

Editor

César Antônio Pereira

Conflito de interesses

Os autores declaram não haver conflito de interesses.

Recebido

31 ago. 2022

Aprovado

10 abr. 2023

Arquitetura da Informação Multimodal: contribuições no desenvolvimento de Inteligência Artificial¹

Multimodal Information Architecture: contribution on Artificial Intelligence developments

George Hideyuki Kuroki Júnior² , Cláudio Gottschalg-Duque² 

¹ Artigo elaborado a partir da tese de G.H. KUROKI JÚNIOR, intitulada “*Multimodal Information Architecture and Artificial Intelligence: applicability and architectural models*”. Universidade de Brasília, Brasília, 2022.

² Universidade de Brasília, Faculdade de Ciência da Informação, Campus Darcy Ribeiro. Brasília, DF, Brasil. Correspondência para/Correspondence to: G.H. KUROKI JÚNIOR. E-mail: kurokijr@gmail.com.

Como citar este artigo/How to cite this article: Kuroki Júnior, G.H.; Gottschalg-Duque, C. Arquitetura da Informação Multimodal: contribuições no desenvolvimento de Inteligência Artificial. *Transinformação*, v. 35, e226729, 2023. <https://doi.org/10.1590/2318-0889202335e226729>

Resumo

Apresentar contribuições da Arquitetura da Informação Multimodal na organização da informação para treinamento de redes neurais artificiais, visando posicionar a Ciência da Informação como área do conhecimento atuante em problemas de Inteligência Artificial. Utilizaram-se as definições da Arquitetura da Informação Multimodal seguindo a fase tecnológica com abordagem explicativa e qualitativa. Propõe-se procedimento de cinco passos para delineamento, análise e transformação do espaço informacional a ser utilizado em métodos de treinamento e aprendizagem de redes neurais, de forma a complementar lacunas identificadas por autores voltados a implementações da ciência da computação. Verifica-se potencial para desenvolvimento de um método estruturado de Arquitetura da Informação Multimodal que forneça instrumentos para a organização do pré-processamento de dados a serem utilizados como massa de teste e aprendizado em redes neurais artificiais. Esse método posicionaria a ciência da informação como atuante e produtora de soluções de inteligência artificial, substituindo o papel atual de mera consumidora de soluções pré-fabricadas pela Ciência da Computação.

Palavras-chave: Arquitetura da Informação Multimodal. Inteligência Artificial. Técnicas de Organização da Informação.

Abstract

To present Multimodal Information Architecture contributions on information organizing on training artificial neural networks, aiming to position information science as an active body of knowledge in artificial intelligence problems. The definitions of Multimodal Information Architecture were used following the technological phase with an explanatory and qualitative approach. A five-step procedure is proposed for delineating, analyzing and transforming the

informational space to be used in neural network training and learning methods, in order to complement gaps identified by authors focused on computer science implementations. Great potential for developing a structured method of Multimodal Information Architecture was observed, which would provide instruments for organizing data pre-processing which are used as test and learning sample in artificial neural networks. This method could place information science as an actor and producer of artificial intelligence solutions, replacing its current role as consumer of prefabricated solutions made by computer science.

Keywords: *Multimodal Information Architecture. Artificial Intelligence. Information Organizing Techniques.*

Introdução

A crescente utilização de modelos de Inteligência Artificial em atividades cotidianas de classificação e tratamento de informação coloca um novo prisma de observação à questão levantada por Hjørland (2008). Segundo o autor, a organização do conhecimento como área de estudo teria como peças centrais a Ciência da Informação e a Biblioteconomia, todavia sendo seriamente desafiada pela ciência da computação.

Ao tempo em que tal afirmativa foi realizada, uma proposição de arquitetura e implementação de redes neurais artificiais, desenvolvida por Hinton, Osindero e Teh (2006), possibilitou superar um obstáculo histórico enfrentado pela computação, dando passo a mais ao que foi denominado como “o juízo final para a Ciência da Informação e Biblioteconomia”.

Até a proposta de Hinton, Osindero e Teh (2006), a construção de redes neurais artificiais padecia de falta de profundidade em suas implementações: notoriamente, o cérebro humano, base para o desenvolvimento de modelos de inteligência, possui diversas camadas de análise, o que possibilita o tratamento de problemas com maior complexidade. Com o advento da proposta em questão, o número de camadas de tratamento ultrapassou o limite de duas ou três.

O perpassar da limitação computacional deu origem à grande variedade de implementações tecnológicas, originando inumeráveis desenhos arquiteturais de redes neurais que aplicam múltiplos algoritmos matemáticos para se obter uma medida de inteligência por meio da verificação de padrões.

Ainda que haja avanços por parte da ciência da computação, uma crítica feita por Hjørland ainda é passível de discussão:

Existem muitas comunidades separadas que trabalham com diferentes tecnologias, mas muito poucas pesquisas sobre seus pressupostos básicos e méritos e lados fracos. O problema não é apenas formular uma teoria, mas descobrir suposições teóricas em diferentes práticas, formular esses pressupostos de forma tão clara quanto possível, para possibilitar a comparação das abordagens (Hjørland, 2008, p. 87).

Posiciona-se, neste ensaio, a Arquitetura da Informação Multimodal (AIM) como uma contribuição inicial da ciência da informação como uma contrapartida teórica a modelos de Inteligência Artificial (IA), mais especificamente no Processamento de Linguagem Natural (PLN).

Procedimentos Metodológicos

Para analisar, de forma estruturada, os impactos da aplicação da AIM sobre problemas de PLN, propõe-se a utilização do percurso metodológico para construção de uma visão de mundo (M³) criada por van Gigch e Moigne (1989).

Tal proposta considera a construção do conhecimento ao longo de três etapas que guardam íntima relação entre elas: um nível metafísico, anterior à formalização do objeto do conhecimento;

um nível do objeto do conhecimento em si; e um nível da aplicação do conhecimento construído. Nesse sentido, este artigo adaptará tal metodologia da seguinte forma:

- i) No nível metafísico: identificar as questões fundamentais do atual estágio do PLN;
- ii) No nível do objeto do conhecimento: propor formas de aplicação da AIM em problemas de PLN;
- iii) No nível da aplicação do conhecimento: gerar produtos de AIM para implementação em PLN.

Deep learning: aplicações, desenvolvimento e desafios em processamento de linguagem natural

Os ditames fundamentais para a construção de redes neurais artificiais foram sedimentados ao longo das décadas de 60 a 90. Com a entrada dos anos 2.000 e a proposta de Hinton, Osindero e Teh (2006), uma nova gama de implementações passou a se valer da profundidade de camadas de análise, dando origem ao termo *deep learning*.

