



ANÁLISE AUTOMATIZADA DE COMPETÊNCIAS EM DOCUMENTAÇÃO DE PLANEJAMENTO EDUCACIONAL A PARTIR DA TAXONOMIA DE BLOOM

DOI: 10.37702/2175-957X.COBIENGE.2025.6291

Autores: YURI CARLOS ACHNITIZ BATISTA BELINSKI, VINICIUS FERREIRA GOMES, ANDRÉIA DAMASIO DE LELES, TALITA DOS REIS LOPES BERBEL, JOSÉ ROBERTO CARDOSO

Resumo: Este artigo aborda a análise de competências presentes nos documentos de planejamento educacional de uma instituição de ensino superior (IES), de acordo com a Taxonomia de Bloom. Para tal, foram empregadas, técnicas de mineração de texto, processamento de linguagem natural e aprendizado de máquina, ao passo que foi realizado um estudo de caso na IES, visando testes e validação. A análise dos modelos de inteligência artificial mostraram resultados eficientes, de acordo com métricas. Pesquisa qualitativa com profissionais da IES, apontaram satisfação e abordagem útil. Este artigo contribui diretamente à área educacional, ao propor uma abordagem para rastreio e comparação de competências cognitivas nos documentos de planejamento.

Palavras-chave: Taxonomia de Bloom, Mineração de Texto, Aprendizagem de Máquina, Mineração de Texto, Aprendizado de Máquina

ANÁLISE AUTOMATIZADA DE COMPETÊNCIAS EM DOCUMENTAÇÃO DE PLANEJAMENTO EDUCACIONAL A PARTIR DA TAXONOMIA DE BLOOM

1 INTRODUÇÃO

O planejamento educacional pode ser entendido como um instrumento para a concretização de ações conscientes e intencionais para alcançar as metas estabelecidas a partir da definição de condições, materiais e métodos pertinentes ao âmbito educacional (Delgado Junior, 2022). Neste contexto, Silva *et al.* (2024) pontuam que o planejamento é um processo que deve contemplar os diversos elementos presentes no desenvolvimento acadêmico, como a escolha das bases teóricas e metodológicas pertinentes ao curso, o contexto social, econômico e cultural que a instituição de ensino está inserida, bem como a definição de objetivos educacionais claros e mensuráveis.

De tal forma, o processo de planejamento culmina na criação de três documentos principais: o Projeto Pedagógico do Curso (PPC), os Planos de Ensino (PE) e os Planos de Aula (PA). Os PPC são o documento base para a concepção de todos os cursos superiores no Brasil, neles são definidos os recursos utilizados e as habilidades a serem desenvolvidas pelos estudantes ao longo de todo o curso, com base nas Diretrizes Curriculares Nacionais estabelecidas pelo Ministério da Educação (Brasil, 2019).

Os Planos de Ensino e de Aula são elaborados para cada disciplina disponível em determinado curso, seguindo os parâmetros especificados no PPC. Os PE contemplam os objetivos referentes ao que se planeja que os estudantes desenvolvam ao longo da disciplina, ao passo que os PA clarificam o que é abordado individualmente em cada aula. À vista disso, a elaboração destes documentos proporciona um espaço para analisar os recursos pedagógicos utilizados e incentiva uma reavaliação ativa da efetividade das práticas propostas (Delgado Junior, 2022; Silva *et al.*, 2024).

Uma das possíveis formas para se explicitar e avaliar as competências e habilidades que se esperam ser desenvolvidas por educandos é através da Taxonomia de Bloom, criada por Bloom *et al.* (1956) e posteriormente revisada por Anderson e Krathwohl (2001). Por meio desta, são definidos domínios correspondentes às competências a serem ensinadas, sendo que cada domínio possui sua própria gama de categorias hierarquicamente organizadas. Para que um estudante possa avançar para a próxima categoria, ele antes deve apresentar um resultado satisfatório na categoria anterior (Ferraz e Belhot, 2010). O Domínio Cognitivo é comumente o mais estudado e utilizado por educadores de diversas áreas, tal que se baseia no ato de aprender e obter conhecimento, possuindo seis categorias que descrevem o desenvolvimento intelectual. São elas: “Lembrar”, “Entender”, “Aplicar”, “Analizar”, “Avaliar” e “Criar” (Ferraz e Belhot, 2010; Freitas e Fortes, 2020).

Considerando a possível vasta quantia de documentos pertinentes ao planejamento de uma instituição de ensino, com este artigo objetiva-se apresentar uma abordagem para a realização de sua análise, viabilizada através da implementação de um sistema que utiliza técnicas de mineração de texto e aprendizado de máquina a fim de se extrair métricas relevantes do planejamento de acordo com o Domínio Cognitivo da Taxonomia de Bloom. Deste modo, pretende-se que os dados obtidos possam ser utilizados por docentes e coordenadores de forma que sejam capazes de observar a rastreabilidade entre o todo escopo de planejamento e aprimorá-lo quando cabível.

