

Dimensão	CrITÉrios de avaliaÇão	Peso	Nota
Bag of Words	Os dados foram vetorizados com BoW e foram calculados termos mais frequentes do corpus, incluindo análises com visualizações?	0,1	
TF-IDF	a. Os dados foram vetorizados com TF-IDF e foram calculadas as similaridades entre os documentos? b. Foram analisados clusters, sendo visualizados e interpretados com PCA ou outra técnica semelhante?	0,2	
Bônus	Foram empregados n-gramas?	0,1	
Embeddings	a. Foram aplicados word embeddings e técnicas relacionadas? (similaridade, sistemas de recomendação ou busca, análise de tópicos, identificação de entidades)	0,3	
Interpretação dos resultados	Produção de texto analisando os resultados obtidos e a adequação das diferentes técnicas e algoritmos à base de dados escolhida.	0,4	