Wason (2018) realiza levantamento sobre a utilização das descobertas realizadas por Hinton, Osindero e Teh (2006), verificando sua utilização de forma ampla em variada gama de domínios, como reconhecimento de voz independente da fonte sonora, redes neurais recorrentes, reconhecimento de caligrafia, redes de crença profundas, auto-encodificadores, modelagem acústica, detectores de características de classes, síntese de caligrafia, modelagem de linguagens, melhoria e desenvolvimento de modelos, entre outros. Conclui-se que três grandes desafios ainda perduram na maioria das aplicações de IA:

- 1) Volume de dados: a massa de dados necessária para se obter aprendizado satisfatório seria da natureza de dez vezes a quantidade de parâmetros (neurônios) da rede desenhada;
- 2) Fenômeno de *overfitting*: quanto maior o tamanho da rede, em termos de número de parâmetros, maior a probabilidade de que o aprendizado esteja superdimensionado, resultando em uma baixa capacidade de generalização (mudanças pequenas nos objetos de entrada resultam em um resultado insatisfatório);
- 3) Natureza frágil: redes neurais tendem a serem especializadas, de forma que, ao serem treinadas em uma determinada tarefa, seu desempenho em outra tarefa é extremamente insatisfatório.

Da junção dos dois primeiros desafios citados por Wason (2018), ainda se identifica problema anteriormente mapeado por Arel, Rose e Kanowski (2010) denominado de “Problema da Dimensionalidade dos Dados”, no qual a complexidade de aprendizado cresce de forma exponencial em detrimento do aumento linear do número de dimensões dos dados.

Segundo Minaee *et al.* (2021), as mais recentes tentativas de obtenção de melhores resultados em PLN baseiam-se em transformadores e Modelos Pré-Treinados (MPT). Desde as primeiras implementações de redes neurais para PLN, como redes convolucionais, redes recorrentes e redes LSTM (*Long Short-Term Memories*, ou memórias longas de curto-período), percebe-se a dificuldade em capturar as relações entre palavras dentro de uma frase. Com o advento de modelos baseados em Mecanismos de Atenção proposto inicialmente por Bahdanau, Cho e Bengio (2014), redes neurais passaram a tratar diversos objetos de forma agrupada. Com base neste avanço, Vaswani *et al.* (2017) propuseram uma nova arquitetura denominada “transformadores”, que trouxe duas inovações relevantes: atribuição de uma pontuação de atenção que avalia a influência de uma palavra sobre outra; e melhoria nos métodos de paralelização, reduzindo o tempo de treinamento. A partir de 2018, observa-se um crescimento em MPT baseados em transformadores, dotados de

arquiteturas mais densas e pré-treinados em grandes volumes de dados textuais, o que, de forma conjunta, acarreta melhor contextualização de palavras e sentenças. Qiu *et al.* (2020) realizaram um levantamento sobre os MPT mais utilizados, classificando-os por meio de quatro categorias:

1) Tipo de representação: forma de representação do idioma, visando à identificação de regramentos linguísticos implícitos e conhecimento de senso comum que não são explícitos em dados textuais;

2) Modelo arquitetural: modo de captura dos contextos, se de forma sequenciada (palavra após palavra) ou não sequenciada (utilizando uma estrutura sintática ou semântica pré-definida);

3) Tipo de tarefa de pré-treinamento: objetivo pretendido ao longo do treinamento. Em aprendizado supervisionado, busca-se uma função capaz de mapear pares de entrada e saída; em aprendizado não supervisionado, busca-se obter conhecimento intrínseco a partir de dados não classificados; em aprendizado auto-supervisionado, há a junção dos tipos anteriores, em que o método de treinamento é baseado em aprendizado supervisionado, mas a classificação dos dados é gerada de forma automática.

4) Extensões ao modelo: MPT geralmente visam representações universais de um idioma para aplicações genéricas. Para aplicações específicas, maior enriquecimento do modelo é desejável como multi-idioma, multimodal ou específico de um domínio ou tarefa.

Qiu *et al.* (2020) também dividem os MPT em duas gerações, conforme seus objetivos. A primeira geração busca bons modelos de mapeamento de palavras, obtendo classificação hierárquica de palavras em detrimento de um modelo da linguagem. São independentes do contexto. Word2vec de Mikolov *et al.* (2013a), GloVe de Pennington, Socher e Manning (2014), assim como CBow e *Continuous Skip-Gram* de Mikolov *et al.* (2013b), são exemplos. A segunda geração busca produzir vetores de palavras a nível de frases, levando em consideração o contexto em que as palavras se encontram. CoVe de McCann *et al.* (2017), ELMo de Peters *et al.* (2018), *OpenAI GPT* de Radford *et al.* (2018) e BERT de Devlin *et al.* (2018) são exemplos.

Dada a amplitude de modelos disponíveis, Minaee *et al.* (2021) propõem um procedimento de cinco passos para a escolha de uma rede neural de PLN:

- 1) Seleção do MPT;
- 2) adaptação ao domínio do problema;
- 3) inserção de camada adaptada à tarefa;
- 4) ajuste de pesos à tarefa;
- 5) compressão do modelo.

Decorrida a análise de mais de 150 modelos voltados a PLN utilizando mais de 40 conjuntos de dados, os autores concluem que, por mais que grandes avanços tenham sido obtidos, algumas questões permanecem desafiadoras:

1) Ausência de dados para tarefas mais complexas: embora a quantidade de dados coletados ao longo dos anos seja expressiva, há tarefas, como perguntas e respostas com raciocínio de múltiplos passos, classificação de textos para documentos com múltiplos idiomas e classificação de texto para documentos longos;

2) Modelos de conhecimentos de senso comum: a falta de modelos com conhecimentos de senso comum limita a capacidade de análise de redes neurais, como responder a perguntas sobre o mundo real ou lidar com a incompletude de informações;

3) Modelos com uso eficiente de memória: a maioria dos modelos modernos requer grande quantidade de memória, o que leva à necessidade de compressão dos mesmos;

4) Aprendizado com menor esforço: a maioria dos modelos de *deep learning* são treinados por meio de aprendizado supervisionado. Na prática, coletar e classificar dados para um novo domínio é uma tarefa complexa e desafiadora.

Arquitetura da Informação Multimodal: contribuições para o desenvolvimento de *deep learning*

Segundo Kuroki Júnior (2018), define-se a AIM como a construção e distinção de mundos arquiteturais, por meio de suposição de modelos relacionais agrupados por contextos espaço-tempo de estados de informação correlacionados ou não.

As arquiteturas produzidas com tal direcionamento podem ser aplicadas a diversos modos de significação das coisas, ou seja, uma forma multimodal de interpretação. O modo segue a definição dada por Kress e Van Leeuwen (2001) e Kress (2009) como qualquer recurso socialmente e culturalmente moldado para se construir significados. Para os autores, qualquer modo, incluindo a língua (na concepção de idioma escrito e falado e suas possibilidades), possui limitações e potencialidades.

A AIM poderia, em tese, produzir visões ou agrupamentos de dados que possam expressar, de forma mais eficaz, um domínio ou um contexto de informação para facilitar (ou atenuar as medidas de erro e/ou incompletude do conjunto de dados de teste) o reconhecimento de padrões por meio de redes neurais. Na AIM, um mundo arquitetural é um contexto de relações entre sujeitos e objetos, ou seja, um domínio semântico.

De forma resumida, os passos propostos para obtenção de uma arquitetura da informação para PLN são:

- 1) Identificação de entidades de contexto;
- 2) identificação de correlações entre entes;
- 3) distinção de domínios;
- 4) proposição de relações entre domínios;
- 5) agrupamento por contextos espaço-tempo.