15 a 18 DE SETEMBRO DE 2025
CAMPINAS - SP

2 TRABALHOS CORRELATOS

A utilização da Taxonomia de Bloom pode ser encontrada na literatura para a classificação de questões de provas e exames (Zhang *et al.*, 2021; Huang *et al.*, 2022; Ifham *et al.*, 2022; Rawat, Kumar e Samant, 2023) ou de objetivos de aprendizagem (Shaikh, Daudpotta e Imran, 2021; Li *et al.*, 2022; Hernández-Carrasco *et al.*, 2022). Notavelmente, não há um foco nos estudos acerca da classificação de componentes presentes no escopo geral de planejamento, tampouco como estes componentes se correlacionam.

Por instância, o estudo publicado por Li *et al.* (2022), se propõe a definir qual a melhor abordagem para a realização da classificação de objetivos de aprendizagem considerando o Domínio Cognitivo empregando técnicas de inteligência artificial. No estudo, os autores coletaram 21380 objetivos de aprendizagem provenientes de 5558 cursos de uma universidade australiana, os quais foram manualmente classificados em cada categoria do Domínio Cognitivo por pessoas capacitadas. De tal modo, os autores empregaram técnicas de Processamento de Linguagem Natural (PLN) e experimentaram modelos de aprendizado de máquina para satisfazer o objetivo proposto.

Similarmente, Hernández-Carrasco *et al.* (2022) desenvolveram um *framework* com o objetivo de auxiliar no planejamento didático no nível de graduação. O *framework* tem instrutores iniciantes como público-alvo principal, possuindo foco em validar a construção de atividades de aprendizagem ao passo que é capaz de fornecer *feedback* para o docente em caso da ausência de algum elemento. O método proposto utiliza técnicas de PLN para identificar os objetos de ensino presentes nos textos que descrevem os objetivos de aprendizagem, de forma que cada componente encontrado é classificado utilizando as métricas de Similaridade de Cosseno e Distância de Levenshtein. A análise realizada conta com uma base de dados autoral, e por conta dessa dependência, limita-se apenas ao domínio específico da área de computação.

3 METODOLOGIA APLICADA

Para a concretização do presente trabalho, foi conduzida uma pesquisa exploratória com a abordagem de um estudo de caso, realizado no âmbito de um centro universitário localizado no interior do estado de São Paulo. A instituição de ensino superior (IES) conta com cursos em diversas áreas de engenharia.

A princípio, foram estudados os elementos presentes nos artefatos de planejamento pedagógico, abrangendo a construção do referencial teórico a respeito da composição dos mesmos. Especificamente, foi estudado o que se refere ao desenvolvimento do perfil profissiográfico do discente. O estudo centrou-se na identificação e avaliação das competências delineadas nos PPC, PE e PA. A definição das competências representa o conjunto de conhecimentos, atitudes e capacidades que os educandos devem adquirir, refletindo diretamente na construção do perfil do discente ao longo do curso (Carvalho e Tonini, 2017; Cordeiro *et al.*, 2020; Leles, Zaina e Cardoso, 2024).

Neste contexto, foi observada a capacidade do planejamento em guiar o desenvolvimento dessas competências, qualificado a partir da adoção da distribuição da Taxonomia de Bloom como base para a avaliação dos objetivos de aprendizagem, nos quesitos complexidade e profundidade. Paralelamente, procedeu-se também ao estudo de técnicas de mineração de texto e aprendizado de máquina com objetivo de identificar e qualificar, de forma automatizada, a rastreabilidade das competências presentes escopo geral do planejamento. Neste contexto, destacam-se a análise estrutural dos documentos e a experimentação de técnicas de mineração de texto.

15 a 18 DE SETEMBRO DE 2025
CAMPINAS - SP

Por conseguinte, a consolidação dos estudos anteriores se deu através da construção de um sistema capaz de realizar a análise dos documentos de planejamento de ensino a partir da identificação e metrificação da rastreabilidade do Domínio Cognitivo de Bloom por meio de técnicas de mineração de texto e aprendizado de máquina. As métricas de rastreabilidade referem-se à distribuição, em termos quantitativos, dos níveis cognitivos da Taxonomia de Bloom de acordo com as competências presentes. Deste modo, é considerado o alinhamento entre os objetivos de aprendizagem dos Planos de Aula e Ensino e as competências previstas, medindo tanto a quantidade de objetivos classificados por nível cognitivo quanto o grau de correspondência com as competências esperadas.

Posteriormente, para a análise e interpretação dos resultados adquiridos foram utilizados métodos quantitativos e qualitativos. Ao que se refere à criação do sistema, foram observadas métricas de execução. Referente aos modelos de aprendizado de máquina, foram consideradas métricas de acurácia e F1-Score. Por fim, a ferramenta foi disponibilizada para membros do corpo docente do centro universitário, composto por especialistas, professores e coordenadores, com o intuito de se realizar uma validação qualitativa fundamentada por profissionais capacitados.

4 IMPLEMENTAÇÃO

A implementação do sistema foi estruturada considerando etapas sequenciais para extração, classificação e rastreamento das informações contidas nos documentos de planejamento acadêmico, empregando técnicas de mineração de texto e aprendizado de máquina na identificação e classificação das competências e habilidades de acordo com os níveis cognitivos da Taxonomia de Bloom.

Mineração de texto é o processo de obtenção de informações relevantes de um texto a partir do reconhecimento de padrões de tendências através de métodos estatísticos ou de aprendizagem de máquina, e por fim, armazenando esses dados de forma estruturada para tomada de decisão (Zhai e Massung, 2018).

