | agricultura, pecuária e programas |
| Cidadania | assistência social, direitos humanos, proteção e defesa do cidadão, rede de proteção da mulher e voluntariado |
| Comunicação | meios de comunicação |
| Cultura e lazer | biblioteca, dança, museus, música, programação, publicações, teatro e turismo |
| Desenvolvimento urbano | municípios |
| Direitos | assistência jurídica gratuita, direitos autorais e ouvidoria |
| Documentos | documentação escolar, documentos pessoais e emissão |
| Economia e finanças | empresas e impostos |
| Educação e ciência | ação educativa, documentação escolar, educação infantil, educação profissional, ensino fundamental, ensino médio, ensino superior e idiomas |
| Habitação | infraestrutura e programas habitacionais |
| Indústria, comércio e serviços | cotações e preços, defesa do consumidor, empreendedorismo, instituto de pesos e medidas |
| Meio ambiente | licenciamento ambiental, monitoramento ambiental, plantio e recuperação florestal, polícia ambiental, proteção dos animais, qualidade ambiental; taxas e emissão de boletos |
| Saúde | atenção à saúde, covid-19, doações, medicamentos, serviços de emergência, sistema de saúde e vigilância em saúde |
| Segurança | atestado e certidões, corpo de bombeiros, defesa civil, denúncia, desaparecidos, furtos e roubos, polícia, prevenção e repressão ao crime |
| Trabalho | piso salarial, seguro desemprego e vagas |
| Transporte e trânsito | acidente de trânsito, infraestrutura, motorista, multas, portos, transporte de passageiro e veículos |

Fonte: Elaborado pelos autores (2020)

Após estas inclusões, iniciou-se o processo de descoberta de padrões em base de dados, aos quais podem ser “[...] válidos quando aplicados a novos dados (isto é, dados não considerados na construção do modelo ou determinação do padrão); e [...] úteis para o utilizador, auxiliando o processo de tomada de decisão” (SANTOS, 2001, p.84). O objetivo na busca do padrão é para estimar cenários de atendimento, melhorando a organização e distribuição dos colaboradores para a resolução do solicitado, e conseqüentemente é possível diminuir o tempo de resposta para a dúvida ou problema do usuário.

MINERAÇÃO DOS DADOS

Conforme apresentado por Castro e Ferrari (2016), mineração de dados é um dos elementos presentes em um processo de descoberta de conhecimento em

bases de dados (*Knowledge Discovery in Databases* - KDD). Este processo de descoberta se utiliza da aplicação de alguns algoritmos específicos para extração de padrões de dados (FAYYAD, PIATETSKY-SHAPIRO e SMYTH, 1996).

A descoberta de conhecimento em bases de dados envolve algumas etapas (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996), (ADRIAANS; ZANTINGE, 1996), as quais são sintetizadas por Castro e Ferrari (2016) e apresentadas no quadro 3.

Quadro 3 – Etapas do processo de descoberta de conhecimento em base de dados

| Etapas | Denominação | O que será realizado |
|----------|--|---|
| 1ª Etapa | Base de dados | Dados organizados, nível mais básico de abstração |
| 2ª Etapa | Preparação ou pré-processamento de dados | Etapa anterior a mineração, inclui limpeza (remoção de ruídos e dados inconsistente), integração (combinação de dados obtidos de diversas bases), a seleção ou redução (definição dos dados relevantes à análise) e a transformação (consolidação dos dados em formatos apropriados para a mineração) |
| 3ª Etapa | Mineração de dados | Aplicação de algoritmos para extrair conhecimentos. Os algoritmos podem ser agrupamento, predição, associação e detecção de anomalias. |
| 4ª Etapa | Avaliação ou validação do conhecimento | Avaliação dos resultados objetivando identificar conhecimentos. |

Fonte: Adaptado de CASTRO; FERRARI (2016)

A 3ª etapa “mineração de dados” é usada para especificar o tipo de informação a ser obtida na tarefa, ao qual são divididas nas categorias: descritiva e preditiva, quadro 4.