Identificação de entidades de contexto

Para PLN e MPT, um contexto pode ser visto tão somente como um grupo de textos agrupados por semelhança linguística, semântica, factual, senso comum ou qualquer outra característica. Tal assertiva não se aplica para a AIM. Um contexto só se torna um espaço arquitetural, quando é considerado o ponto de vista de um sujeito de ao menos um objeto. Em contrapartida, um objeto pode ser classificado de forma diferente por múltiplos sujeitos, assim como determinada sequência de texto pode exprimir significados distintos em contextos distintos. Redes neurais de PLN visam sobrepor essa barreira por meio de volume de dados, o que, conforme Minaee *et al.* (2021), é restrito para tarefas mais complexas. Nesse sentido, a primeira intervenção da AIM visa definir os sujeitos e objetos de um contexto, sendo:

- a) “Sujeito” uma entidade dotada de capacidade de produzir e manipular informação;
- b) “Objeto” uma entidade com potencial de significação, dotada de atributos que possam ser interpretados por sujeitos de forma comum;
- c) Uma “Correlação” quando um sujeito transforma um objeto por meio de “definição”, “comparação”, “fusão” ou “decomposição”, e o produto dessa operação é aceito dentro do corpo de conhecimento compartilhado pelos sujeitos que compõem o contexto.

Identificação de correlações entre entes

Para Kuroki Júnior (2018), as relações conectam instâncias de um contexto ou os próprios contextos em si. Uma correlação se forma entre um sujeito e um objeto por meio de uma relação em um determinado contexto. No contexto de PLN, à luz da AIM, as relações fundamentais propostas são quatro:

- a) “Definição” é uma correlação realizada por um sujeito que transforma o estado de um ente em um contexto para objeto, abrindo a possibilidade de agregar outros entes como atributos.
- b) “Comparação” só é aplicável a objetos definidos por um sujeito. Qualquer nível de comparação se dá por meio de análise dos atributos assinalados a objetos distintos.
- c) “Fusão” é a junção de dois objetos para a formação de um terceiro.
- d) “Decomposição” é a operação oposta à fusão, em que um objeto dá origem a outros dois distintos.

Distinção de domínios

Em uma AIM aplicada a PLN, um “Domínio” é um grupo de atributos de objetos que podem ser identificados, de forma comum, por diferentes sujeitos por meio correlações similares. Dessa forma, sujeitos e objetos podem compor diversos domínios. Três formas possíveis de estabelecimento de domínios são:

- 1) Descrição: partindo de um conjunto de atributos potenciais, verifica-se o acolhimento semântico desses por sujeitos, para então encontrar tais atributos em determinados objetos, agrupando os mesmos;
- 2) Inspeção: analisando um conjunto de objetos, agrupam-se os mesmos por atributos comuns, e verifica-se o reconhecimento comum em um determinado grupo de sujeitos;
- 3) Verificação: inquirindo determinado grupo de sujeitos, identificam-se atributos percebidos de forma comum pelos indivíduos desse grupo, e agrupam-se objetos que contenham esses atributos.

Proposição de relações entre domínios

As três primeiras operações visam identificar entes e domínios de um modelo. Tratam-se essas definições como o lado informacional do problema. As relações entre esses domínios dão o caráter arquitetural da proposta, no sentido de uma imposição de ordem. Para que uma AIM impacte de alguma forma em um contexto ou até um domínio, alguma alteração nesse espaço informacional deve ser realizada. Isso se dá por meio de relações entre domínios.

- 1) Para Kuroki Júnior (2018), relações são dotadas de regras que as restringem. A proposta original da AIM utiliza lógica modal para expressar relações. Três relações básicas de manipulação de domínio são propostas para alterar esse domínio ou produzir um novo:
 - 2) Identidade: uma relação de identidade é obtida quando todos os objetos de um domínio podem ser encontrados em um outro domínio. Corresponde ao operador modal de necessidade;
 - 3) Proximidade: uma relação de proximidade é identificada quando parte dos objetos de um domínio pode ser encontrado em um outro domínio. Corresponde ao operador modal de possibilidade;
 - 4) Incidental: relações incidentais nem sempre são perceptíveis, com certa medida de aleatoriedade em suas incidências. A forma mais simples de defini-las seria como uma relação de segunda ordem.

Quando ocorre a extensão das relações, o autor utiliza estruturas modais lógicas citadas por Carnielli e Pizzi (2008):

- a) Reflexiva: uma estrutura reflexiva é identificada quando uma relação proposta é aplicável de um domínio para ele mesmo;
- b) Serial: uma estrutura serial é identificada quando uma relação proposta é aplicável de um domínio para outro;
- c) Simétrica: uma estrutura simétrica é identificada quando uma relação proposta é aplicável mutuamente entre dois domínios;
- d) Transitiva: uma estrutura transitiva é identificada quando, supondo três domínios [A, B, C]; caso A tenha a relação proposta com B, e B possui a relação proposta com C, então A possui a relação proposta com C;
- e) Euclidiana: uma estrutura euclidiana é identificada quando uma relação proposta é reflexiva, simétrica e transitiva.

Agrupamento por contextos espaço-tempo

Aplicar todo o regramento possível a um domínio ou conjunto de domínios não é o objetivo da AIM. Uma medida de economia das relações deve ser levada em consideração; caso contrário, qualquer configuração tenderia a mapear a realidade objetiva da forma mais próxima possível. Para Kuroki Júnior (2018), distinções espaço-tempo podem ser identificadas por meio de estruturas deônticas, que exprimem uma lógica de obrigações e permissões. Distinguem-se essas de estruturas epistêmicas, as quais tratam de conhecimento. A principal reside na impossibilidade de estruturas deônticas assumirem uma verdade imutável: tão somente consideram a possibilidade de uma ocorrência. Um exemplo simples citado por Portner (2009) seria o regramento moral “não assassinar”. Ainda que esse seja elencado como necessário (deve existir em todos os contextos possíveis), assassinatos ocorrem ainda assim.

Todos os regramentos listados até esse ponto endereçaram questões espaciais de uma arquitetura da informação: o quão abrangente um modelo é no tocante as relações, objetos e atributos que considera. A questão temporal torna-se, de fato, um limitador para qualquer modelo estático, o que leva à necessidade de um modelo cíclico, conforme ilustrado a seguir.

Implementação de uma Arquitetura da Informação Multimodal

Seguindo o percurso metodológico proposto, sugere-se uma aplicação de AIM em um problema de PLN a título de exemplificação. A situação selecionada refere-se à classificação de textos. A dificuldade reside tanto na ausência de dados suficientes para aprendizado quanto na abrangência semântica desses dados. Em resumo, trata-se de uma análise de tendência positiva ou negativa de um conjunto de texto segundo uma legislação de incentivos a pesquisa, desenvolvimento e inovação. Anualmente, são submetidos mais de 10.000 textos que podem ser classificados em 16 categorias de conhecimento. Até o presente momento, tão somente os dados de 2014 e 2015 encontram-se classificados.

