Pós-Graduação Lato Sensu
MODELAGEM EM CIÊNCIA E TECNOLOGIA

CONVITE DE DEFESA

Agrupamento em texto: análise semântica latente e k-means aplicados em dados de avaliação institucional

Carlos Bino de Souza

13 de junho de 2025 - 14h

meet.google.com/iog-ggje-iuh

UnifalMG
Universidade Federal de Alfenas

Pós-Graduação Lato Sensu
MODELAGEM EM CIÊNCIA E TECNOLOGIA

Agrupamento em texto: análise semântica latente e k-means aplicados em dados de avaliação institucional

Carlos Bino de Souza

Banca:

Deive Ciro Oliveira (UNIFAL-MG)

Gislene Araújo Pereira (UNIFAL-MG)

Danilo Machado Pires (UNIFAL-MG)

Unifal MG
Universidade Federal de Alfenas

Pós-Graduação Lato Sensu
MODELAGEM EM CIÊNCIA E TECNOLOGIA

Agrupamento em texto: análise semântica latente e k-means aplicados em dados de avaliação institucional

Este artigo aborda o uso de técnicas de Processamento de Linguagem Natural (NLP) na análise automatizada de comentários qualitativos submetidos à Comissão Própria de Avaliação (CPA) da Universidade Federal de Alfenas (Unifal-MG). O objetivo é realizar a análise semântica automatizada dos comentários qualitativos enviados à CPA da Unifal-MG, aplicando técnicas de NLP e LSA, com SVD, seguidas da aplicação do algoritmo de clusterização k-means. Justifica-se a pesquisa pela dificuldade enfrentada pela CPA em sistematizar e interpretar, de forma objetiva e escalável, grandes volumes de dados textuais, geralmente tratados de maneira manual e subjetiva. Para tanto, foram aplicadas técnicas de NLP, Análise Semântica Latente (LSA), vetorização TF-IDF, redução de dimensionalidade com Decomposição por Valores Singulares (SVD), agrupamento por meio do algoritmo \textit{k-means} e a visualização dos cluster por meio de nuvens de palavras. A definição do número de \textit{clusters} baseou-se na análise do coeficiente de silhueta e de um dendrograma. Os resultados revelaram dez agrupamentos que permitem identificar os assuntos recorrentes de cada cluster.

Unifal MG
Universidade Federal de Alfenas