

Uso de técnicas de processamento de linguagem natural para identificação de similaridade de serviços públicos

Sandro Luís Brandão Campos¹, Josiel Maimone de Figueiredo²

¹ Programa de Pós-Graduação em Propriedade Intelectual e Transferência de Tecnologia para a Inovação (PROFNIT)
Universidade Federal de Mato Grosso (UFMT) – Cuiabá, MT – Brasil
Governo do Estado de Mato Grosso

²Instituto de Ciência da Computação
Programa de Pós-Graduação em Propriedade Intelectual e Transferência de Tecnologia para a Inovação (PROFNIT)
Universidade Federal de Mato Grosso (UFMT) – Cuiabá, MT – Brasil

sandrobrandao@gmail.com, josiel@ic.ufmt.br

***Abstract.** The demand and supply of digital public services is spreading globally and has been boosted by the effects of the pandemic. In general, the description of a public service has keywords and specific data that characterize its purpose. This work uses natural language processing techniques on this textual information and performs similarity analysis between public services to generate content-based recommendations for citizen services. The results showed that it is possible to find similarity between the 71 services analyzed that make it possible to personalize content to the citizen, recommending public services similar to the one he usually uses.*

***Resumo.** A procura e a oferta por serviços públicos digitais estão se difundindo globalmente e foram potencializados com os efeitos da pandemia. De forma geral, a descrição de um serviço público possui palavras chaves e dados específicos que caracterizam o seu propósito. Este trabalho usa técnicas de processamento de linguagem natural nessas informações textuais e realiza análise de similaridade entre os serviços públicos para gerar recomendação, baseada em conteúdo, dos serviços ao cidadão. Os resultados demonstraram ser possível encontrar similaridade entre os 71 serviços analisados que possibilita personalizar conteúdo ao cidadão recomendando serviços públicos similares ao que ele geralmente utiliza.*

1. Introdução

Os sistemas digitais de serviços públicos estão se difundindo cada vez mais nas três esferas (municipal, estadual e federal). Os efeitos da pandemia demandaram e aceleraram mais ainda a adoção desse tipo de serviço. Devido à quantidade, variedade e independência dos vários serviços, é muito difícil alcançar a padronização das descrições e conseqüentemente de um catálogo com os vários serviços públicos disponíveis.

O mercado privado utiliza diversas técnicas para melhor conhecer e se relacionar com seu público-alvo [Sohrabi e Khanlari 2007; Gomez-Uribe e Hunt 2016] e a administração pública precisa se atentar de vez para esse modelo estabelecido.

Hoje, maior integração, coerência e a horizontalidade são centrais para a transformação provocada pelo paradigma do governo digital, que abraça a ideia de digitalização total do setor público. É imperativo de governo digital alavancar dados e tecnologias para uma transformação de serviço mais coesa, colaboração transversal e compartilhamento de dados para produzir organizações do setor público proativas, abertas e voltadas para o usuário [OCDE 2000].

O serviço público é uma atividade administrativa ou de prestação direta ou indireta de bens ou serviços à população, exercida por órgão ou entidade da administração pública [Brasil 2017]. É o que transforma os dados em algo de significado para o cidadão.

A demanda de serviço do cidadão moderno é diferenciada no contexto atual, pois hoje já consome, nas diversas plataformas digitais disponíveis, serviços de forma mais cômoda e intuitiva. E a oferta de serviços públicos pode ser elemento transformador na relação com o cidadão, desde que esteja disponível e seja conhecido por quem precisa utilizar.

E a tecnologia da informação, especialmente com as técnicas de inteligência artificial, pode ser considerada uma excelente estratégia para suportar as demandas da sociedade e elemento chave para influenciar a capacidade de um governo eletrônico em entregar valor público, que não poderá mais separar a eficiência de sua atuação dos objetivos de outras políticas no seu ambiente.

Assim, o propósito deste artigo é realizar um experimento para avaliar uma alternativa de oferta de serviços públicos ao cidadão, com base nos serviços públicos similares ao que ele normalmente consome, utilizando uma abordagem de filtragem baseada em conteúdo pela descrição dos serviços públicos disponibilizados, tratando o texto com técnicas de processamento de linguagem natural.

2. Referencial Teórico

Um dos desafios da tecnologia é conseguir interpretar os dados descritivos em linguagem natural sobre os serviços públicos e torná-los decifráveis para a máquina a fim de que seja possível encontrar semelhança entre os serviços. E neste ponto o Processamento de Linguagem Natural (PLN) é elemento essencial para essa tarefa.

O PLN é a subárea da Inteligência Artificial (IA) que estuda a capacidade e as limitações de uma máquina em entender a linguagem dos seres humanos. O objetivo do PLN é fornecer aos computadores a capacidade de entender o sentido de um texto, reconhecimento de voz, e geração de resposta a questões [Beysolow II 2018].