Aplicação de PLN em domínio não tratado por AIM

Os textos classificados nos anos de 2014 e 2015 foram submetidos a treinamento, validação e teste em uma rede neural de classificação de textos. Para tal tarefa, foi utilizado o modelo

BERTimbau de Souza, Nogueira e Lotufo (2020), treinado por meio do *corpus* brWaC, de Filho *et al.* (2018), o qual possui 3.5 milhões de documentos e 2.68 bilhões de tokens. O modelo utilizado separa os dados em três partes: treinamento, validação e teste. Para cada conjunto, duas variáveis são observadas. Perda (*loss*) representa a diferença entre os resultados esperados e os resultados obtidos pela máquina. Por meio da perda, obtêm-se os ajustes dos pesos da rede neural, o que possibilita o avanço no aprendizado ao longo do experimento. Menores valores de perda indicam melhor aprendizado da rede. Acurácia (*acc*) representa o percentual de acertos obtidos em cada etapa do experimento. Essa variável apresenta o quão assertivo é o modelo com base nos dados apresentados.

Para isolar os produtos da AIM de qualquer interferência advinda de técnicas vinculadas à ciência da computação (enriquecimento da base de dados, mudança no algoritmo de aprendizado, aumento do escopo de análise), nenhum procedimento de melhoria será aplicado tanto ao ambiente quanto ao conjunto de dados originais, o que garante que qualquer resultado esteja vinculado única e exclusivamente à AIM.



Figura 1 – Ciclo de construção de Arquitetura da Informação Multimodal.

Fonte: Elaborada pelos autores (2022).

Inicialmente, foram realizados 10 experimentos com 20 ciclos de treinamento cada para os dados de 2014 e 2015 analisados de forma conjunta, com os seguintes resultados médios:

- *Loss* (perda) em treinamento: 0,6592374.
- Acurácia em treinamento: 63,10%.
- *Loss* (perda) em validação: 0,6829368.
- Acurácia em validação: 58,39%.
- *Loss* (perda) em teste: 0,6336318.
- Acurácia em teste: 62,05%.

A assertividade decaiu drasticamente quando utilizados os dados dos anos de 2014 e 2015, de forma separada. Para o ano de 2014, o resultado médio é de 53,64% de acurácia em treinamento, 53,56% em validação e 55,72% em teste. Para o ano de 2015, os números são melhores, com 76,86% em treinamento, 74,39% em validação e 77,21% em teste.

Identificação de entidades de contexto

O primeiro passo para transformar o ambiente informacional em questão é a identificação de entidades de cada contexto original. Os sujeitos ativos na configuração inicial analisam textos submetidos em 16 áreas do conhecimento. Como a classificação desses é dada por meio de vários indivíduos (pessoas naturais), aplicando-se a AIM de Kuroki Júnior (2018), o conjunto de conhecimentos expressos em cada área pode ser considerado um sujeito, obtendo-se, por conseguinte, 16 sujeitos.

De forma reflexa, o *corpus* de objetos também é definido por essa distinção de sujeitos, dado que há um acordo semântico entre as pessoas que analisam os textos em cada área. A diferença reside no fato de que cada área de conhecimento possui duas partições de valor binário – aprovado ou reprovado –, dotado de 3 agrupamentos semânticos – elemento inovador, barreira tecnológica e método –, totalizando 96 contextos semânticos. Nesse sentido, dado que objetos são expressos por meio de atributos, tão somente substantivos são elegíveis como entidades, devido à sua capacidade de absorção de atributos por meio de outros termos semânticos que os modifica. A Figura 2 demonstra os quantitativos obtidos por contexto para o ano de 2015.

Identificação de correlações entre entes

A segunda fase para produção de uma AIM é a identificação de correlações entre os sujeitos e objetos do domínio. Nesse sentido, uma técnica denominada Frequência Inversa em Documentos (FID), originariamente proposta por Jones (1973), foi utilizada. Trata-se de uma medida logarítmica de relevância de um termo perante um conjunto de documentos: quanto menor a incidência de determinada palavra em um texto, maior é a probabilidade de sua relevância. O procedimento de seleção de entidades deve identificar aquelas que sejam relevantes ao modelo como um todo, mantendo-se a relação de relevância do ente potencial frente ao contexto de origem. Nesse sentido, 5 etapas de análise são propostas:

- 1) Cálculo do FID de cada ente perante cada um dos 96 domínios semânticos;
- 2) Obtenção da média FID de cada ente, considerando a soma dos valores dos 96 domínios semânticos;