Quadro 4 – Categoria e tarefa da mineração de dados

| Categorias | Definição | Objetivo | Tarefa |
|-------------|--|--|--|
| Descritivas | Propriedades gerais dos dados | Encontrar padrão | - Especializações - Agrupamentos - Sumarização - Modelagem de dependências - Detecção de desvios |
| Preditivas | Inferências dos dados objetivando predições. | Utilizar valores dos atributos descritivos para prever valores futuros ou desconhecidos de outros atributos de interesse | - Classificação - Regressão ou estimação |

Fonte: Adaptado de SILVA; PERES; BOSCAROLI (2016)

Utilizou-se nesta pesquisa a planilha eletrônica *Microsoft Excel* na criação e preparação, etapas 1 e 2, e o aplicativo *Weka* (*software* livre para mineração de dados) na mineração dos dados, etapa 3.

Na 1ª etapa, denominada de obtenção de dados, foram disponibilizados em 14 arquivos do *Microsoft Excel*®, sendo:

- um arquivo, com a denominação “serviços acessados 2019”, contemplando: denominação dos serviços (nome), quantidades de acessos no ano (4.789.704) e o percentual de que cada serviço apresenta em relação ao total;

- doze arquivos, cada um representando um mês do ano, denominado “serviços acessados 2019 horas mês”, apresentando as informações conforme quadro 1, e;

- um arquivo, com os “serviços acessados 2019 mês a mês”, contendo: denominação do serviço, tipo de execução do serviço, total de acessos de 2019, e a separação mês a mês o total de vezes o serviço.

Os dados dos doze arquivos “serviços acessados 2019 hora mês” foram reunidos em uma planilha do *Microsoft Excel*, e incluídas duas colunas com a denominação “grandes áreas” e “novos menus”, conforme apresentados no quadro 2. Cada serviço prestado foi responsável por identificar qual área e menu pertencem e, portanto, foram inseridos nos 436 serviços disponibilizados no programa. Dentre todos os serviços disponíveis, vinte e dois não foram localizados, considerando que os serviços ou foram substituídos ou desativados.

Uma vez que os dados estão disponíveis, passou-se a 2ª etapa no processo de descoberta de conhecimento em bases de dados.

RESULTADOS E DISCUSSÕES

Nesta seção busca-se demonstrar os principais resultados obtidos por esta pesquisa. Os resultados são apresentados conforme estabelecido no descrito quadro 3. Após a realização da 1ª etapa, a 2ª etapa consiste na preparação ou pré-processamento de dados

Com a finalidade de preparar a base de dados para a mineração, excluíram-se os 22 serviços que foram substituídos ou desativados. Desta forma, a base de dados apresenta 4980 linhas, com 414 serviços diferentes. O serviço “solicitar seguro-desemprego”, na grande área “assistência” e novos menus “trabalhos” apresenta no tipo de execução de serviço “on line e/ou presencial” – 24.563 acessos, e também somente “on line” – 99 acessos, contemplando 24 linhas, os demais 413 apresentam 12 linhas cada.

A variável hora (conforme quadro 1) foi alterada de acordo com BRASIL (2014) manual de comunicação da Secretaria Especial de Comunicação Social do Poder Executivo Federal, categorizando-a como: madrugada (da zero hora às 6h, exclusive); manhã (das 6h às 12h, exclusive); tarde (das 12h às 18h, exclusive). Após esta categorização, a quantidade de serviços prestados por hora foi unificada de acordo com as categorias criadas. Isto posto, o quadro 5 apresenta como ficou constituída a base de dados após a criação das categorias: madrugada, manhã, tarde e noite.