O processamento de linguagem natural (PLN) é um subcampo da inteligência artificial que se concentra na interação entre computadores e seres humanos através da linguagem natural. O objetivo do PLN é permitir que as máquinas comprehendam, interpretem e respondam as entradas de linguagem humana de maneira significativa e útil. Isso envolve diversas tarefas, como análise sintática, reconhecimento de entidades nomeadas, análise de sentimentos, tradução automática, geração de texto e resumo automático (Caseli e Nunes, 2023).

4.1 Mineração de Texto

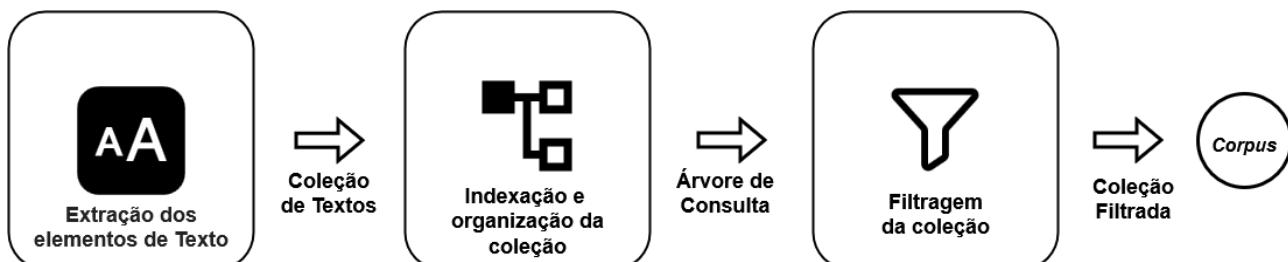
Os documentos de planejamento representam a entrada do sistema e são compostos por objetivos educacionais, competências e habilidades organizadas nos níveis de curso, disciplina e aula. A extração dessas informações é baseada no modelo conceitual de recuperação de informação proposto por Belkin e Croft (1992) e tem como resultado final o *corpus* preparado para a etapa de pré-processamento dos modelos de classificação.

O processamento do sistema (Figura 1) comprehende três etapas principais: extração, indexação e filtragem. Na fase de extração, o conteúdo é identificado e separado com base na estrutura do documento, permitindo a criação de um conjunto de segmentos textuais distintos. Em seguida, na fase de organização, esses segmentos são armazenados em uma estrutura hierárquica, onde cada componente recebe um rótulo para consulta posterior (Figura 2). Por fim, a etapa de filtragem faz a estruturação do *corpus* a partir da aplicação de critérios de seleção de acordo com os metadados conhecidos, como tipo de documento, nomenclatura das sessões e hierarquia dos elementos textuais.

15 a 18 DE SETEMBRO DE 2025
CAMPINAS - SP

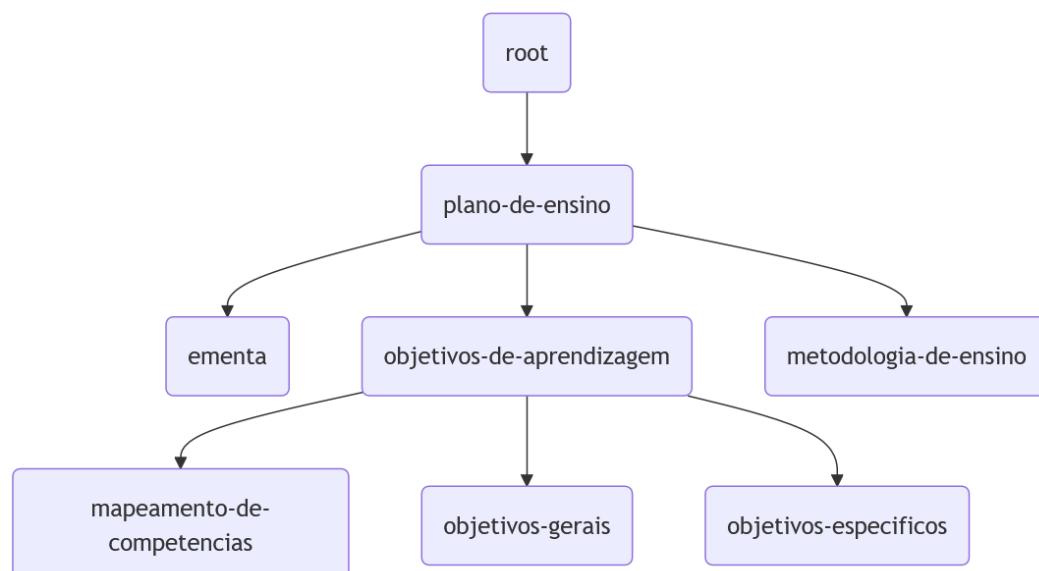
As etapas de pré-processamento foram implementadas utilizando a linguagem *Python*, com o uso da biblioteca *Pandoc* para a conversão dos arquivos no formato *DOCX*. A orquestração das etapas foi realizada por meio do Apache *Airflow*, que gerenciou a execução das tarefas de forma sequencial.

Figura 1 – Preparação do Corpus



Fonte: Autoria própria (2025).

