



Ministério da Educação
UNIVERSIDADE FEDERAL DO PARANÁ
SEPT – Setor de Educação Profissional e Tecnológica
Especialização em Inteligência Artificial Aplicada



TRABALHO FINAL – IAA015 – Tópicos de Inteligência Artificial

Equipe 16 (<https://equipe16-iaa.github.io/>)

Github: <https://github.com/Equipe16-IAA>

Repositório da disciplina:
<https://github.com/Equipe16-IAA/IAA015>

Ana Beatriz Kindinger
Daniel Victor Andrade
Igor Buess Atala Y Mansour
Marlon Mateus Prudente de Oliveira

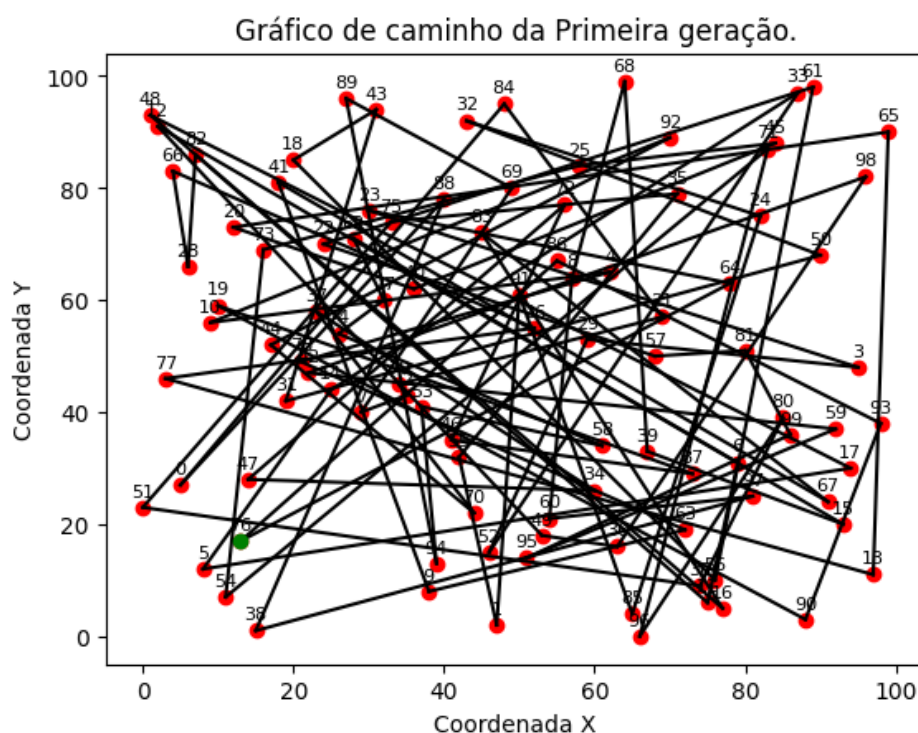
1 – Algoritmo Genético

O Problema do Caixeiro Viajante (PCV) é um problema clássico de otimização combinatória, que busca o menor caminho para visitar um conjunto de cidades. Por ser um problema NP-Completo, sua complexidade aumenta superpolinomialmente com o número de cidades. Neste trabalho, um algoritmo genético foi usado para propor uma solução otimizada em relação a uma solução aleatória, dado que o PCV ainda não possui uma solução ótima para um número muito grande de cidades.

Para o cruzamento, foi utilizada a técnica de Crossover de Ordem (OX) com um ponto de corte, devido à restrição de que os valores das cidades não podem se repetir. Para a mutação, o indivíduo foi dividido em dois segmentos, e os alelos do segundo segmento foram embaralhados.

Na população inicial, foram criados 100 indivíduos aleatoriamente, e o indivíduo com o melhor fitness foi selecionado para a segunda geração, implementando o elitismo. Foram observadas 30.000 gerações, com uma taxa de cruzamento de 90% e uma taxa de mutação de 1%.