E para modelar a língua e possibilitar o entendimento pela máquina, necessita-se que sejam realizados pré-processamentos que abstraem a complexidade da língua, a fim de que reste apenas o essencial para ser processado. Com o pré-processamento há uma redução no vocabulário para tornar os dados mais simples e higienizados facilitando o processamento pelo computador.

Ainda segundo Beysolow II (2018) o pré-processamento contempla etapas como a tokenização, que visa separar um único corpo de texto para que o computador avalie cada palavra como um elemento individual. Outra operação importante é a eliminação de palavras de parada (*stopwords*), que geralmente são definidas como palavras muito comuns em uma dada linguagem. O processo de stemização (*stemming*) consiste em reduzir uma palavra ao seu radical e a lematização reduz a palavra ao seu lema, que é a forma no masculino e singular e no caso de verbos, o lema é o infinitivo. Há ainda outros filtros como remoção de acentuações, símbolos, numerais e outros caracteres ou palavras indesejadas, onde no idioma português é uma oportunidade de melhoria.

Outras técnicas podem ser aplicadas como para procurar saber quantas vezes uma determinada palavra aparece num texto, como o modelo *Bag-of-Words*. Essa contagem pode ser estruturada pelo conjunto de palavras avaliadas como n-gramas, e quando o $n=2$ sendo um bigrama, que é uma sequência de dois elementos adjacentes de uma sequência de tokens, e com $n=3$, trigrama, que no mesmo conceito do bigrama, é uma sequência de 3 palavras.

Outra técnica importante trata-se de contagem com importância da palavra no texto como a TF-IDF (*Term Frequency Inverse Document Frequency*) [Beysolow II 2018], que é uma medida estatística que tem o intuito de indicar a importância de uma palavra de um documento em relação a uma coleção de documentos ou em um corpus linguístico.

As técnicas de aprendizado de máquina, área da Inteligência Artificial cujo objetivo é a construção de sistemas capazes em adquirir conhecimentos de forma automática [Monard e Baranauskas 2003], fazem parte da grande revolução que proporciona a máquina a pensar muito próxima a um ser humano. Essas técnicas de aprendizado de máquina podem ter seu uso empregado em diferentes áreas, especialmente nos sistemas de recomendação que trata de uma técnica de software que fornece sugestões de informações que sejam úteis para o usuário, quanto a processos de tomada de decisão, como quais produtos comprar, qual música ouvir, a que filme assistir ou quais notícias online para ler (RICCI, 2015). Segundo Aggarwal (2016), os sistemas de recomendação baseados em conteúdo se propõem a combinar itens que são similares com o que o usuário consumiu, com base no histórico seu de itens consumidos. Para Ricci et al. (2015) as medidas de similaridade são ferramentas primordiais nos cálculos de semelhanças entre entidades. A similaridade de cosseno é basicamente aplicada para determinar a semelhança entre as transações e entre itens.

A aplicação de sistemas de recomendações em serviços públicos é um ponto importante na discussão deste experimento. E alguns autores já discutem alternativas importantes relacionadas a esse tema, como em Ayachi et al. (2016) que propõe o uso de sistemas de recomendação baseados em restrições (*Constraint-based recommender systems*), Meo et al. (2007) que considera tanto o perfil do cidadão quanto às características dos dispositivos utilizado por ele. Em Sabucedo et al. (2012), que explorou a descoberta de serviços públicos usando folksonomias. Em Martinguerrero et al. (2016), buscou-se desenvolver um recomendador com filtro de colaboração. Já em Iaquina et al. (2013) apresentou uma abordagem para segmentação de usuários no domínio de governo eletrônico com base em uma técnica de agrupamento parcial que explora dados demográficos e geográficos. Não foi identificada a aplicação de filtragem baseada em conteúdo para recomendação de serviços públicos ao cidadão.

3. Resultados

A base de dados principal utilizada pelo experimento é sobre Serviços Públicos. Essa fonte de dados foi estruturada especificamente para o experimento baseado em conteúdo textual coletado de portais e documentos sobre serviços de instituições públicas no Estado de Mato Grosso. Conforme descreve a Tabela 1, a base de dados possui a relação de 71 (setenta e um) serviços ofertados para o cidadão, bem como suas características básicas, contendo o ID do serviço (*servico_id*), o nome do serviço (*servico_nome*), a descrição sobre o serviço (*servico_descricao*), a área de negócio relacionada (*servico_area*). Os serviços públicos preparados para o experimento estão distribuídos em diversas áreas como: educação, saúde, segurança pública, trânsito, jurídico, meio ambiente, tributário, ouvidoria, aposentadoria, trabalho, social, gestão, cultura, esporte e lazer.