Figura 2 – Arvore de consultas Parcial



Fonte: Autoria própria (2025).

4.2 Modelos classificadores da Taxonomia de Bloom

Os dados extraídos na etapa de mineração de texto são processados por modelos classificadores binários, que individualmente associam as competências e habilidades às categorias cognitivas da Taxonomia de Bloom. Por se tratar de um problema multi-rótulo, a criação de modelos diferentes corrobora com a possibilidade de uma competência corresponder a mais de uma categoria. Após uma análise empírica dos documentos, foi observado que as competências presentes nos PPC e PE tendem a ocorrer em frases mais longas em relação às competências nos PA. Dessa forma, foram criados dois grupos de modelos: para os PA, que utilizam até os primeiros 20 *tokens* de sua entrada e para os PPC e PE, que utilizam até os primeiros 40 *tokens*.

Para a realização de seu treinamento e teste, foram utilizados dados provenientes do estudo de Li *et al.* (2022), como já mencionado, contendo 21.380 objetivos de aprendizagem previamente categorizados. O conjunto de dados foi traduzido automaticamente de sua língua original (inglês) para português e posteriormente expandido por meio da técnica de back-

15 a 18 DE SETEMBRO DE 2025
CAMPINAS - SP

translation, aumentando o volume de dados em 96% e mantendo a distribuição proporcional entre as categorias.

Todas as etapas foram realizadas no ambiente Google Colab, utilizando a linguagem *Python* 3. A vetorização dos textos foi realizada com o modelo *BERTimbau*, especializado para a língua portuguesa. Esse modelo gera *embeddings* densos, permitindo capturar relações semânticas entre palavras. A classificação foi realizada utilizando redes *Long Short-Term Memory* (LSTM) bidirecionais, implementadas com a biblioteca *TensorFlow*. O treinamento dos modelos seguiu uma divisão de 70% dos dados para treino e 30% para teste.

Para otimizar a classificação, foi realizada uma busca de hiperparâmetros com o método *Hyperband* da biblioteca *KerasTuner*, com foco no aumento do F1-Score. Os parâmetros ajustados incluíram quantidade de camadas, unidades por camada, *dropout*, *learning rate*, *weight decay* e *batch size*. O Quadro 1 apresenta a configuração final utilizada.

Quadro 1 – Descrição dos parâmetros utilizados

Parâmetro	Descrição
Quantidade de camadas	Quantidade de camadas LSTM Bidirecionais e de <i>dropout</i> utilizadas em sequência. Varia entre 1 e 4.
Unidades de cada camada	Quantidade de neurônios de cada camada LSTM Bidirecional. Varia entre 16 e 256, com passo de 16.
<i>Dropout</i> de cada camada	Valor de <i>dropout</i> de cada camada. Varia entre 0,0 e 0,5, com passo de 0,001.
Valor de <i>learning rate</i>	Parâmetro de <i>learning rate</i> do otimizador. Varia entre 0,0001 e 0,01, com passo de 0,0001.
Valor de <i>weight decay</i>	Parâmetro de <i>weight decay</i> do otimizador. Varia entre 0,0001 e 0,01, com passo de 0,0001.
<i>Batch size</i>	Quantidade de dados do lote para o treinamento. Varia entre 16 e 256, com passo de 16.

Fonte: Autoria própria (2025).

5 RESULTADOS

Para testes e validação do sistema, a IES do estudo de caso disponibilizou um total de 64 documentos de planejamento de ensino, sendo eles: 1 Projeto Pedagógico de Curso, 30 Planos de Ensino, 33 Planos de Aulas. Todos os documentos foram fornecidos no formato DOCX e foram construídos seguindo os padrões de formatação próprios da instituição de ensino, compactados em um arquivo zip antes do processamento. Como resultado da execução das rotinas de mineração de texto, os dados extraídos foram obtidos de forma estruturada, resultando em 457 competências e habilidades, 1 curso, 89 disciplinas e 906 aulas.

5.1 Métricas de Acurácia e F1-Score

Como *baseline*, foram considerados os valores de acurácia e F1-Score disponibilizados por (Li *et al.*, 2022) no mesmo estudo referente ao conjunto de dados utilizado. Os valores estão presentes na Tabela 1. As métricas relativas aos modelos desenvolvidos com entrada de tamanho 20 e 40 estão dispostas na Tabela 2.

15 a 18 DE SETEMBRO DE 2025
CAMPINAS - SP
Tabela 1 – Resultados *baseline*

Categoria do modelo	Acurácia	F1-Score
Lembrar	0,987	0,878
Entender	0,971	0,947
Aplicar	0,961	0,931
Analisar	0,975	0,922
Avaliar	0,974	0,929
Criar	0,962	0,888

Fonte: Li *et al.* (2022).