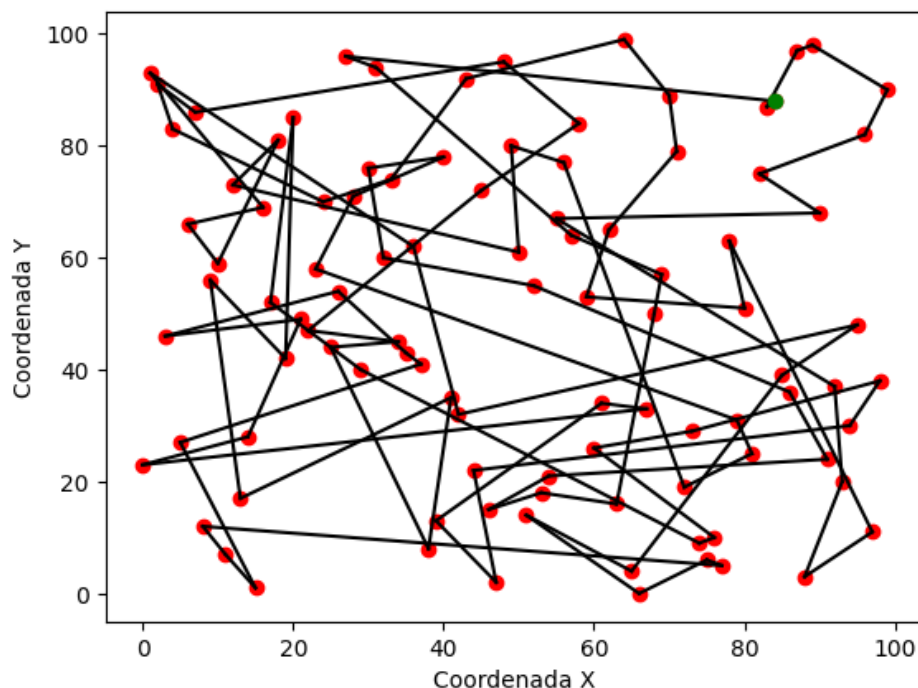
Melhor indivíduo da primeira geração:



Distância total percorrida: 5569.9 pixels

Melhor indivíduo da última geração:

Gráfico de caminho encontrado.



Distância total percorrida: 2231.6 pixels

Os resultados demonstram a eficácia do AG em reduzir a distância total das rotas. Além da observação da menor distância euclidiana, a plotagem das rotas otimizadas revela uma significativa diminuição no número de cruzamentos entre os caminhos percorridos. Essa redução de cruzamentos sugere uma menor complexidade da rota, o que corrobora a diminuição da distância total percorrida.

2 – Comparando a representação de dois modelos vetoriais

Para visualizar a representação vetorial de textos, selecionamos dez fragmentos curtos, abrangendo temas como clima, culinária e animais, com similaridades intencionais para evidenciar o agrupamento vetorial. O pré-processamento incluiu conversão para minúsculas, remoção de pontuação, números e stopwords, seguido de tokenização.

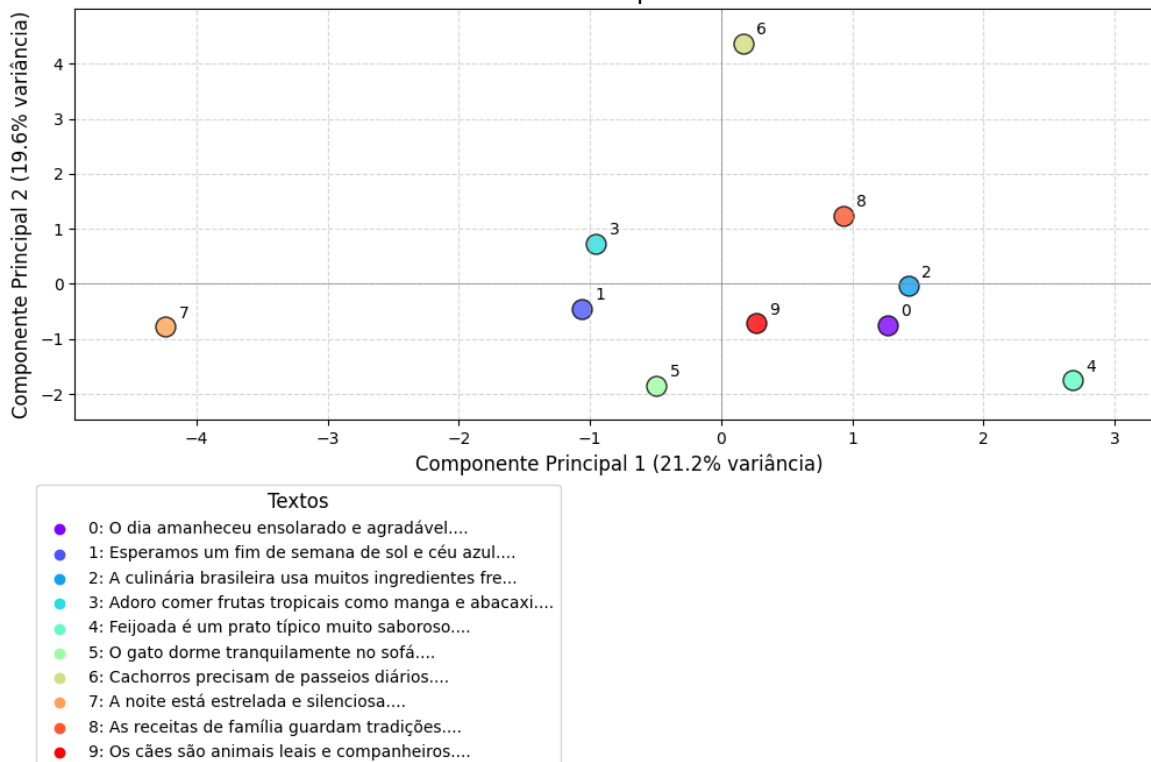
A representação vetorial foi obtida via Word2Vec, treinado com 100 épocas para capturar nuances semânticas em nosso pequeno corpus. A representação de cada sentença foi calculada pela média dos vetores das palavras, resultando em vetores de alta dimensionalidade.