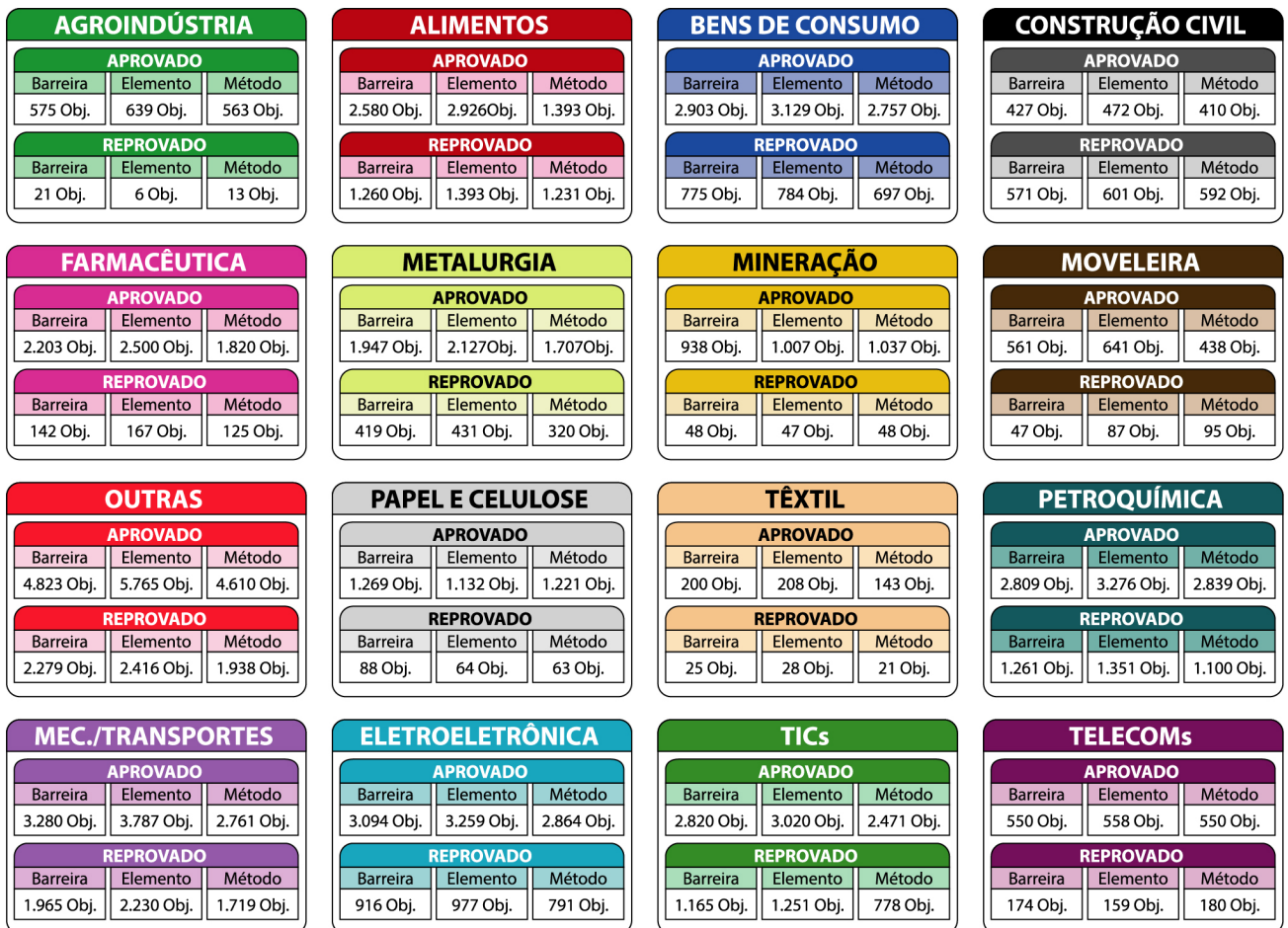


Figura 2 – Objetos identificados por contexto – ano base: 2015.

Fonte: Elaborada pelos autores (2022).

3) Seleção dos entes cuja média FID (calculado no item 2 acima) seja maior que o desvio padrão, considerando todas as médias FID;

4) Identificação de objetos por meio de definição, comparação, fusão e decomposição.

Para o ano de 2015, foram identificadas 21.142 entidades potenciais. Ao aplicar o sequenciamento dos passos 1 a 3, este número decai para 513. Entre as entidades potenciais, foram identificados os conjuntos semânticos [método, metodologia], [fabricação, produção], [necessário, necessidade], [produtivo, produtividade], [fim, final, resultado], [sistema, *software*]. Os atributos de tais pares foram analisados por meio de COMPARAÇÃO, a fim de verificar a necessidade de DEFINIÇÃO de dois termos ou de FUSÃO em apenas um termo. Os resultados das relações potenciais em questão foram:

- Conjunto semântico [método, metodologia]: percentual de semelhança entre atributos de 1,69%. Relação de DEFINIÇÃO;
- Conjunto semântico [fabricação, produção]: percentual de semelhança entre atributos de 1,85%. Relação de DEFINIÇÃO;
- Conjunto semântico [necessário, necessidade]: percentual de semelhança entre atributos de 5,26%. Relação de DEFINIÇÃO;

- Conjunto semântico [produtivo, produtividade]: percentual de semelhança entre atributos de 8,47%. Relação de DEFINIÇÃO;
- Conjunto semântico [fim, final, resultado]: percentual de semelhança entre atributos de 4,81%. Relação de DEFINIÇÃO;
- Conjunto semântico [sistema, *software*]: percentual de semelhança entre atributos de 1,96%. Relação de DEFINIÇÃO.

Destarte, as 513 entidades potenciais obtidas ao longo das quatro etapas de seleção são reconhecidas e correlacionadas como objetos do domínio.

Distinção de domínios

Identificados os 16 sujeitos e 513 objetos atuantes no domínio original, passa-se a alteração da configuração desse espaço informacional por meio de descrição, inspeção ou verificação. Dado que o percurso para se obter essa configuração iniciou da análise de um conjunto de textos por pessoas naturais, o procedimento de verificação torna-se a escolha mais assertiva para a distinção dos domínios. O procedimento é dotado de 3 passos:

- 1) Inquirir um grupo de sujeitos;
- 2) Identificar atributos comuns;
- 3) Agrupar objetos que possuam tais atributos.

A primeira etapa foi realizada de forma prévia à aplicação da AIM, quando ocorreu a análise dos textos por pessoas naturais, ou seja, foi realizada ao se obter o conjunto de dados original classificado por área de conhecimento e aprovação/reprovação de cada texto individualmente. A segunda etapa foi realizada no item

Identificação de correlações entre entes, onde foram obtidos os 513 objetos reconhecidos pelos 16 sujeitos do contexto inicial. Para a realização da terceira etapa, quatro procedimentos foram realizados:

a) Cálculo da relevância dos objetos para cada um dos 16 sujeitos: cada área possui duas classificações de mérito (aprovado ou reprovado) para três contextos semânticos (elemento inovador, barreira tecnológica e método), totalizando seis parâmetros de análise. Somam-se os valores da FID de cada objeto nos seis parâmetros de análise, obtendo-se o valor de relevância do objeto para cada um dos 16 sujeitos. Esse valor representa o quão relevante é cada objeto para os sujeitos;

b) Índice de adesão do sujeito ao ambiente: dotado do valor de relevância dos objetos, a soma desses valores representa o quão aderente está o escopo de conhecimento do sujeito ao contexto analisado;

c) Obtenção do índice de dispersão do contexto informacional: por meio do cálculo do desvio padrão dos índices de adesão calculados no procedimento anterior, verifica-se o quão uniforme é o ambiente informacional.

d) Concepção de domínios com base no índice de dispersão do ambiente informacional: quanto maior o índice de dispersão, maior a quantidade de agrupamentos, observando-se a necessidade de compensação entre os índices de adesão dos sujeitos ao ambiente.

O índice de dispersão calculado com base no passo “c” para o ano de 2015 foi de 562,38, o que divide o espectro de valores dos índices de adesão em 4 faixas:

- 1) 0 a 562,38: composto pelos sujeitos metalurgia, farmacêutica, papel e celulose, mineração, movelaria, construção civil, agroindústria, telecomunicações e têxtil;
- 2) 562,39 a 1.124,76: composto pelos sujeitos petroquímica, bens de consumo, tecnologias da informação e comunicação (TICs), alimentos e eletroeletrônica;
- 3) 1.124,77 a 1.687,14: composto pelo sujeito mecânica e transporte;
- 4) 1.687,15 a 2.249,52: composto pelo sujeito outros.

O menor nível de distinção/agregamento possível no referido contexto informacional, defeso considerar a totalidade dos 16 sujeitos, é a divisão em dois domínios. Tal divisão deve considerar um equilíbrio no índice de adesão do sujeito ao contexto informacional. Nesse sentido, os agrupamentos [1, 4] e [2, 3] apresentam-se como os mais equilibrados, dando origem a:

- Domínio potencial 1, composto pelos sujeitos metalurgia, farmacêutica, papel e celulose, mineração, movelaria, construção civil, agroindústria, telecomunicações, têxtil e outros;
- Domínio potencial 2, composto pelos sujeitos petroquímica, bens de consumo, TICs, alimentos, eletroeletrônica, e mecânica e transporte.

Relações entre domínios

Dotados de dois domínios potenciais encontrados na etapa anterior, passa-se ao estabelecimento de relações entre as áreas de conhecimento e esses domínios, bem como entre os domínios em si. Nesse sentido, a Figura 3 demonstra as relações de identidade e proximidade que originaram os domínios potenciais, bem como a extensão das relações entre esses domínios.