Tabela 2 – Resultados dos grupos de tamanho 20 e 40

Categoria do modelo	Acurácia (entrada 20)	F1-Score (entrada 20)	Acurácia (entrada 40)	F1-Score (entrada 40)
Lembrar	0,929	0,929	0,935	0,935
Entender	0,930	0,930	0,926	0,926
Aplicar	0,917	0,917	0,921	0,921
Analisar	0,933	0,933	0,932	0,932
Avaliar	0,949	0,949	0,923	0,923
Criar	0,908	0,908	0,907	0,907

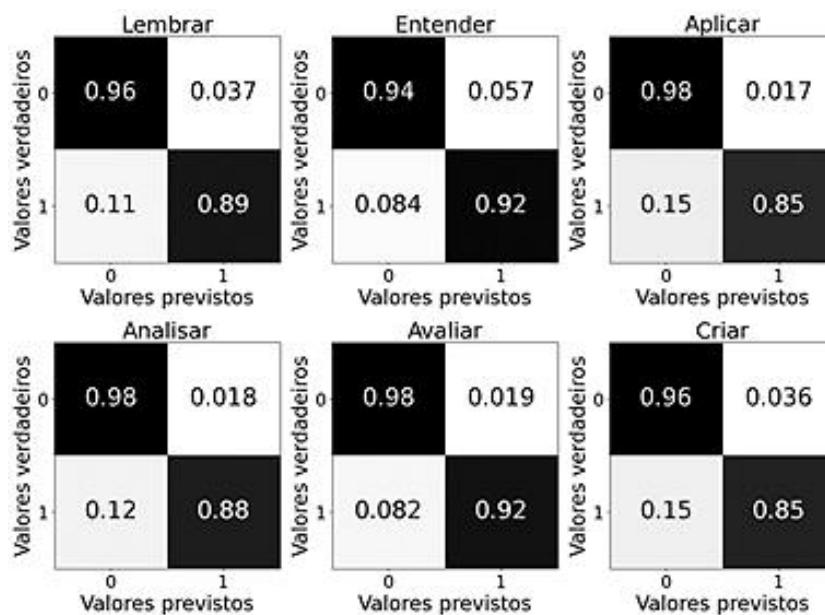
Fonte: Autoria própria (2025).

Em todos os casos, todas as métricas obtidas passaram de 90%. Como estão relativamente próximas às consideradas na *baseline*, foram consideradas aceitáveis. Devido à separação dos dados de teste, criada com uma quantia igual de rótulos verdadeiros e falsos, os valores de acurácia e F1-Score de um mesmo grupo de modelos são iguais. Portanto, uma análise mais específica acerca das previsões e erros cometidos é feita no seguinte subcapítulo, com a utilização da plotagem de matrizes de confusão.

5.2 Matrizes de Confusão

Com as matrizes de confusão, é possível observar em quais classes os modelos apresentaram maior facilidade ou dificuldade de identificação. As matrizes de cada um dos modelos com entrada 20 e 40 podem ser observadas, respectivamente, nas Figuras 3 e 4.

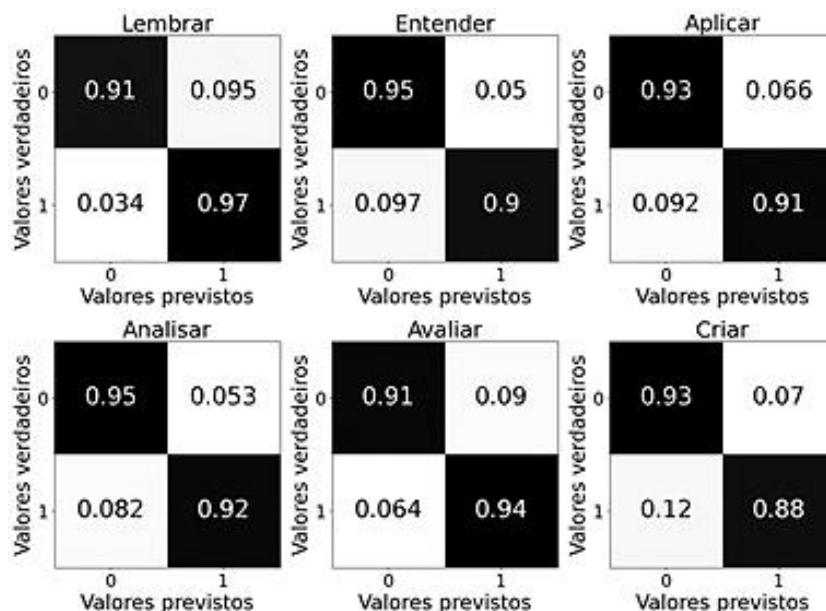
Figura 3 – Matrizes de confusão dos modelos com entrada 20



Fonte: Autoria própria (2025).

15 a 18 DE SETEMBRO DE 2025
CAMPINAS - SP

Figura 4 – Matrizes de confusão dos modelos com entrada 40



Fonte: Autoria própria (2025).

No caso dos modelos com entrada 20, todos demonstraram maior facilidade para as classes negativas de cada categoria, sendo que a dificuldade para identificação das classes positivas não foi considerada insatisfatória. Em relação aos modelos com entrada de tamanho 40, os da categoria “Lembrar” e “Avaliar” apresentaram maior facilidade na previsão de classes positivas, enquanto os outros na das negativas, sendo todas consideradas aceitáveis. Em uma visão geral, é possível concluir que nenhum dos modelos apresenta falhas graves na previsão dos dados de teste.