A Análise de Componentes Principais (PCA) foi aplicada para reduzir a dimensionalidade, facilitando a visualização em gráficos 2D. Os resultados, multiplicados por 100 para clareza visual, demonstram o agrupamento de sentenças semanticamente similares, validando a eficácia da representação vetorial e da PCA na identificação de relações textuais.

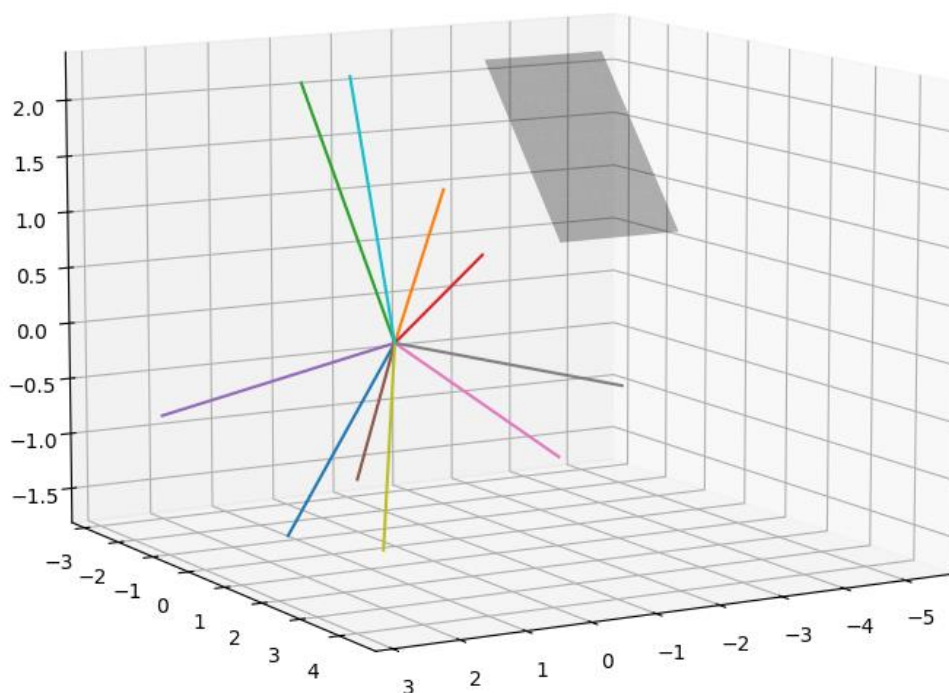
Frases utilizadas		relações
1	O dia amanheceu ensolarado e agradável.	Clima 1
2	Esperamos um fim de semana de sol e céu azul.	Clima 2 (similar a 1)
3	A culinária brasileira usa muitos ingredientes frescos.	Culinária 1
4	Adoro comer frutas tropicais como manga e abacaxi.	Culinária/frutas 2
5	Feijoada é um prato típico muito saboroso.	Culinária 3 (similar ao 3)
6	O gato dorme tranquilamente no sofá.	Animais 1
7	Cachorros precisam de passeios diários.	Animais 2
8	A noite está estrelada e silenciosa.	Clima/Noite 3 (ligeiramente similar ao 1/2)
9	As receitas de família guardam tradições.	Culinária 4 (similar ao 3/5)
10	Os cães são animais leais e companheiros.	Animais 3 (similar ao 7)

Visualização dos vetores

Visualização 2D dos Vetores de Sentenças (Word2Vec + PCA)
Variância Total Explicada: 55.3%