Tabela 1 - Fonte de dados de Serviços Públicos

Coluna	Descrição
servico_id	Identificação do serviço público
servico_nome	Nome do serviço público que aparece para o cidadão
servico_descricao	Descrição com detalhe sobre o propósito do serviço público
servico_area	Área onde o serviço é aplicado dentro do governo

3.1 Tratamento dos dados sobre os serviços públicos

Os dados sobre os serviços públicos, após combinados as características textuais relacionadas aos serviços, foram higienizados da seguinte forma:

- Separação do texto em palavras (Tokenização)
- Transformação do texto em minúscula (Sensibilidade ao caso)
- Remoção de palavras de parada (*Stop Words*)
- Remoção de palavras específicas (Palavras excluídas)
- Radicalização das palavras (Lematização)
- Remoção das palavras curtas (palavras até dois caracteres)
- Remoção de pontuações
- Concatenação das palavras (*Join*) após limpeza

3.2 Comportamento dos dados sobre os serviços públicos

Sobre o campo de descrição dos serviços, *servico_descricao*, foi realizada uma análise antes e após a higienização (limpeza) de seu conteúdo e as informações foram registradas conforme Tabela 2.

Tabela 2. Contagem de palavras do campo *servico_descricao*

Contagem	Antes de higienizar	Após higienizar
Contagem média de palavras:	59,73	36,11
Contagem mínima de palavras:	18	11
Contagem máxima de palavras:	189	110

Observa-se que reduziu bastante a média de palavras no campo descrição de serviços, demonstrando que muitas palavras poderiam ser desconsideradas no experimento, o que torna mais efetivo o resultado do modelo de filtragem, sendo assim a higienização dos dados foi considerada uma ação de suma importância.

A distribuição de contagem de palavras apresentada pela Figura 1, antes de higienizar/limpar os dados, apresentou a quantidade de palavras (*word_count*) do campo descrição de serviços, por intervalos de 0 a 19 palavras resultando em 4 serviços (eixo *count*), de 20 a 39 resultando em 20 serviços, de 60 a 79 palavras em 12 serviços, de 80 a 99 em 7 serviços, de 100 a 119 em 4 serviços e 1 serviço entre 160 e 179 palavras.

Após a limpeza/higienização dos dados, a distribuição de contagem de palavras reafirmou que muitas descrições dos serviços continham palavras desnecessárias para o contexto e que poderiam influenciar negativamente no processo de filtragem. A Figura 2 apresenta esses resultados, agregando a maioria dos serviços em intervalos de 0 a 60 palavras por descrição de serviço.

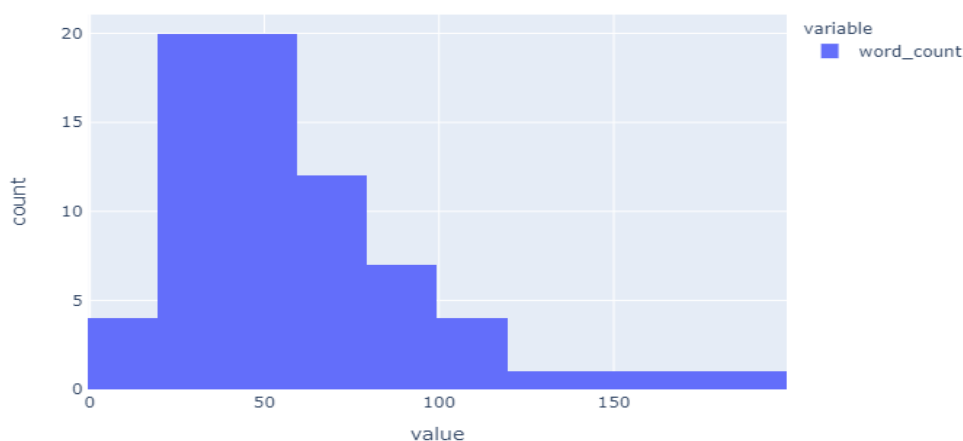


Figura 1. Contagem de palavras antes da limpeza dos dados

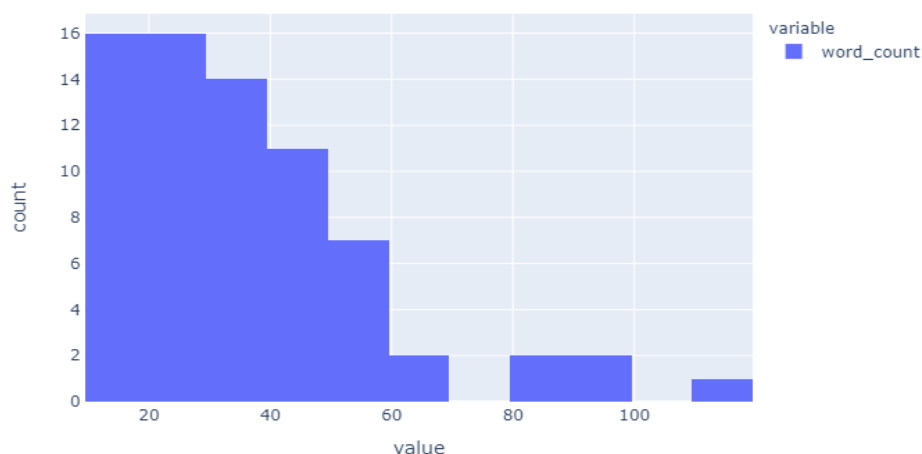


Figura 2. Contagem de palavras após limpeza dos dados

A Figura 3 apresenta o resultado mais aprofundado das principais palavras usadas pelas descrições dos serviços incluindo *stopwords*:

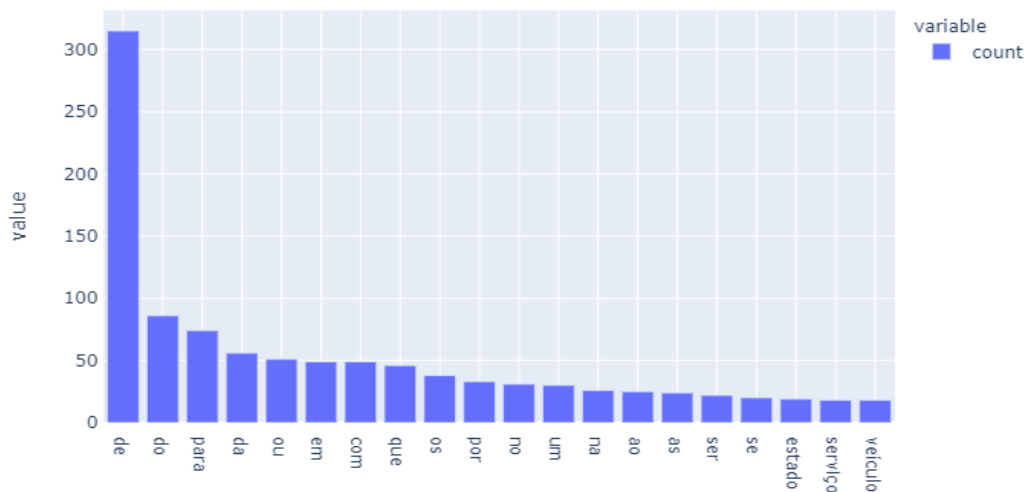


Figura 3. Palavras mais usadas antes da limpeza dos dados

Nota-se a presença de muitas palavras curtas, que embora sejam fortemente usadas para dar sentido ao texto, no contexto de pré-processamento de dados geram ruídos quando analisados. Então, verificando a mesma informação após a limpeza dos dados, percebe-se um resultado mais condizente com o conteúdo a ser tratado pelo modelo, conforme aponta a Figura 4.

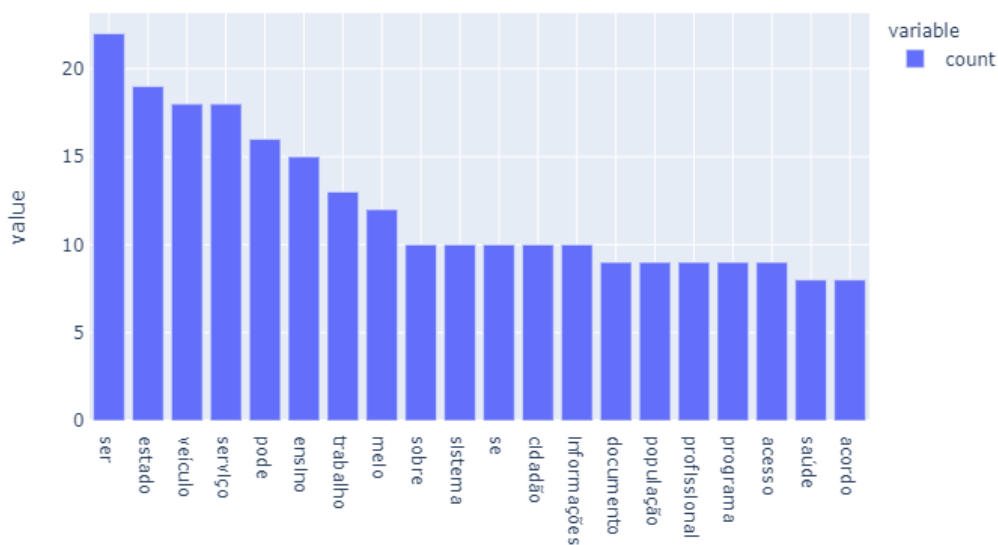


Figura 4. Palavras mais usadas após limpeza

Da mesma forma foram analisados os bigramas antes de higienizar os dados, conforme apresenta a Figura 5 e depois da higienização conforme a Figura 6. Percebe-se que bigramas com *stopwords* também poderiam influenciar nos resultados, reforçando a necessidade da remoção das palavras comuns.