Em um quadro geral, foi observado que o fator determinante principal para a classificação de cada categoria se dá pela existência de uma ou mais palavras-chave (comumente verbos) geralmente presentes no início dos excertos de texto. Tal fato é corroborante com o padrão de escrita esperado pela Taxonomia de Bloom, na qual verbos específicos são associados a determinadas categorias do Domínio Cognitivo.

5.3 Construção dos Relatórios de Rastreabilidade

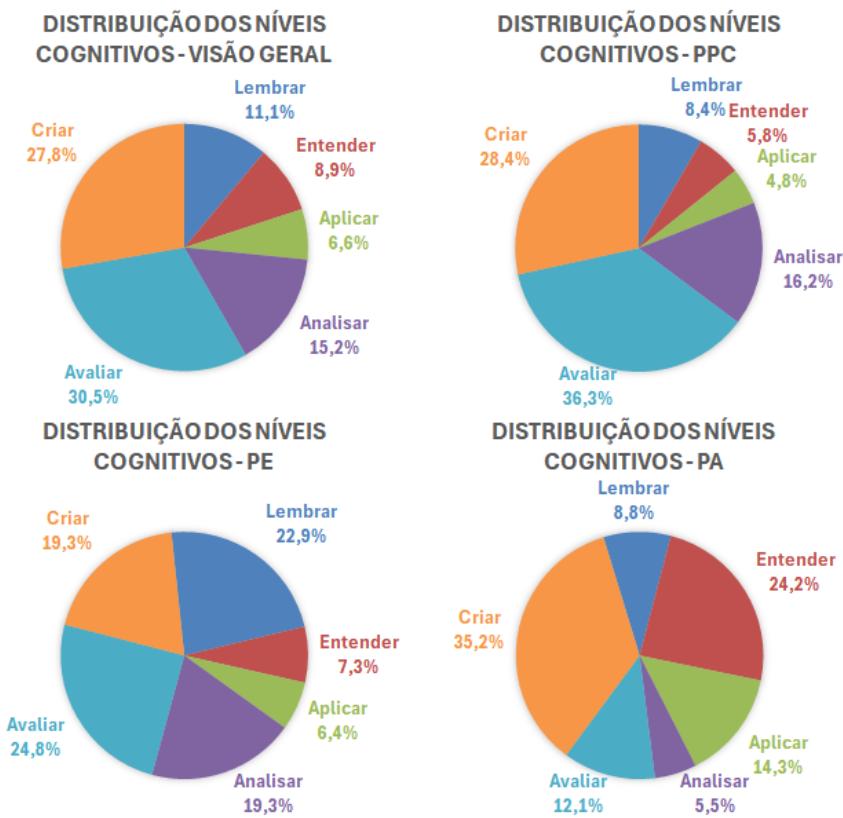
A partir dos resultados obtidos com o sistema, foram elaborados relatórios que analisam a rastreabilidade dos níveis cognitivos da Taxonomia de Bloom nos documentos de planejamento de ensino. A Figura 5 apresenta a distribuição consolidada dos níveis cognitivos em todos os documentos analisados. Observa-se uma predominância dos níveis superiores, com “Avaliar” e “Criar” representando a maior parte das competências, enquanto “Lembrar” e “Entender” aparecem com menor frequência. Essa tendência indica um foco no desenvolvimento de habilidades analíticas e criativas, em detrimento das competências mais básicas.

No nível do Projeto Pedagógico de Curso, há uma concentração nos níveis “Avaliar” e “Criar”, com menor presença dos níveis da base da pirâmide, sugerindo que as habilidades mais básicas são assumidas implicitamente na formação inicial. Os Planos de Ensino seguem um padrão semelhante, mas apresentam maior incidência do nível “Lembrar”, o que pode indicar uma preocupação adicional com a fixação de conteúdos fundamentais.

Os Planos de Aula exibem uma distribuição mais equilibrada entre os níveis cognitivos, com destaque para “Criar” e “Entender”, refletindo a operacionalização das competências de maneira prática nas atividades e metodologias de ensino.

15 a 18 DE SETEMBRO DE 2025
CAMPINAS - SP

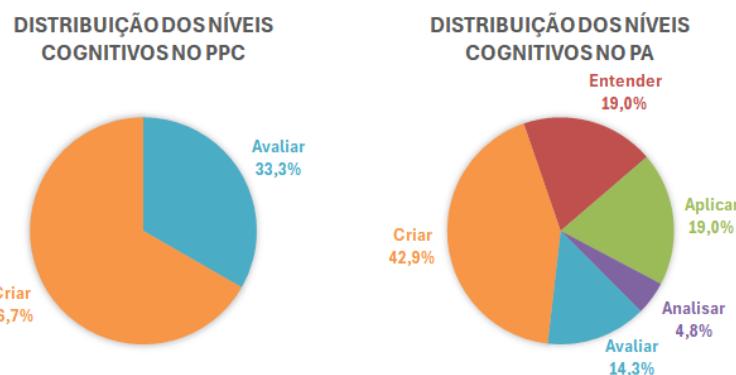
Figura 5 – Distribuição dos níveis cognitivos nos documentos de planejamento



Fonte: Autoria própria (2025).

