Information Retrieval and Web Search

Project

Semantic Game Finder

Main Idea and Overview:

The main goal of the project is to create a search system that does more than just keyword matching, by giving it the capability to understand the context of the query as well.

The core idea is to leverage this knowledge of intent behind the user query through natural language understanding and combine it with effective information retrieval models to provide a specifically tailored set of results for the user.

Use cases include providing a list of games with similar titles or a list of games with similar description even when the users don't use specific keywords.

Implementation Components:

Natural Language processing: Contextual Analysis.

- A BERT or GPT to get contextual/relationship information from queries – query expansion

Key Outcome: Expanded query with keywords that align with the context provided by user query.

Data: (probable): Expanded query fed to retrieve relevant documents.

- Data API keys from IGDB – international Game database.
- Data API keys from VGG.

Information Retrieval: Expanded query and Documentation are consumed here.

- Reverse indexing for quick term retrieval documents with keywords.
- TF-IDF to calculate the (weights) importance of the terms in document relative to the entire collection.
- Vector space model to create ranking.
- Optional probabilistic ranking layer for validation

Key Outcome: Sorted list of game titles in ranked system that are provided for user's query.

EXAMPLE USE CASE: -> query = "Find adventure game like Zelda"

RESULTS: -> list[Okami, Darksiders, Horizon Zero Dawn, Elder scrolls]