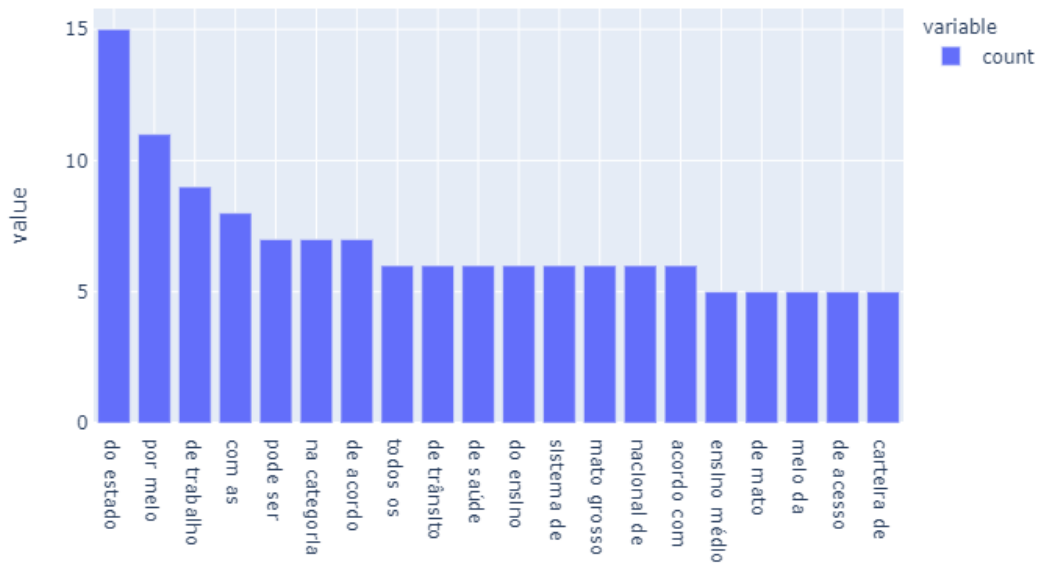


Figura 5. Bigramas mais usados antes da limpeza dos dados

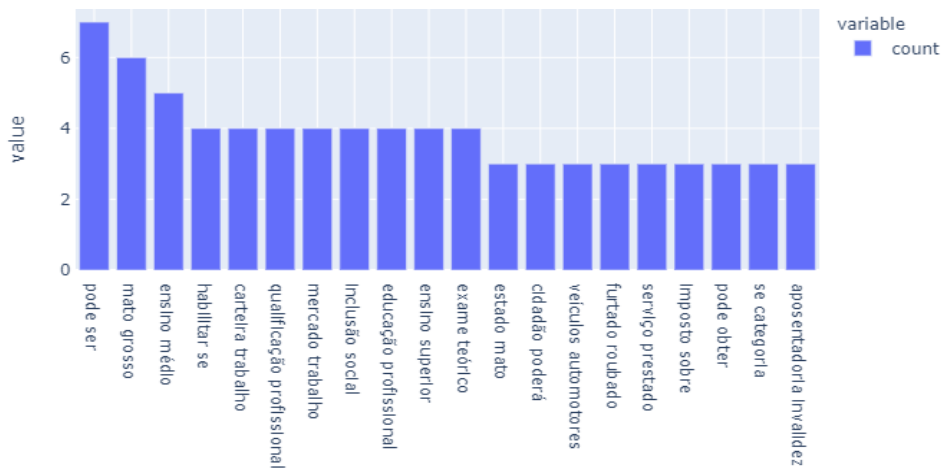


Figura 6. Bigramas mais usados após a limpeza dos dados

A mesma abordagem foi realizada com trigramas. A Figura 7 demonstra o resultado antes de higienizar os dados e na Figura 8 após a higienização.