Os relatórios gerados foram apresentados a profissionais da IES do estudo de caso envolvidos na construção dos documentos de planejamento, que avaliaram a eficácia do sistema de análise através de uma entrevista não estruturada, enfatizando a importância da comparação entre PPC e PA. Segundo os entrevistados, a presença de mais níveis cognitivos em um PA do que no PPC não o desqualifica, no entanto, a ausência de níveis destacados no PPC em seus respectivos PA é um indicador potencial de desalinhamento. As Figuras 6 e 7 exemplificam respectivamente os cenários citados. Também foi destacado que o teor dos documentos justifica uma maior granularidade nos níveis cognitivos presentes nos PA em relação aos PPC.

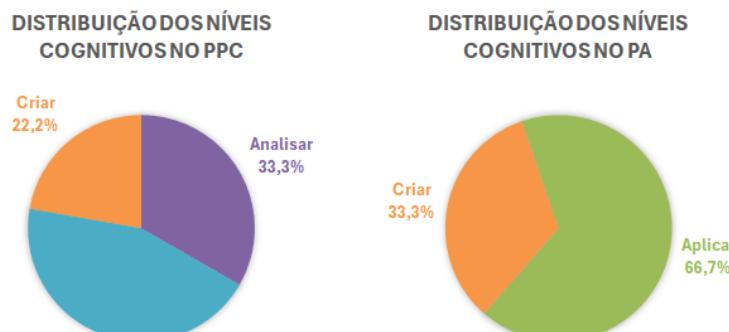
Figura 6 – Distribuição alinhada dos níveis cognitivos para um mesmo componente curricular



Fonte: Autoria própria (2025).

15 a 18 DE SETEMBRO DE 2025
CAMPINAS - SP

Figura 7 – Distribuição desalinhada dos níveis cognitivos para um mesmo componente curricular



Fonte: Autoria própria (2025).

6 LIMITAÇÕES

A ferramenta apresenta limitações relacionadas ao tipo de arquivo e estrutura dos documentos processados, restringindo a análise apenas arquivos DOCX formatados seguindo o padrão de documentos de planejamento da IES do estudo de caso, o que consequentemente levantou exceções em relação ao processamento de documentos com elementos textuais inconsistentes.

Ambos os grupos de modelos, com entradas de 20 e 40, apresentam limitações similares. Foi notado que a presença de palavras específicas agrega um peso indesejado para a inferência, induzindo em falsos positivos. Exemplos incluem palavras referentes à categoria “Criar”, como “formar” ou “desenvolver”; e à categoria “Aplicar”, com palavras como “realizar” e “atuar”. Tais palavras podem ser utilizadas como auxiliares em outros contextos, de forma que apenas sua presença não consolida uma competência em uma categoria.

Os modelos com entrada de tamanho 40, utilizados para os PPC e PE, apresentam bons resultados na interpretação de competências mais curtas, porém demonstram dificuldades para as competências com mais termos. Uma possível causa é o fato de que competências mais longas trazem consigo mais palavras que podem causar ruídos indesejados durante a análise pelo modelo, implicando na presença de falsos positivos. Por mais que uma entrada maior de tokens possa agravar esta situação, sua utilização se justifica pelo fato de que certas palavras-chave podem ocorrer no final de um excerto, que caso contrário não seriam captadas.

7 CONSIDERAÇÕES FINAIS

O presente artigo teve como objetivo empregar técnicas de aprendizado de máquina e mineração de texto para a análise documentos de planejamento educacional, sendo realizado um estudo de caso em um centro universitário do interior paulista. A proposta visou facilitar o processo de planejamento, promovendo uma análise mais objetiva e detalhada dos documentos envolvidos. Após a análise dos resultados obtidos, foi atestado que a ferramenta desenvolvida auxilia educadores a identificar e incorporar competências desejadas em suas disciplinas, contribuindo para um planejamento educacional mais alinhado com as demandas contemporâneas. Este estudo contribui diretamente para a área educacional ao propor uma abordagem para o rastreio e comparação de competências cognitivas nos documentos de planejamento.

Por fim, o presente trabalho pode servir de base para a criação de sistemas posteriores com objetivos semelhantes, que possam se utilizar de métricas diferentes para análise ou um

15 a 18 DE SETEMBRO DE 2025
CAMPINAS - SP

escopo diferente de documentos existentes nas instituições de ensino. Também há a possibilidade de ser desenvolvido outro sistema que consuma os dados disponibilizados para exibi-los de acordo com as preferências dos profissionais envolvidos, assim como facilite o *upload* dos documentos.

REFERÊNCIAS

ANDERSON, L. W.; KRATHWOHL, D. R. **A taxonomy for learning, teaching, and assessing:** A revision of Bloom's taxonomy of educational objectives: complete edition. [S.I.]: Addison Wesley Longman, Inc., 2001.

BELKIN, N. J.; CROFT, W. B. Information filtering and information retrieval: two sides of the same coin? **Commun. ACM**, Association for Computing Machinery, New York, NY, USA, v. 35, n. 12, p. 29–38, dez. 1992.

BLOOM, B. S. et al. **Cognitive domain. taxonomy of educational objectives.** In: Handbook. [S.I.]: David McKay New York NY, 1956. v. 1.

BRASIL. Institui a orientação às diretrizes curriculares nacionais dos cursos superiores. **Diário Oficial da União**, Brasília, DF, 2019. Disponível em: http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=119811-pces334-19&category_slug=agosto-2019-pdf&Itemid=30192. Acesso em: 02 jun. 2025.