Quadro 5 – 10 Exemplo dos dados utilizados para a mineração dos dados

| Grandes áreas | Novos menus | Serviço | Tipo Exec. Serviço | Mês | Ano | Total Madrugada | Total Manhã | Total Tarde | Total noite |
|-----------------------|----------------------|---|------------------------|-------------------|----------|-----------------|-------------|-------------|-------------|
| Comunicação | Meios de Comunicação | Acessar a Agência de Notícias do Paraná | Somente online | Jan eiro | 201 9 | 17 | 99 | 154 | 58 |
| Assistência | Assistência Social | Acionar o Disque Idoso Paraná | Online e/ou presencial | Fev erei ro | 201 8 | 34 | 34 | 54 | 33 |
| Direitos | Direitos Autorais | Receber orientações sobre direitos autorais | Por Telefone | Mar ço | 201 7 | 6 | 7 | 21 | 11 |
| Transporte e Trânsito | Portos | Agendar visita no Porto de Paranaguá | Diversas formas | Ago sto | 201 9 | 2 | 5 | 7 | 2 |

Fonte: Elaborado pelos autores (2020)

Para a utilização do algoritmo do APRIORI, os totais das categorias: madrugada, manhã, tarde e noite, até então variáveis quantitativas discretas, foram agrupados por uma medida de posição separatriz. A utilização dos destas medidas são interessantes em casos onde “[...] existe maior concentração ou dispersão” dos dados (NOVAES, QUEIROZ, COUTINHO, 2013, p.54). A separatriz utilizada foi quartis, a qual “[...] divide a distribuição ordenada em quatro partes iguais” (BRUNI, 2010, p. 80).

Para o cálculo de cada quartil foi utilizada a fórmula (1), de acordo com Bruni (2010, p.80):

$$Q_{nq} = x_{\left[\frac{nq \cdot n}{4} + \frac{1}{2}\right]} \quad (1)$$

Onde:

Q = quartil que se deseja obter

nq = número o quartil que se deseja obter (1,2,3)

x = elemento da série ordenada

n = tamanho da amostra

Os valores que dividem a série ordenada para cada um dos períodos são representados na figura 1.

Figura 1 – Variável período separada em quartil utilizada para a mineração dos dados

| | Min. | 1º quartil | 2º quartil | 3º quartil | Máx. |
|-----------|------|------------|------------|------------|-------|
| Madrugada | 0 | 2 | 7 | 23 | 5090 |
| Manhã | 0 | 5 | 19 | 83 | 31048 |
| Tarde | 0 | 8 | 26 | 107 | 39015 |
| Noite | 0 | 5 | 16 | 68 | 30151 |

Fonte: Elaborado pelos autores (2020)

Verifica-se que no período da tarde (representado no segundo quartil), também denominado de mediana (medida de tendência central), que este período é o que apresenta maior número de acessos aos serviços, seguido pelo período da manhã. Os períodos da manhã e da tarde, são os períodos que abrangem maior número de consultas aos serviços do programa Piá, superior a 83 e 107 acessos o 3º quartil.

Com o apoio da medida separatriz dos quartis, foram desenvolvidas as categorias: baixo, médio e alto. Desta forma, os dados quantitativos discretos associados às variáveis madrugada, manhã, tarde e noite, foram substituídos pelas categorias baixo, médio e alto, conforme apresentado no quadro 6.

Quadro 6 – Categorias utilizada na variável período para a mineração dos dados

| Categorias | Madrugada | Manhã | Tarde | Noite |
|------------|-----------------------|-----------------------|------------------------|-----------------------|
| Baixo | Até 7 exclusive | Até 19 exclusive | Até 26 exclusive | Até 16 exclusive |
| Médio | 7 até 23 exclusive | 19 até 83 exclusive | 26 até 107 exclusive | 16 até 68 exclusive |
| Alto | Acima de 23 inclusive | Acima de 83 inclusive | Acima de 107 inclusive | Acima de 68 inclusive |

Fonte: Elaborado pelos autores (2020)

Observa-se que a categoria: “Baixo” para o período da tarde é considerado inferior a 26 solicitações, “Médio” de 26 a 107 solicitações e “Alto” acima de 107 solicitações. Isto posto, após a preparação dos dados, estes seguiram para a etapa da mineração.