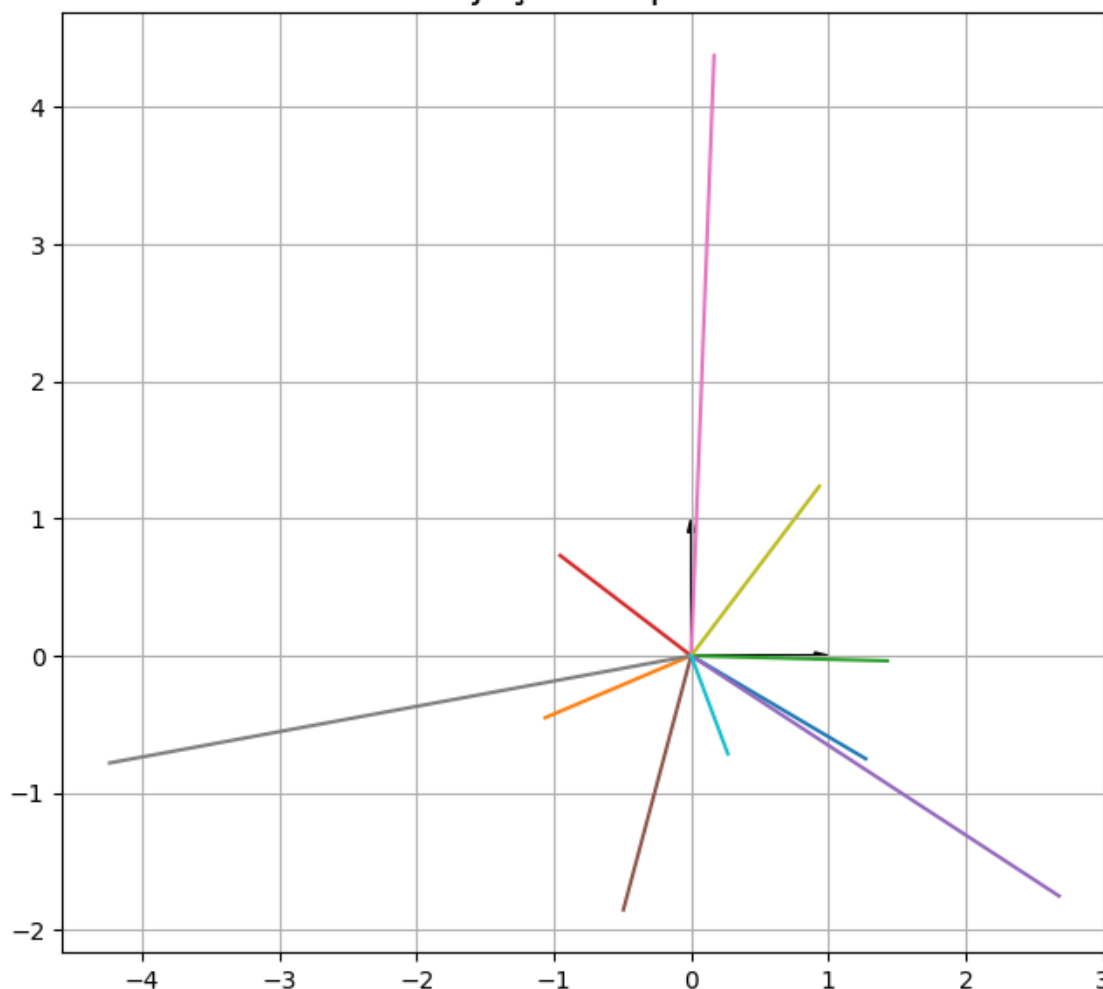
O gráfico acima mostra cada texto original como um ponto no espaço 2D. A posição de cada ponto é determinada pelo vetor da sentença (média dos vetores das palavras) após a redução com PCA. Textos com conteúdo semanticamente similar (segundo o modelo Word2Vec) devem aparecer mais próximos uns dos outros. Com um corpus pequeno, as relações capturadas pelo Word2Vec podem ser fracas ou baseadas em coocorrências simples de palavras.



Textos	
0:	O dia amanheceu ensolarado e agradável....
1:	Esperamos um fim de semana de sol e céu azul....
2:	A culinária brasileira usa muitos ingredientes fre...
3:	Adoro comer frutas tropicais como manga e abacaxi....
4:	Feijoada é um prato típico muito saboroso....
5:	O gato dorme tranquilamente no sofá....
6:	Cachorros precisam de passeios diários....
7:	A noite está estrelada e silenciosa....
8:	As receitas de família guardam tradições....
9:	Os cães são animais leais e companheiros....

O gráfico acima representa a visualização dos vetores das sentenças em três dimensões, essa técnica é uma alternativa para gerar uma visualização 2D desses vetores com a dimensionalidade menor e, portanto, mais compreensível para a interpretação humana, como segue abaixo:

Projeção no plano



Textos

- 0: O dia amanheceu ensolarado e agradável....
- 1: Esperamos um fim de semana de sol e céu azul....
- 2: A culinária brasileira usa muitos ingredientes fre...
- 3: Adoro comer frutas tropicais como manga e abacaxi....
- 4: Feijoada é um prato típico muito saboroso....
- 5: O gato dorme tranquilamente no sofá....
- 6: Cachorros precisam de passeios diários....
- 7: A noite está estrelada e silenciosa....
- 8: As receitas de família guardam tradições....
- 9: Os cães são animais leais e companheiros....