CARVALHO, L.; TONINI, A. Uma análise comparativa entre as competências requeridas na atuação profissional do engenheiro contemporâneo e aquelas previstas nas diretrizes curriculares nacionais dos cursos de engenharia. **Gestão & Produção**, v. 24, 2017.

CASELI, H. d. M.; NUNES, M. d. G. V. **Processamento de linguagem natural:** conceitos, técnicas e aplicações em português. [S.I.]: BPLN. 2023.

CORDEIRO, F. et al. Production engineers profiling: competences of the professional the market wants. **Production**, v. 30, 2020.

DELGADO JUNIOR, C. M. D. A metodologia e o planejamento de aula no ensino superior. **Revista Científica Aceritte - ISSN 2763-8928**, v. 2, n. 9, p. e2993, ago. 2022.

FERRAZ, A. P. d. C. M.; BELHOT, R. V. Taxonomia de bloom: revisão teórica e apresentação das adequações do instrumento para definição de objetivos instrucionais. **Gestão & Produção**, Universidade Federal de São Carlos, v. 17, n. 2, p. 421–431, 2010.

FREITAS, E. J. D. R.; FORTES, L. S. Aprendizagem ativa aplicada à engenharia: Um estudo sobre a percepção do aprendizado. **Revista de Ensino de Engenharia, Revista de Ensino em Engenharia**, v. 39, n. 1, p. 295–311, 2020.

HERNÁNDEZ-CARRASCO, D. A. et al. A framework to assist in didactic planning at undergraduate level. **Mathematics**, v. 10, n. 9, 2022.

HUANG, J. et al. Automatic classroom question classification based on bloom's taxonomy. In: Proceedings of the 13th International Conference on Education Technology and Computers.

15 a 18 DE SETEMBRO DE 2025
CAMPINAS - SP

13, 2022, New York, NY. **Proceedings** [...] Association for Computing Machinery, 2022. p. 33–39.

IFHAM, M. et al. Automatic classification of questions based on bloom's taxonomy using artificial neural network. In: 2022 International Conference on Decision Aid Sciences and Applications (DASA). [S.I.: s.n.], **Proceedings** [...] DASA, 2022. p. 311–315.

LELES, A.; ZAINA, L.; CARDOSO, J. R. Challenge- based learning for competency development in engineering education, a prisma-based systematic literature review. **IEEE Transactions on Education**, v. 67, n. 5, p. 746–757, 2024.

LI, Y. et al. Automatic classification of learning objectives based on bloom's taxonomy. In: Proceedings of the 15th International Conference on Educational Data Mining, 15, 2022. Durham, United Kingdom. **Proceedings** [...] International Educational Data Mining Society, 2022. p. 530–537.

RAWAT, A.; KUMAR, S.; SAMANT, S. S. A systematic review of question classification techniques based on bloom's taxonomy. In: 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), 14. [S.I.: s.n.], **Proceedings** [...] ICCCNT, 2023. p. 1–7.

SHAIKH, S.; DAUDPOTTA, S. M.; IMRAN, A. S. Bloom's learning outcomes' automatic classification using lstm and pretrained word embeddings. **IEEE Access**, v. 9, p. 117887–117909, 2021.

SILVA, N. et al. Desvendando o planejamento educacional: Uma análise da aplicação prática nos contextos do plano de ensino e plano de aula. **Revista Ibero-Americana de Humanidades, Ciências e Educação**, v. 10, p. 2287–2296, 03 2024.

ZHAI, C.; MASSUNG, S. **Text Data Management and Analysis: A Practical Introduction to Information Retrieval and Text Mining**. [S.I.]: Association for Computing Machinery and Morgan & Claypool, 2018.

ZHANG, J. et al. Automated classification of computing education questions using bloom's taxonomy. In: Proceedings of the 23rd Australasian Computing Education Conference, 23, 2021. New York, NY. **Proceedings** [...] New York: Association for Computing Machinery, 2021. p. 58–65.

AUTOMATED ANALYSIS OF COMPETENCES IN EDUCATIONAL PLANNING DOCUMENTATION BASED ON BLOOM'S TAXONOMY

Abstract: This article employs an analytical approach to examine the competences delineated in the educational planning documents of a higher education institution (HEI), as delineated by Bloom's Taxonomy. To this end, text mining, natural language processing and machine learning techniques were utilised, while a case study was conducted at the HEI, with the aim of testing and validation. The analysis of the artificial intelligence models yielded efficient results. The findings from qualitative research conducted with HEI professionals highlighted their satisfaction with the approach and its perceived practicality. This article directly contributes to

REALIZAÇÃO



Associação Brasileira de Educação em Engenharia



15 a 18 DE SETEMBRO DE 2025
CAMPINAS - SP

the educational field by proposing an approach for tracking and comparing cognitive competences in planning documents.

Keywords: Bloom's Taxonomy, Text Mining, Machine Learning

ORGANIZAÇÃO



PONTIFÍCIA UNIVERSIDADE CATÓLICA

REALIZAÇÃO



Associação Brasileira de Educação em Engenharia

ORGANIZAÇÃO



PONTIFÍCIA UNIVERSIDADE CATÓLICA