3ª Etapa: Mineração de dados

Nesta etapa foi utilizada a tarefa de associação por intermédio do algoritmo APRIORI, escolhido em virtude de o algoritmo realizar tarefa de descoberta de associação e “[...] alcançar resultados satisfatórios na tarefa de busca por associação entre dados” (ALMEIDA, 2011, p.6). O método “[...] utiliza das características de um padrão frequente já encontrado anteriormente (prior) para buscar mais padrões [...] relacionado à probabilidade de ocorrer uma associação” (ALMEIDA, 2011, p.73). Com confiança de 90% e suporte de 10% os principais resultados estão apresentados no quadro 7.

Quadro 7 – Melhores regras encontradas pelo algoritmo APRIORI no Programa Inteligência Artificial

| | |
|---|--|
| 1 | ano=2019 Total Manha=baixo Total tarde=baixo 755 ==> Total noite=baixo 746 <conf:(0.99)> lift:(2.03) lev:(0.08) [378] conv:(38.78) |
|---|--|

| | |
|----|---|
| 2 | ano=2019 Total Madrugada=baixo Total Manha=baixo Total tarde=baixo 729 ==> Total noite=baixo 720 <conf:(0.99)> lift:(2.03) lev:(0.07) [365] conv:(37.45) |
| 3 | ano=2019 Total Madrugada=baixo Total tarde=baixo 744 ==> Total noite=baixo 733 <conf:(0.99)> lift:(2.03) lev:(0.07) [371] conv:(31.85) |
| 4 | Tipo Exec. Servico=Diversas formas Total Madrugada=baixo Total tarde=baixo Total noite=baixo 584 ==> Total Manha=baixo 575 <conf:(0.98)> lift:(1.99) lev:(0.06) [286] conv:(29.54) |
| 5 | Tipo Exec. Servico=Somente <i>online</i> Total tarde=alto Total noite=alto 641 ==> Total Manha=alto 630 <conf:(0.98)> lift:(3.93) lev:(0.09) [469] conv:(40.04) |
| 6 | Tipo Exec. Servico=Somente <i>online</i> Total Manha=alto Total noite=alto 641 ==> Total tarde=alto 630 <conf:(0.98)> lift:(3.92) lev:(0.09) [469] conv:(40.02) |
| 7 | Tipo Exec. Servico=Somente <i>online</i> Total Madrugada=alto Total tarde=alto Total noite=alto 580 ==> Total Manha=alto 570 <conf:(0.98)> lift:(3.92) lev:(0.09) [424] conv:(39.52) |
| 8 | Tipo Exec. Servico=Somente <i>online</i> Total Madrugada=alto Total Manha=alto Total noite=alto 580 ==> Total tarde=alto 570 <conf:(0.98)> lift:(3.92) lev:(0.09) [424] conv:(39.5) |
| 9 | ano=2019 Total Madrugada=baixo Total Manha=baixo 750 ==> Total noite=baixo 737 <conf:(0.98)> lift:(2.02) lev:(0.07) [372] conv:(27.52) |
| 10 | ano=2019 Total Madrugada=baixo Total tarde=baixo Total noite=baixo 733 ==> Total Manha=baixo 720 <conf:(0.98)> lift:(1.99) lev:(0.07) [357] conv:(26.48) |

Fonte: Elaborado pelos autores (2020)

Observa-se que do número 1 ao 10, estão as melhores regras pelo algoritmo APRIORI, que inicialmente aparece a regra, com a confiança (estimativa da regra), seguido da palavra *lift* (indica quão mais frequente é o lado direito da regra quanto ao lado esquerdo está presente), *lev* - *Leverage* (número adicional coberto pelos lados direito e esquerdo) e *conv* - convicção (relação entre os itens da regra).