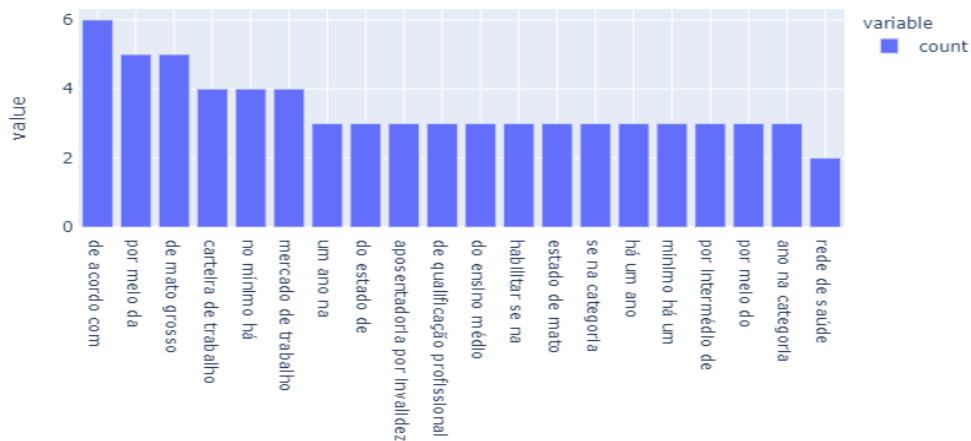


Figura 7. Trigramas mais usados antes de limpeza dos dados

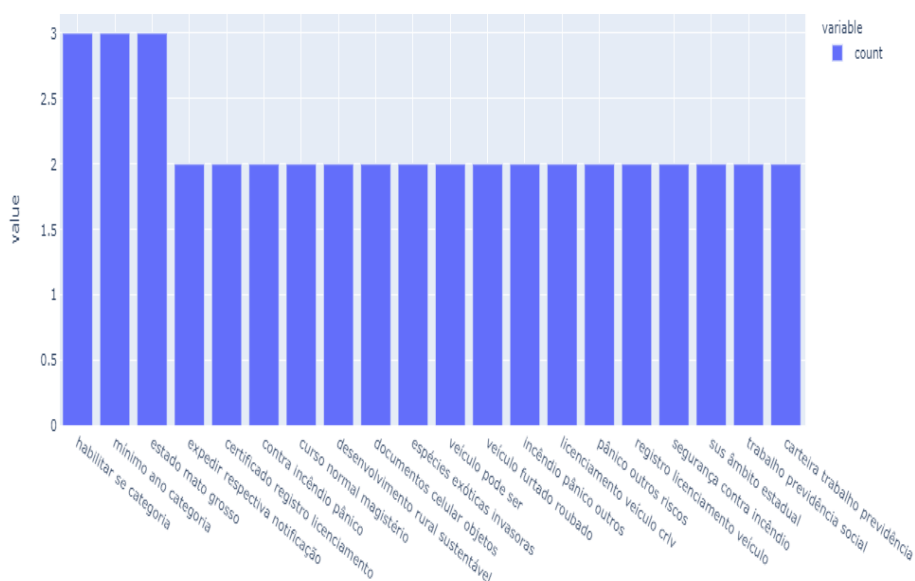


Figura 8. Trigramas mais usados após a limpeza dos dados

3.3 Estruturação dos dados sobre os serviços públicos

Para aproveitar o máximo de informações sobre o serviço, na fonte de dados trabalhada no experimento, foi proposto uma ação de concatenação de atributos com intuito de dar maior relevância para o texto que será utilizado para contagem e análise de similaridade. Foi realizada a concatenação das colunas *SERVICO_NOME*, *SERVICO_DESCRICAO* e *SERVICO_AREA*. Com esta nova estrutura, foram criadas as colunas *FEATURES_COMBINADAS*, contemplando a concatenação sem realizar limpeza dos dados, e a coluna *FEATURES*, com o mesmo valor da coluna *FEATURES_COMBINADAS*, porém com os dados higienizados, conforme Tabela 3.

Tabela 3 - Fonte de dados de Serviços Públicos atualizada

Coluna	Descrição
servico_id	Identificação do serviço público
servico_nome	Nome do serviço público que aparece para o cidadão
servico_descricao	Descrição com detalhe sobre o propósito do serviço público
servico_area	Área onde o serviço é aplicado dentro do governo
features_combinadas	Concatenação dos campos servico_nome, servico_descricao e servico_area
features	Coluna features combinadas após limpeza/higienização

Ou seja, como exemplo para o serviço **VERIFICAR NOTA DE ALUNO**, que está relacionado a área de **EDUCAÇÃO**, e que possui a seguinte descrição: “*Apresentação do boletim escolar ao estudante ou seu responsável legal que foi cadastrado no ato da matrícula. As notas estão disponíveis por ano letivo mantendo o histórico durante o ciclo de estudo do aluno na rede estadual de ensino*”, o texto final inserido na coluna *FEATURES_COMBINADAS* foi: “*VERIFICAR NOTA DE ALUNO EDUCAÇÃO Apresentação do boletim escolar ao estudante ou seu responsável legal que foi cadastrado no ato da matrícula. As notas estão disponíveis por ano letivo mantendo o histórico durante o ciclo de estudo do aluno na rede estadual de ensino*”.

3.4 Uso das técnicas para filtragem baseada em conteúdo

Para a análise baseada em conteúdo foi considerada a coluna concatenada *FEATURES*, pois contempla todo o texto necessário para as análises de similaridade entre os serviços.

Os tratamentos para a higienização realizada sobre a coluna *FEATURES_COMBINADAS* tornaram o texto mais adequados para as etapas seguintes de contagem e similaridades, como podemos observar no texto do serviço **VERIFICAR NOTA DE ALUNO**, conforme Tabela 4, onde na coluna *FEATURES* que comporta o texto tratado, foram realizadas as técnicas como:

- Tokenização: [verificar], [nota], [aluno], [educacao],...
- Transformação em minúsculo: verificar, nota...
- Retirada de stopwords: *de, do, o, a...*
- Remoção de palavras curtas: *ou...*
- Radicalização: *verificar, ser, estar...*
- Remoção de acentuação: *educacao, matricula, apresentacao...*
- Concatenação: *verificar nota aluno educacao...*

Tabela 4. Comparação dos textos antes e após técnicas de PLN

FEATURES_COMBINADAS	FEATURES
VERIFICAR NOTA DE ALUNO EDUCAÇÃO Apresentação do boletim escolar ao estudante ou seu responsável legal que foi cadastrado no ato da matrícula. As notas estão disponíveis por ano letivo mantendo o histórico durante o ciclo de estudo do aluno na rede estadual de ensino	verificar nota aluno educacao apresentacao boletim escolar estudante responsavel legal ser cadastrar ato matricula notas estar disponivel ano letivo manter historico durante ciclo estudo aluno rede estadual ensino

Para validar adequadamente sobre a qual é a melhor métrica a ser utilizada para uso no modelo final, foram realizadas as avaliações com as técnicas para contagem e vetorização das palavras TF-IDF e Contagem simples. E para métricas de similaridade, a Similaridade de Cosseno.

