Roteiro para o Vídeo

1. Introdução (30-60 segundos)

Cumprimente o público e apresente o tema:

"Olá, meu nome é João Nascimento, e hoje vou apresentar o projeto desenvolvido para a disciplina de Processamento de Linguagem Natural. Nosso foco foi criar um sistema de análise de sentimentos para classificar frases em positivo, negativo ou neutro."

Explique a relevância do tema:

"A análise de sentimentos é amplamente utilizada em áreas como marketing, suporte ao cliente e até mesmo para monitorar opiniões públicas nas redes sociais. Este projeto foi motivado pela necessidade de entender o feedback de forma automatizada e eficiente."

2. Descrição do Problema e Objetivo (1-2 minutos)

Explique o problema que você está resolvendo:

"Nosso problema central é identificar automaticamente o sentimento em frases de texto, reduzindo o esforço manual e trazendo insights para negócios e pesquisa."

Contextualize o objetivo:

"O objetivo é desenvolver um modelo de aprendizado de máquina que seja capaz de classificar frases em três categorias: positivo, negativo e neutro."

3. Base de Dados (1-2 minutos)

Descreva brevemente a base de dados utilizada:

"O conjunto de dados contém frases em português, classificadas manualmente em sentimentos positivos, negativos e neutros. Ele foi obtido de [mencionar a fonte]."

Destaque qualquer pré-processamento necessário:

"Realizamos etapas de limpeza nos dados, como a remoção de caracteres especiais, conversão para minúsculas e normalização dos textos."

4. Etapas do Projeto (4-6 minutos)

Divida essa parte em subetapas:

Pré-processamento

"Utilizamos técnicas de limpeza textual para remover ruídos e padronizar o texto. Isso incluiu a remoção de caracteres especiais, números e a normalização para letras minúsculas."

Extração de Características

"Para representar os textos numericamente, utilizamos a técnica de Bag of Words com a biblioteca CountVectorizer. Essa abordagem transforma o texto em uma matriz que reflete a frequência de palavras."

Modelos de Machine Learning

"O modelo principal escolhido foi o Naive Bayes, devido à sua eficiência em problemas de classificação de texto. Também comparamos outros modelos, como SVM e Random Forest, para validar os resultados."

Validação e Métricas

"Utilizamos validação cruzada para avaliar a consistência do modelo. As principais métricas analisadas foram a acurácia, precisão, revocação e F1-score. Além disso, geramos uma matriz de confusão para visualizar melhor os erros."

5. Resultados e Discussão (2-3 minutos)

Mostre os principais resultados:

"Nosso modelo Naive Bayes alcançou uma acurácia de X%, com bons desempenhos nas métricas de precisão e revocação. A matriz de confusão mostrou que a maior parte dos erros ocorreu na classificação de frases neutras."

Discuta limitações e melhorias:

"Os resultados são promissores, mas percebemos que o modelo poderia ser aprimorado com um conjunto de dados maior ou técnicas mais avançadas, como embeddings semânticos, como o Word2Vec ou BERT."

6. Trabalhos Futuros (1-2 minutos)

Proponha melhorias:

"Como próximos passos, pretendemos experimentar modelos mais avançados, como redes neurais profundas, e testar técnicas de embeddings para capturar melhor o contexto das frases. Também planejamos aumentar o conjunto de dados para melhorar a generalização."

7. Encerramento (30-60 segundos)

Conclua a apresentação:

"Com este projeto, conseguimos implementar uma solução eficiente para análise de sentimentos, explorando técnicas de PLN e aprendizado de máquina. Agradeço a oportunidade de apresentar e fico à disposição para perguntas ou comentários. Obrigado!"