A regra 1 representa que: se o ano for 2019, o período da manhã e tarde são “baixo” (com 755 acessos aos serviços), e a noite também é baixa com (746 acessos), regra com 99% de confiança, apresentando dependência positiva entre o consequente (total noite) e o antecedente (total manhã, total tarde), dados os valores *lift* maior que 1 (um), *Leverage* maior que 0 (zero) e valor de convicção 38,78.

Na regra 2, se o ano for 2019, o período da madrugada, manhã e tarde são “baixo” (com 729 acessos aos serviços), e a noite considerado também “baixo (720 acessos), com confiança de 99 %, demonstrando dependência positiva entro o consequente (total noite) e o antecedente (madrugada, manhã e tarde), *lift* maior que 1 (um), *Leverage* maior que 0 (zero) e valor convicção 37,45.

Já na regra 5, se o tipo de execução do serviço for somente *online*, os períodos da tarde e da noite são “alto” (641 acessos), e o da manhã também é alto com 630, com 98% de confiança, com o *lift* de 3,93 (a regra se o tipo de serviço for somente *online*, os períodos de tarde e noite é mais frequente do que no período da manhã, ou seja a regra da direita ocorre com mais frequência nas transações em que o lado esquerdo ocorre).

Na regra 6, se o tipo de execução for somente *online* então o período será manhã alto e tarde alto, ou seja, a regra é determinada (Se X então Y).

Nas regras 5, 6 e 7 observa-se um *lift* alto, 3,93, 3,92 e 3,92, o que significa “alto grau de interesse”, pois uma regra de associação é a relação entre a confiança

da regra e a confiança esperada da regra, conforme o site da *IBM knowledge Center* (2012).

A regra 5 apresenta o *conv* de 40,04 e a regra 10 de 26,48, demonstra a relação entre os itens da regra, ou seja, para a regra 5, se o tipo do serviço for apenas online no período da tarde e a noite (item 1) e no período da manhã (item 2) a relação entre os itens é de 40,04. Na regra 10 (item 1) se for no ano de 2019 no período de madrugada, tarde e noite (item 2) manhã a relação entre os itens é de 26,48.

As informações corroboram no que se refere às demandas do tipo de execução do serviço que, em sua maioria são decorrentes de diversas formas em que o serviço *online* tem as demandas pelo serviço nos períodos da tarde e da noite.

4ª Etapa: Avaliação ou validação do conhecimento

A mineração de dados evidenciou que as demandas de solicitações dos serviços têm maior procura no período da tarde e da noite pelo tipo de execução “somente *online*”. Verificou-se também que em dias em que a procura foi baixa em dois períodos, manhã e noite por exemplo, a procura se mantém baixo nos demais períodos. Assim, observa-se que a gestão pública pode realizar uma campanha maior para os serviços disponíveis *online* para que a população tenha o conhecimento maior e saiba como proceder na respectiva plataforma (site), demonstrando a abrangência desses mais de 400 serviços.

CONSIDERAÇÕES FINAIS

O governo do Paraná é o primeiro estado do Brasil a utilizar a inteligência artificial focada na prestação de serviço à população, com o Programa Inteligência Artificial (PIÁ). Representa na plataforma e no aplicativo mais de 400 serviços do governo apenas em um único local, trabalhando como meio de comunicação entre o cidadão e as demandas para o governo.

Com o levantamento da disponibilização do serviço (data de publicação), foi possível verificar a evolução da preocupação da administração pública em fornecer ao cidadão a possibilidade em encontrar em um único canal a informação dos serviços solicitados.

Foi possível constatar que a demanda de maior acesso aos serviços é no período da tarde (12h inclusive às 18h exclusive), indicando aos gestores públicos que neste período os serviços descritos devem estar disponíveis, independente do “tipo de exibição” (*online* e/ou presencial, por telefone, somente *online*, somente presencial).