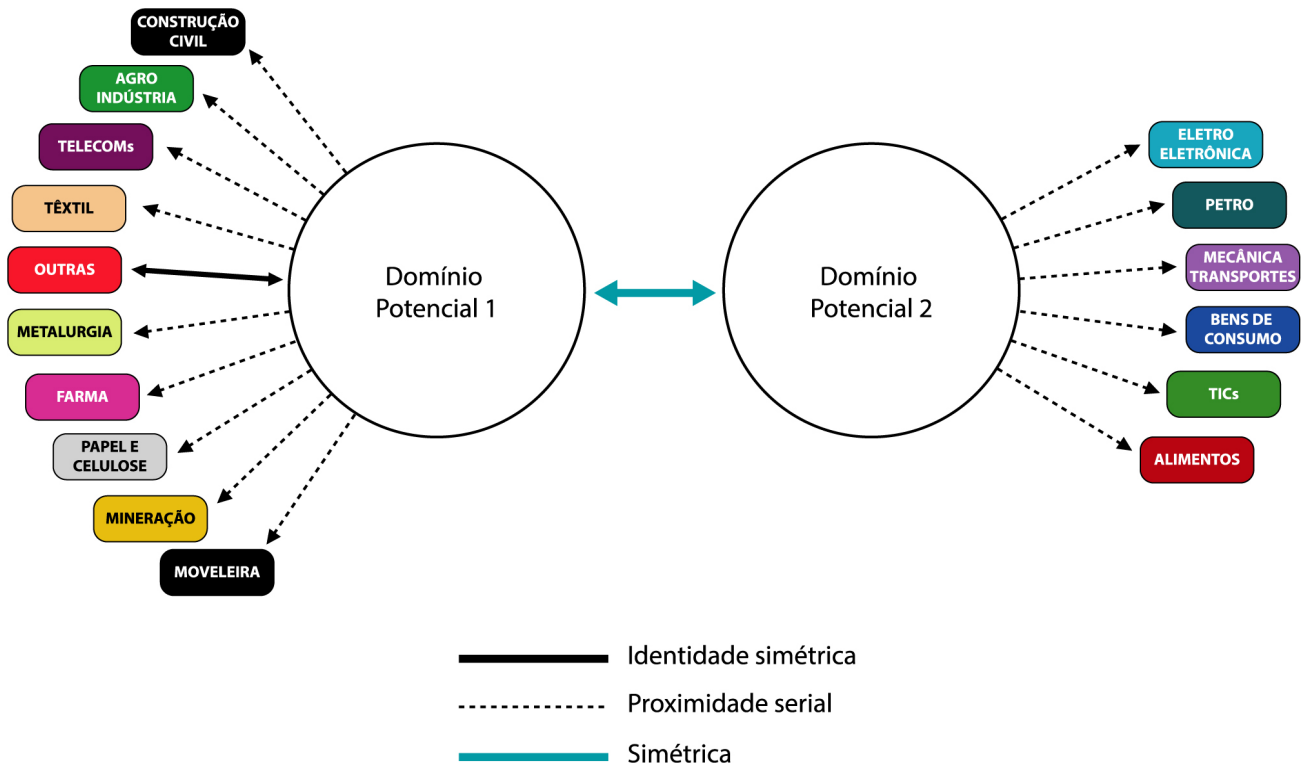


Figura 3 – Relações entre áreas de conhecimento e domínios potenciais, 2015.
 Fonte: Elaborada pelos autores (2022).

Observa-se que, em sua formação, tão somente o domínio potencial 1 possui uma relação de identidade simétrica em sua formação, dado que a área de conhecimento “outros” é a única que possui todos os objetos presente no domínio. Todas as relações identificadas para formação dos domínios potenciais 1 e 2 são reflexivas, uma vez que tal operação parte da identificação de objetos comuns, o que, necessariamente, requer a verificação de existência desse objeto no próprio domínio para só então proceder à verificação de existência do referido objeto em outro domínio.

No tocante às relações entre os domínios potenciais construídos 1 e 2, verifica-se uma única relação simétrica [1,2], dado que todos os objetos podem ser encontrados em qualquer configuração possível de ambos os domínios, o que demonstra que ambos coexistem de forma independente, sendo micro-organizações do contexto informacional original.

Agrupamento por contextos espaço-tempo

Conforme descrito na Identificação de entidades de contexto, foram utilizados os dados do ano base de 2015 para se conceber a distribuição de domínios obtida na Distinção de domínios. De forma a verificar extensão temporal da alteração da arquitetura proposta ao longo dos anos, realizou-se o ciclo da AIM exposto na Figura 1 – Ciclo de construção de Arquitetura da Informação Multimodal, juntamente com os procedimentos descritos na Identificação de entidades de contexto até as Relações entre domínios para o ano base de 2014, obtendo-se configuração distinta de domínios.

Para o passo de

Identificação de correlações entre entes, o número de entidades potenciais passa a ser 480 em 2014, em detrimento das 513 obtidas em 2015.

O índice de dispersão do contexto informacional para o ano de 2014 foi de 798,84. Tal alteração resultou em uma agregação de sujeitos ligeiramente distinta do ano de 2015:

- 1) 0 a 798,84: composto pelos sujeitos metalurgia, farmacêutica, papel e celulose, mineração, movelaria, construção civil, agroindústria, telecomunicações e têxtil;
- 2) 798,85 a 1.597,68: composto pelos sujeitos petroquímica, bens de consumo, TICs, alimentos e eletroeletrônica;
- 3) 2.396,53 a 3.195,37: composto pelos sujeitos mecânica e transporte e outros.

As três mudanças mais significativas são: a separação dos sujeitos mecânica e transporte e outros em duas faixas distintas; reclassificação do sujeito TICs para faixa abaixo do índice de dispersão do contexto; e a reordenação das faixas de agregação. Apesar de as mudanças serem aparentemente desprezíveis, há que se considerar o equilíbrio entre os índices de adesão dos sujeitos. Nesse sentido, são propostos 3 domínios potenciais para o ano de 2014:

- Domínio potencial 3, composto pelo sujeito mecânica e transporte e de parte dos sujeitos que compõem a primeira faixa de agregação do contexto original para o ano de 2014, a saber: agroindústria, movelaria, papel e celulose, farmacêutica e TICs;
- Domínio potencial 4, composto pelo sujeito outros e a parte restante dos sujeitos que compõem a primeira faixa de agregação do contexto original para o ano de 2014, a saber: têxtil, telecomunicações, construção civil, mineração e metalurgia;
- Domínio potencial 5, composto pela totalidade dos sujeitos que compõem a segunda faixa de agregação, a saber: química e petroquímica, bens de consumo, eletroeletrônica e alimentos.

Verifica-se a alta sensibilidade do problema em relação à separação espaço-temporal: uma AIM utilizada em um ano não pode ser tomada, de início, como aplicável a um novo contexto temporal. Confirma-se tal premissa quando se procede à análise dos dados, de 2014 e 2015, em conjunto. O número de entidades potenciais identificados é de 1.192.

O índice de dispersão do contexto informacional elevou para 10.243,65, criando 3 domínios diferentes dos identificados anteriormente:

- Domínio potencial 6, composto pelos sujeitos mecânica e transporte, telecomunicações, construção civil, papel e celulose, farmacêutica e metalurgia;
- Domínio potencial 7, composto pelos sujeitos outros, têxtil, agroindústria, movelaria, mineração e bens de consumo;
- Domínio potencial 8, composto pelos sujeitos química e petroquímica, alimentos, TICs e eletroeletrônica.