Assim, foram realizados os seguintes passos:

1. Criação de uma matriz de contagem com o número ponderado de ocorrências de cada palavra no texto das características do serviço (FEATURE), sendo considerada a avaliação de palavras, bigramas e trigramas.
2. Definição da semelhança de cosseno entre cada texto sobre os serviços a fim de descobrir como eles são semelhantes entre si.

Para calcular a semelhança de cosseno foi utilizada a contagem de palavras em cada documento, por meio das funções *CountVectorizer* ou o *TfidfVectorizer* da biblioteca *Scikit Learn* na linguagem *Python*. Ambos são métodos para converter dados de texto em vetores, pois o modelo pode processar apenas dados numéricos.

No *CountVectorizer*, apenas o número de vezes que uma palavra aparece no documento é considerado, o que resulta em parcialidade em favor das palavras mais

frequentes, o que acaba ignorando palavras raras que podem ajudar o processamento dos dados com mais eficiência.

No *TfidfVectorizer*, a ponderação geral do documento de uma palavra é destacada e considera a palavra contada de acordo com a frequência com que aparecem nos documentos. Foi considerado o uso do *TfidfVectorizer* em vez do *CountVectorizer*, porque reduz as palavras ponderadas que ocorrem com frequência nos documentos e dá maior significado para os resultados, pois muitos serviços podem ter descrição parecida de acordo com o esforço que o gestor do negócio vai caracterizar os serviços na base de dados. Então, palavras que possam identificar especificamente um serviço podem ter destaque nesta opção.

Com o *TfidfVectorizer* o valor aumenta proporcionalmente para contar, mas é inversamente proporcional à frequência da palavra no corpus; essa é a parte da frequência inversa de documentos (IDF). A frequência inversa do documento se ajusta ao fato de que algumas palavras aparecem com mais frequência em geral.

A similaridade do cosseno foi escolhida, pois é uma métrica usada para determinar a semelhança entre os documentos, independentemente do tamanho, o que também pode variar de acordo com o esforço do gestor em caracterizar o serviço. Quando plotada em um espaço multidimensional, onde cada dimensão corresponde a uma palavra no documento, a semelhança de cosseno captura a orientação (o ângulo) dos documentos e não a magnitude. Se a magnitude fosse o mais importante, o cálculo da similaridade baseada na distância euclidiana seria uma melhor opção.

A Figura 9 apresenta um exemplo de como o tratamento acima analisa o conteúdo, em serviços da área de saúde, para identificação da similaridade entre os textos. Após a contagem, vetorização e cálculo de similaridade um serviço apresenta maior ou menor similaridade com outro serviço de acordo com as palavras mais significativas. Assim, um cidadão que utiliza com frequência o serviço CONSULTA MÉDICA, o modelo proposto poderia, baseado na forte similaridade que existe entre os conteúdos textuais dos serviços, recomendar os serviços TELE CONSULTA MÉDICA e CONSULTA MEDICAMENTOS. Ambos os serviços recomendados podem ser de interesse do cidadão e ele poderia nem saber que são disponibilizados.