Uma limitação do presente estudo se deve ao fato de que somente dados do ano de 2019 foram disponibilizados. Outra se refere ao fato da obtenção apenas de quantitativos (somatórios) em relação aos serviços disponibilizados, agrupados por hora.

Para estudos futuros sugere-se a utilização de dados de outros anos, aumentando assim o escopo da pesquisa (incluindo mais anos). Paralelamente, poderia se realizar um estudo comparativo entre as denominadas “grandes áreas”, em que se tem maior procura. Além disso, outro estudo interessante seria

comparar os valores investidos em cada área com as demandas pelo serviço, propondo talvez a criação de uma ferramenta para monitorar a gestão do estado. Incluir características, dados, dos solicitantes trariam maior riqueza as informações geradas por intermédio do processo de mineração de dados, bem como seria importante para conhecer o perfil dos demandantes.

Data in support of the state public management: getting to know the “Artificial Intelligence Program of Paraná – PIÁ”

ABSTRACT

The State of Paraná has created an initiative for the use of artificial intelligence called Paraná Artificial Intelligence (PIÁ), responsible for the centralization of services provided by the government. The consultation of information by users, makes the capture of this data a possibility to use it for possible analysis in public management. With this purpose, we proceeded to data mining in the base of the program, which was made available by Celepar - Information Technology and Communication of Paraná. The data refer to 2019, the year in which the platform unifying more than 400 services was launched at the end of June. Microsoft Excel® software was used for data preparation and Weka for mining via the APRIORI algorithm. The research began with the partnership between the institutions Companhia de Informação do Paraná and the Universidade Federal do Paraná, to build the database for the system analysis. From the availability occurs the statistical analysis and data mining steps, using the association algorithm (APRIORI). The analyses revealed that in the year 2019, the afternoon period (12h to 18h exclusive) presents higher demand, requiring that services should be available regardless of the "type of display" (online and/or face-to-face, by phone, online only, face-to-face only), passing on to managers that this time slot should present employees available to serve the population. Future research can increase the scope of the research, more periods of analysis, comparing investments to the agencies with the demand of the service.

KEYWORDS: APRIORI Algorithm. Paraná's Government Site. State Public Management..

NOTAS

¹ HIRATSUKA, E. L. **Solicitação de dados para artigo científico** [mensagem pessoal]. Mensagem recebida por: <maiamarcelomaia@hotmail.com>. 22 jun. 2020.

REFERÊNCIAS

ADRIAANS, P.; ZANTINGE, D. **Data Mining**. England: Longman, 1996.

AGÊNCIA DE NOTÍCIAS DO PARANÁ (AEN). **Paraná inteligência artificial é apresentado em fórum da smart city**. Foz do Iguaçu, 2019. Disponível em: <<http://www.AGÊNCIA.pr.gov.br/modules/noticias/article.php?storyid=104719&tit=Parana-Inteligencia-Artificial-e-apresentado-em-forum-da-Smart-City>>. Acesso em: 14 jun. 2020.

ALMEIDA, D. C. **Descoberta automatizada de associações com o uso do algoritmo como técnica de mineração de dados**. 173 f. Dissertação (Mestrado em Engenharia da Computação) – Universidade Federal de Goiás, Goiânia, 2011. Disponível em: <<https://repositorio.bc.ufg.br/tede/bitstream/tde/966/1/Dissertacao%20Derciley%20Cunha%20de%20Almeida.pdf>>. Acesso em: 28 set. 2021.

BARRETO, J. F. **A impossibilidade de aplicação do CDC (lei nº 8.078/1990) no âmbito dos serviços públicos**. 49f. 2014. Trabalho de Conclusão de Curso (Bacharelado em Direito) – Centro de Ciências Sociais Aplicadas, Universidade Federal do Rio Grande do Norte, Natal. Disponível em: <https://monografias.ufrn.br/jspui/bitstream/123456789/7416/1/JuzianneFB_Monografia.pdf>. Acesso em: 16 jun. 2020.