Aplicação de PLN com tratamento de AIM

Identificada a impossibilidade de se proceder à produção de um modelo preditivo para o problema selecionado, com base na seleção indistinta de dados e dotado dos produtos de AIM, obtido ao longo dos passos de Identificação de entidades de contexto a Agrupamento por contextos espaço-tempo, procedeu-se à validação do modelo obtido. Para tal intento, os dados de 2014 e 2015 foram divididos e concatenados, conforme os domínios potenciais construídos e treinados por 10 vezes, mantendo-se as condições de treinamento descritas na Aplicação de PLN em domínio não tratado por AIM.

As médias dos resultados dos experimentos o domínio potencial 1, do ano de 2015, foram:

- *Loss* (perda) em treinamento: 0,5026263.
- Acurácia em treinamento: 78,13%.
- *Loss* (perda) em validação: 0,5137055.
- Acurácia em validação: 80,04%.
- *Loss* (perda) em teste: 0,5178427.
- Acurácia em teste: 82,63%.

As médias dos resultados dos experimentos o domínio potencial 2, do ano de 2015, foram:

- *Loss* (perda) em treinamento: 0,51442408.
- Acurácia em treinamento: 75,81%.
- *Loss* (perda) em validação: 0,58853166.
- Acurácia em validação: 71,93%.
- *Loss* (perda) em teste: 0,625711.
- Acurácia em teste: 68,43%.

As médias dos resultados dos experimentos o domínio potencial 3, do ano de 2014, foram:

- *Loss* (perda) em treinamento: 0,7437411.
- Acurácia em treinamento: 55,95%.
- *Loss* (perda) em validação: 0,6598377.
- Acurácia em validação: 57,48%.

- *Loss* (perda) em teste: 0,6948857
- Acurácia em teste: 47,93%.

As médias dos resultados dos experimentos o domínio potencial 4, do ano de 2014, foram:

- *Loss* (perda) em treinamento: 0,6966854.
- Acurácia em treinamento: 55,70%.
- *Loss* (perda) em validação: 0,6778311.
- Acurácia em validação: 55,11%.
- *Loss* (perda) em teste: 0,7213331.
- Acurácia em teste: 56,68%.

As médias dos resultados dos experimentos o domínio potencial 5, do ano de 2014, foram:

- *Loss* (perda) em treinamento: 0,7036218.
- Acurácia em treinamento: 51,93%.
- *Loss* (perda) em validação: 0,6929909.
- Acurácia em validação: 52,14%.
- *Loss* (perda) em teste: 0,6704475.
- Acurácia em teste: 53,67%.

As médias dos resultados dos experimentos o domínio potencial 6, dos anos de 2014 e 2015, em conjunto, foram:

- *Loss* (perda) em treinamento: 0,6786096.
- Acurácia em treinamento: 62,66%.
- *Loss* (perda) em validação: 0,6503350.
- Acurácia em validação: 63,79%.
- *Loss* (perda) em teste: 0,7030873.
- Acurácia em teste: 64,95%.

As médias dos resultados dos experimentos o domínio potencial 7, dos anos de 2014 e 2015, em conjunto, foram:

- *Loss* (perda) em treinamento: 0,69630314.
- Acurácia em treinamento: 59,12%.
- *Loss* (perda) em validação: 0,67950455.
- Acurácia em validação: 55,92%.
- *Loss* (perda) em teste: 0,73877194.
- Acurácia em teste: 53,24%.

As médias dos resultados dos experimentos o domínio potencial 7, dos anos de 2014 e 2015, em conjunto, foram:

- *Loss* (perda) em treinamento: 0,6438083.
- Acurácia em treinamento: 67,55%.
- *Loss* (perda) em validação: 0,6392064.
- Acurácia em validação: 63,90%.

- Loss (perda) em teste: 0,7030873.
- Acurácia em teste: 64,95%.

Discussão dos resultados

Verifica-se variação nos resultados obtidos após o tratamento do conjunto informacional original e sua separação em domínios de relevância. Alguns domínios apresentam melhora na acurácia de predição, e outros apresentam uma piora na acurácia da predição. Notoriamente, o domínio 1, referente ao ano de 2015, apresenta os melhores resultados, o que confirma o índice estatístico de dispersão menor em relação ao ano de 2014 e o conjunto de ambos os anos (562,38 para 2015, comparado a 798,84 para 2014 e 10.243,65 para 2014 e 2015) (Tabelas 1, 2 e 3).

Tabela 1 – Comparação de resultados, 2014.

Variável	2014	2014 – Domínio 3	2014 – Domínio 4	2014 – Domínio 5
<i>Train Loss</i>	0,7009084	0,7120126	0,6966854	0,7036218
<i>Train Acc</i>	53,64%	56,43%	55,70%	51,93%
<i>Val Loss</i>	0,6876946	0,6994961	0,6778311	0,6929909
<i>Val Acc</i>	53,56%	57,35%	55,11%	52,14%
<i>Test Loss</i>	0,5698085	0,6522141	0,7213331	0,6704475
<i>Test Acc</i>	55,72%	57,11%	56,68%	53,67%

Fonte: Elaborada pelos autores (2022).

Tabela 2 – Comparação de resultados, 2015.

Variável	2015	2015 – Domínio 1	2015 – Domínio 2
<i>Train Loss</i>	0,5415669	0,5026263	0,51442408
<i>Train Acc</i>	76,86%	78,13%	75,81%
<i>Val Loss</i>	0,5755080	0,5137055	0,58853166
<i>Val Acc</i>	74,39%	80,04%	71,93%
<i>Test Loss</i>	0,6152779	0,5178427	0,625711
<i>Test Acc</i>	77,21%	82,63%	68,43%

Fonte: Elaborada pelos autores (2022).

Tabela 3 – Comparação de resultados, 2014 e 2015.

Variável	Geral	Geral – Dom. 6	Geral – Dom. 7	Geral – Dom. 8
<i>Train Loss</i>	0,6592374	0,6786096	0,69630314	0,6438083
<i>Train Acc</i>	63,10%	62,66%	59,12%	67,55%
<i>Val Loss</i>	0,6829368	0,6503350	0,67950455	0,6392064
<i>Val Acc</i>	58,39%	63,79%	55,92%	63,90%
<i>Test Loss</i>	0,6336318	0,7030873	0,73877194	0,7030873
<i>Test Acc</i>	62,05%	64,95%	53,24%	64,95%

Fonte: Elaborada pelos autores (2022).

Nesse sentido, verifica-se que a AIM proposta identificou as áreas de conhecimento de maior e menor conexão e potencial de extração de padrões de IA, além de indicar as áreas que apresentam maior déficit de potencial de aprendizado, conforme distinção temporal. Para exprimir esse resultado, a Tabela 4 apresenta um resumo de contribuição, onde é assinalado valor 1 caso a área tenha figurado em um domínio com aumento da acurácia de predição, e valor -1 caso a área tenha figurado em um domínio com redução da acurácia de predição.

Tabela 4 – Análise de potencial de aprendizado por área do conhecimento.