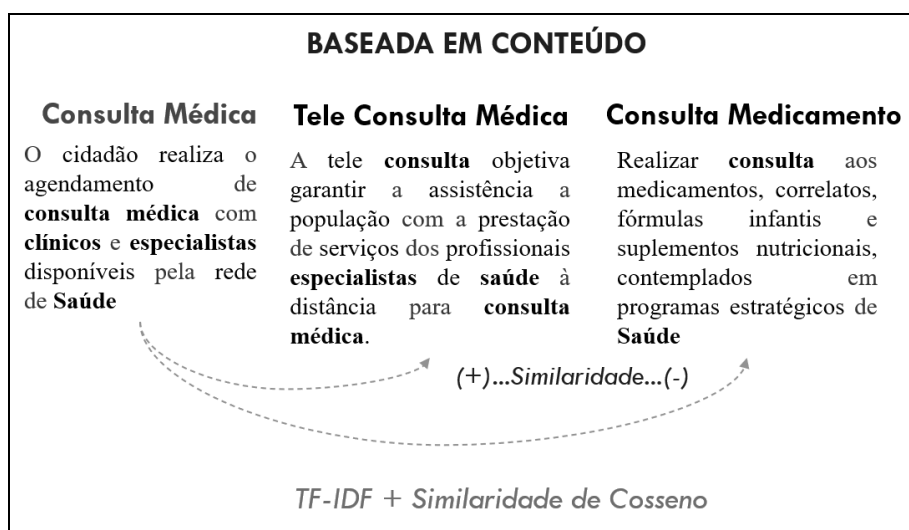


Figura 9. Análise baseada no conteúdo da descrição do serviço público

4. Considerações Finais

Normalmente a administração pública de qualquer esfera possui um grande portfólio de serviços públicos que podem ser ofertados como uma carta de serviços. No entanto, o desconhecimento pelo cidadão dos itens dessa carta de serviço é algo também comum. Assim, a proposta de processamento de texto e cálculo para identificar a similaridade entre os serviços comprovados no experimento é um mecanismo importante para que um serviço possa ser oferecido ao cidadão considerando a relevância com que ele acessa e usa determinados serviços. Ou seja, para um cidadão que acessa serviços de “consulta médica” pode lhe ser ofertado serviço de “tele consulta médica”, ou de “solicitação de medicamento”, ou de “vacinação”, pois o conteúdo descritivo deles apresenta similaridade por conter palavras relacionadas a área de saúde, que com a higienização/limpeza usando técnicas de processamento de linguagem natural foi possível identificar essa similaridade e de forma proativa a administração pública poderá ofertar os demais serviços que podem ser desconhecidos pelo cidadão, mas que poderão ser úteis.

Os experimentos aqui apresentados focam na questão da PLN usada na preparação do processamento, contudo para que a experiência do cidadão não fique fechada apenas ao conteúdo que ele acessa e ficar limitado naquele mesmo padrão de preferências e não conseguir ter acesso a conteúdo novo, causando um efeito bolha [Kelly, 2017], a filtragem baseada em conteúdo poderá ser utilizada de forma híbrida com outros tipos de filtragem, como a colaborativa, demográfica, entre outros.

Como trabalhos futuros pretende-se efetivar o experimento com a incorporação de informações baseadas na Carta de Serviço do Governo de Mato Grosso, contudo, as etapas de processamento de texto e cálculo de similaridade serão as mesmas aqui mostradas, pois o objetivo foi alcançado com a avaliação do experimento demonstrando forte correlação entre os conteúdos textuais que descrevem os serviços públicos e que as técnicas de processamento de linguagem natural são fundamentais para um tratamento adequado desse conteúdo para maior precisão no cálculo de similaridade.

Referências

- Aggarwal, C. C. (2016). *Recommender Systems*. Springer International Publishing.
- Ayachi, R., Boukhris, I., Mellouli, S., Ben Amor, N., & Elouedi, Z. (2016). Proactive and reactive e-government services recommendation. *Universal Access in the Information Society*, 15(4), 681–697.
- Beysolow II, T. (2018). *Applied Natural Language Processing with Python: Implementing Machine Learning and Deep Learning Algorithms for Natural Language Processing*. Apress.
- Brasil. (2017). Dispõe sobre participação, proteção e defesa dos direitos do usuário dos serviços públicos da administração pública, no LEI No 13.460, DE 26 DE JUNHO DE 2017. http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2017/lei/l13460.htm
- Gomez-Uribe, C. A., & Hunt, N. (2016). The Netflix Recommender System: Algorithms, Business Value, and Innovation. *ACM Transactions on Management Information Systems*, 6(4), 1–19.

- Iaquinta, L., Torsello, M. A., Comerio, M., Fanelli, A. M., & Semeraro, G. (2013). *User Segmentation in e-Government Services*. 6.
- Kelly, K. (2017). *Inevitável: As 12 forças tecnológicas que mudarão nosso mundo* (1ª). HSM.
- Martinguerrero, J., Palomares, A., Balaguerballester, E., Soriaolivas, E., Gomezsanichis, J., & Sorianoasensi, A. (2006). Studying the feasibility of a recommender in a citizen web portal based on user modeling and clustering algorithms. *Expert Systems with Applications*, 30(2), 299–312.
- Meo, P. D., Quattrone, G., Terracina, G., & Ursino, D. (2007). *Utilization of intelligent agents for supporting citizens in their access to e-government services*. 39.
- Monard, M. C., & Baranauskas, J. A. (2003). Conceitos sobre Aprendizado de Máquina. In *Sistemas Inteligentes—Fundamentos e Aplicações* (p. 525). Manole.
- OCDE. (2020). *The OECD Digital Government Policy Framework* (OECD Public Governance Policy Papers N° 3; OECD Public Governance Policy Papers, Vol. 3). OECD Publishing.
- Ricci, F., Rokach, L., Shapira, B., & Kantor, P. B. (Orgs.). (2015). *Recommender Systems Handbook*. Springer US.
- Sabucedo, L. A., Barreiros, R. S., Gago, J. M. S., & Iglesias, M. F. (2012). A hybrid Semantic driven recommender for services in the eGovernment domain. *2012 Second International Conference on Digital Information and Communication Technology and It's Applications (DICTAP)*, 409–414.
- Sohrabi, B., & Khanlari, A. (2007). *Customer Lifetime Value (CLV) Measurement Based on RFM Model*. 14(47), 14.