BRASIL. Constituição (1988). **Constituição**: República Federativa do Brasil. Brasília, DF: Senado Federal, 1988. Disponível em: <http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm>. Acesso em: 23 set. 2020.

BRASIL. Senado Federal. **Manual de Comunicação da Secretaria Especial de Comunicação Social**, 2014. Disponível em: <[https://www12.senado.leg.br/manualdecomunicacao/redacao-e-estilo/estilo/hora#:~:text=A%20madrugada%20vai%20da%20zero,\(ou%20meia%20noite\).>](https://www12.senado.leg.br/manualdecomunicacao/redacao-e-estilo/estilo/hora#:~:text=A%20madrugada%20vai%20da%20zero,(ou%20meia%20noite).>)>. Acesso: 13 jul. 2020.

BRUNI, A. L. **Estatística aplicada à gestão empresarial**. 2. ed. São Paulo: Atlas, 2010.

CASTRO, L. N. de.; FERRARI, D. G.. **Introdução à mineração de dados: conceitos básicos, algoritmos e aplicações**. São Paulo: Saraiva, 2016.

CELEPAR. Soluções para o cidadão. Curitiba, 2019. Disponível em: <<http://www.celepar.pr.gov.br/Pagina/Solucoes-para-cidadao-0>>. Acesso em: 21/06/2020.

CÓBE, R. M.O.; *et al.* Rumo a uma política de Estado para inteligência artificial. **Revista USP**, São Paulo, n. 124, p.37-48, jan./mar. 2020. DOI: <https://www.revistas.usp.br/revusp/article/view/167914>. Disponível em: <<http://www.periodicos.usp.br/revusp/article/view/167914/159995>>. Acesso em: 14/06/2020.

DUARTE, P. S; *et al.* A avaliação da acessibilidade web em portais governamentais: O Paraná Inteligência Artificial (PIÁ). **Revista Brasileira de Administração Científica**, v. 11, n.1, p.53-66, 2000. Disponível em: <<http://sustenere.co/index.php/rbadm/article/view/CBPC2179-684X.2020.001.0005/1881>>. Acesso em: 23/09/2020.

FAYYAD, U; PIATETSKY-SHAPIRO, G; SMYTH, P. The KDD Process for Extracting Useful Knowledge from Volumes of Data. **Communications of the ACM**, v.39, n.11, p.27–34, nov. 1996. Disponível em: <https://scweb.uhcl.edu/boetticher/ML_DataMining/p27-fayyad.pdf>. Acesso em: 14/07/2020.

HOJDA, A.; FARINIUK, T. M. D.; SIMÃO, M. M. B..Building a smartcitywithtrust: the case of '156 central' of Curitiba-Brazil. **Economía, Sociedad y Territorio**, v. XIX, n. 60, p. 79-108, 2019. DOI: <http://dx.doi.org/10.22136/est20191298>. Disponível em: <<https://est.cmq.edu.mx/index.php/est/article/view/1298>>. Acesso em: 6/06/2020.

IBGE. Resolução nº 3, 28 de agosto de 2019. Disponível em: <<http://www.in.gov.br/en/web/dou/-/resolucao-n-3-de-26-de-agosto-de-2019-212912380>>. Acesso em: 6/06/2020.

IBM KNOWLEDGE CENTER. 2012. Disponível em: <https://www.ibm.com/support/knowledgecenter/en/SSEPGG_9.7.0/com.ibm.im.model.doc/c_lift_in_an_association_rule.html>. Acesso em: 15/07/2020.

LAYNES, A. **Paraná lança programa de inteligência artificial para serviços públicos**. 2019. Disponível em: <<http://www.AGÊNCIA.pr.gov.br/modules/debaser/visualizar.php?audiovideo=1&xfid=82650&tit=Governo-lanca-programa-de-inteligencia-artificial-para-servicos-publicos>>. Acesso em: 06/06/2020.