Área do conhecimento	2014	2015	2014/2015	Potencial
Agroindústria	1	1	-1	1
Alimentos	-1	-1	1	-1
Bens de consumo	-1	-1	-1	-3
Construção civil	1	1	1	3
Eletroeletrônica	-1	-1	1	-1
Farmacêutica	1	1	1	3
Mecânica e transporte	1	-1	1	1
Metalurgia	1	1	1	3
Mineração	1	1	-1	1
Movelaria	1	1	-1	1
Papel e celulose	1	1	1	3
Química e petroquímica	-1	-1	1	-1
TICs	1	-1	1	1
Telecomunicações	1	1	1	3
Têxtil	1	1	-1	1
Outros	1	1	-1	1

Fonte: Elaborada pelos autores (2022).

Verifica-se que as áreas que constam no domínio 1, referente ao ano de 2015, apresentam valores potenciais positivos (agroindústria, construção civil, farmacêutica, metalurgia, mineração, movelaria, papel e celulose, telecomunicações, têxtil e outros), bem como a totalidade de áreas que contribuem positivamente em todos os domínios, também figurando nesse rol.

Conclusão

Ainda que os resultados tenham sido discretos sob o ponto de vista de acurácia de predição, percebe-se potencial de melhoria em grande parte dos domínios distinguidos. Dado que nenhum procedimento de enriquecimento de dados ou aprimoramento do modelo linguístico foi realizado, é plausível a conclusão de que a AIM, por si só, indicou o melhor agrupamento de dados possível em cada momento temporal, partindo-se tão somente dos registros apresentados inicialmente. Outras técnicas de obtenção da Identificação de correlações entre entes são possíveis, porém, não foram objeto deste artigo.

Referências

- Arel, I.; Rose, D. C.; Karnowski, T. P. Deep machine learning—a new frontier in artificial intelligence research [research frontier]. *IEEE Computational Intelligence Magazine*, v. 5, n. 4, p. 13-18, 2010.
- Bahdanau, D.; Cho, K.; Bengio, Y. *Neural machine translation by jointly learning to align and translate*. arXiv:1409.0473, 2014.
- Carnielli, W.; Pizzi, C. *Modalities and multimodalities*. [S.l.]: Springer Science & Business Media, 2008.
- Devlin, J. et al. *Bert*: Pre-training of deep bidirectional transformers for language understanding. arXiv:1810.04805, 2018.
- Filho, J. A. W. et al. *The brwac corpus: a new open resource for brazilian portuguese*. In: Proceedings of the eleventh international conference on language resources and evaluation (LREC 2018). [S.l.: s.n.], 2018.

- Hinton, G. E.; Osindero, S.; Teh, Y.-W. A fast learning algorithm for deep belief nets. *Neural Computation*, v. 18, n. 7, p. 1527-1554, 2006.
- Hjørland, B. What is knowledge organization (ko)? *Knowledge Organization*, v. 35, n. 2/3, p. 86-101, 2008.
- Jones, K. S. Index term weighting. *Information Storage and Retrieval*, v. 9, n. 11, p. 619-633, 1973.
- Kress, G. What is mode? In: Jewitt, C. (ed.). *The Routledge Handbook of Multimodal Analysis*. London: Routledge, 2009.
- Kress, G.; Van Leeuwen, T. *Multimodal discourse: The modes and media of contemporary communication*. London, UK, 1 Ed. Hodder Arnold Publication, 2001. 142 p.
- Kuroki Júnior, G. H. *Sobre uma arquitetura da informação multimodal: reflexões sobre uma proposta epistemológica*. 2018. 236 f. Dissertação (Mestrado em Ciência da Informação) – Universidade de Brasília, Brasília, 2018. Doi: <http://dx.doi.org/10.26512/2018.02.D.31920>.
- McCann, B. et al. Learned in translation: Contextualized word vectors. *Advances in Neural Information Processing Systems*, v. 30, 2017.
- Minaee, S. et al. Deep learning-based text classification: a comprehensive review. *ACM Computing Surveys (CSUR)*, v. 54, n. 3, p. 1-40, 2021.
- Mikolov, T. et al. Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, v. 26, 1, 2013a. Disponível em: <https://proceedings.neurips.cc/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf>. Acesso em 30 de junho de 2023.
- Mikolov, T. et al. Efficient estimation of word representations in vector space. *arXiv:1301.3781*, 2013b. Disponível em: <https://arxiv.org/pdf/1301.3781.pdf%C3%AC%E2%80%94%20%C3%AC%E2%80%9E%C5%93>. Acesso em: 30 jun. 2023.
- Pennington, J.; Socher, R.; Manning, C. D. Glove: Global vectors for word representation. *Proceedings of the Conference on Empirical Methods In Natural Language Processing*, p. 1532-1543, 2014.
- Peters, M. E. et al. Deep contextualized word representations. *Association for Computational Linguistics New Orleans*, 2018. Disponível em: <https://aclanthology.org/N18-1202/>. Acesso em: 30 jun. 2023.
- Portner, P. *Modality*. [S.l.]: Oxford University Press, 2009.
- Qiu, X. et al. Pre-trained models for natural language processing: a survey. *Science China Technological Sciences*, v. 63, n. 10, p. 1872-1897, 2020.
- Radford, A. et al. Improving language understanding by generative pre-training. 2018. Preprint. Disponível em: <https://www.cs.ubc.ca/~amuham01/LING530/papers/radford2018improving.pdf>. Acesso em: 30 jun. 2023.
- Souza, F.; Nogueira, R.; Lotufo, R. BERTimbau: pretrained BERT models for Brazilian Portuguese. In: Brazilian Conference on Intelligent Systems, BRACIS, 9., 2020, Rio Grande do Sul. *Proceedings [...]*. [S.l.:s.n.], 2020.
- van Gigch, J. P.; Moigne, J. L. L. A paradigmatic approach to the discipline of information systems. *Behavioral Science*, v. 34, n. 2, p. 128-147, 1989.
- Vaswani, A. et al. Attention is all you need. *Advances in Neural Information Processing Systems*, v. 30, 2017. Disponível em: <https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf>. Acesso em: 30 de junho de 2023.
- Wason, R. Deep learning: evolution and expansion. *Cognitive Systems Research*, v. 52, p. 701-708, 2018.



Disponível em:

<https://www.redalyc.org/articulo.oa?id=384378477015>

Como citar este artigo

Número completo

Mais informações do artigo

Site da revista em redalyc.org

Sistema de Informação Científica Redalyc
Rede de Revistas Científicas da América Latina e do Caribe,
Espanha e Portugal
Sem fins lucrativos acadêmica projeto, desenvolvido no
âmbito da iniciativa acesso aberto

George Hideyuki Kuroki Júnior, Cláudio Gottschalg-Duque
**Arquitetura da Informação Multimodal: contribuições no
desenvolvimento de Inteligência Artificial¹**
**Multimodal Information Architecture: contribution on
Artificial Intelligence developments**

Transinformação

vol. 35, e226729, 2023

Pontifícia Universidade Católica de Campinas,

ISSN: 0103-3786

ISSN-E: 2318-0889

DOI: <https://doi.org/10.1590/2318-0889202335e226729>