LOBO, L. C. Inteligência artificial, o futuro da medicina e a educação médica. **Revista Brasileira de educação Médica**, Rio de Janeiro, v. 42, n. 3, p. 3-8, 2018. Disponível em: <<https://www.scielo.br/pdf/rbem/v42n3/1981-5271-rbem-42-3-0003.pdf>>. Acesso em: 16/06/2020.

NOVAES, D. V.; QUEIROZ, C. de; COUTINHO, S. **Estatística para Educação profissional e tecnológica**. 2. ed. São Paulo: Atlas, 2013.

OLIVEIRA, R. da C. Gestão pública municipal: um desafio para os gestores. **Revista Científica Multidisciplinar Núcleo do Conhecimento**, v. 1, p. 107 – 116, 2019. Disponível em: <<https://www.nucleodoconhecimento.com.br/administracao/gestao-publica-municipal>>. Acesso em: 23/09/2020.

PAVANELI, A. **Com 11,4 milhões de habitantes, Paraná é o quinto estado mais populoso do país, diz IBGE**. Globo.com., Curitiba, 2019. Disponível:< <https://g1.globo.com/pr/parana/noticia/2019/08/28/com-114-milhoes-de-habitantes-parana-e-o-quinto-estado-mais-populoso-do-pais-diz-ibge.ghtml>>. Acesso em: 6/06/2020.

PROGRAMA INTELIGÊNCIA ARTIFICIAL DO PARANÁ (PIÁ), 2020. Disponível em: <<https://www.pia.pr.gov.br/pagina-39.html>>. Acesso em: 13/06/2020.

SANTOS, M. Y. C. A. **Padrão um sistema de descoberta de conhecimento em base de dados geo-referenciadas**. 2001. 337 f. Tese (Doutorado em Tecnologias e Sistemas de Informação – Escola de Engenharia, Universidade do Minho, Portugal, 2001. Disponível em: < <http://repositorium.sdum.uminho.pt/handle/1822/202>>. Acesso em: 28 set. 2021.

SILVA, L. A. da; PERES, S. M.; BOSCARIOLI, C. **Introdução à análise de dados: com aplicações em R**. 1. ed. Rio de Janeiro: Elsevier, 2016.

SILVEIRA, M. S. **Aplicações de técnicas de inteligência artificial à comunicação alternativa e aumentativa**. 116 f. Dissertação (Mestrado em Ciência da Computação) – Curso de pós-graduação em Ciência da Computação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 1996. Disponível:< <https://www.lume.ufrgs.br/bitstream/handle/10183/23953/000178635.pdf?sequence=1&isAllowed=y>>. Acesso em: 6/06/2020.

SIMÃO, J. B.; RODRIGUES, G. Acessibilidade às informações públicas: uma avaliação do portal de serviços e informações do governo federal. **Ciência da Informação**, Brasília, v. 34, n.2, p. 81-92, maio/ago. 2005. Disponível em: <<https://www.scielo.br/pdf/ci/v34n2/28558.pdf>>. Acesso em: 6/06/2020.

VAZ, L. G. N. **Competências do Gestor Público**. Disponível em: <<http://www.administradores.com.br/producao-academica/competencias-do-gestor-publico/5267/>>. Acesso em: 13/06/2020.

Recebido: 26/09/2020

Aprovado: 19/10/2021

DOI: 10.3895/rts.v18n50.13236

Como citar: MAIA, M. et al. Dados no apoio à gestão pública estadual: conhecendo o "Programa Inteligência Artificial do Paraná – PIÁ". **Rev. Technol. Soc.**, Curitiba, v. 18, n. 50, p.61-79, jan./mar., 2022. Disponível em: <https://periodicos.utfpr.edu.br/rts/article/view/13236>. Acesso em: XXX.

Correspondência:

Direito autoral: Este artigo está licenciado sob os termos da Licença Creative Commons-Atribuição 4.0 Internacional.

