elastic | The Search AI Company

Search Like a Human
Building an AI Semantic Copilot with Elasticsearch

NAVIN BALAJI RK

Introduction

- Problem: Why Traditional Search Falls Short.

- Content:
  Users expect natural language understanding:
  “Show me mind-bending sci-fi thrillers similar to Inception”
  “Find white sneakers with black stripes in size 9”
  “Suggest spicy Indian vegetarian dishes for dinner”

  Keyword-based systems fail to understand meaning Need semantic search that understands intent

- Tools: Javascript , Microsoft Azure Open AI and Elasticsearch

How do search like Semantic?

Elasticsearch
Store and search vectors using kNN

Azure OpenAI
Generate text embeddings from queries
(text-embedding-ada-002)

Node.js API
Connect everything together

HTML/JS Frontend
Display results in real-time

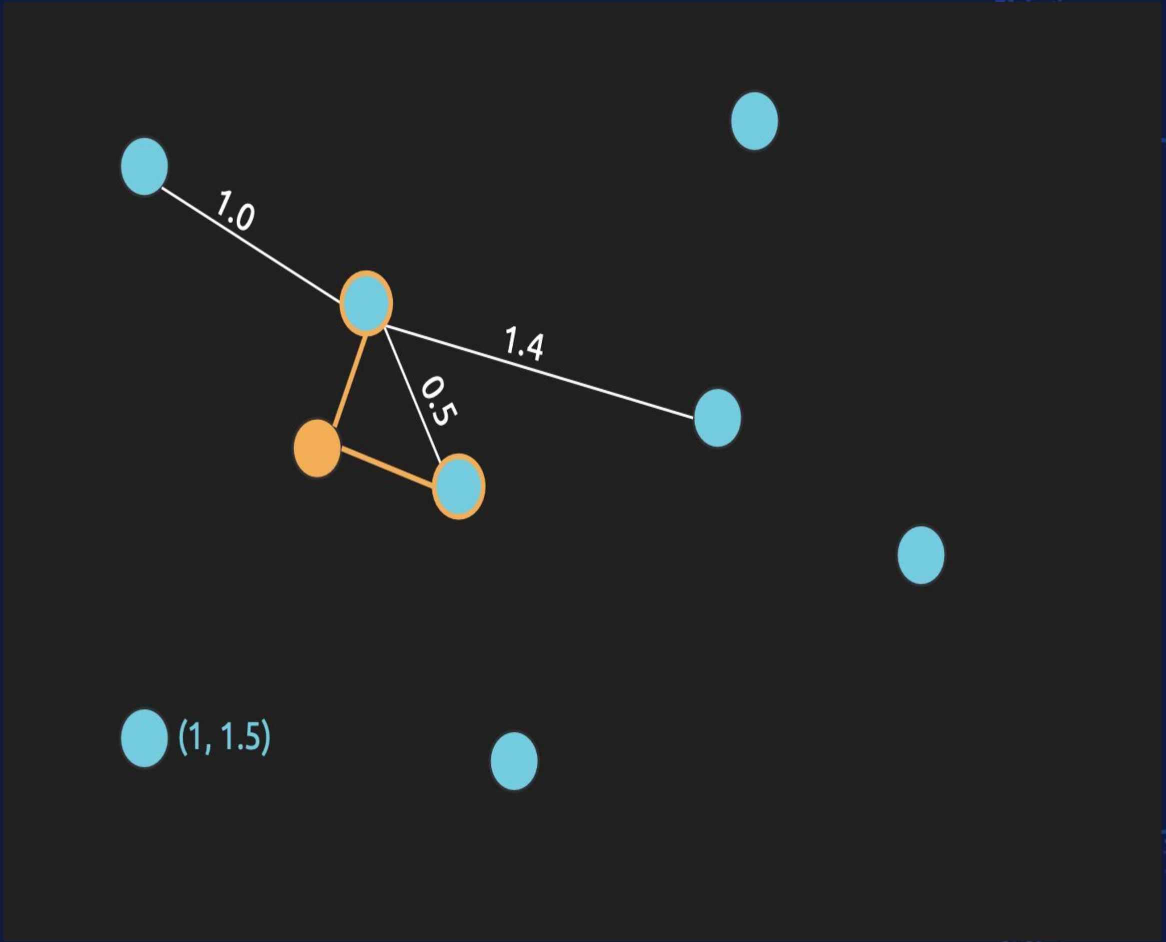
Benefits of Semantic Search

- Better user experience
- Accurate, meaningful results
- Scalable with Elasticsearch
- Future-ready with AI embeddings
- Easy to extend to products, articles, or documents

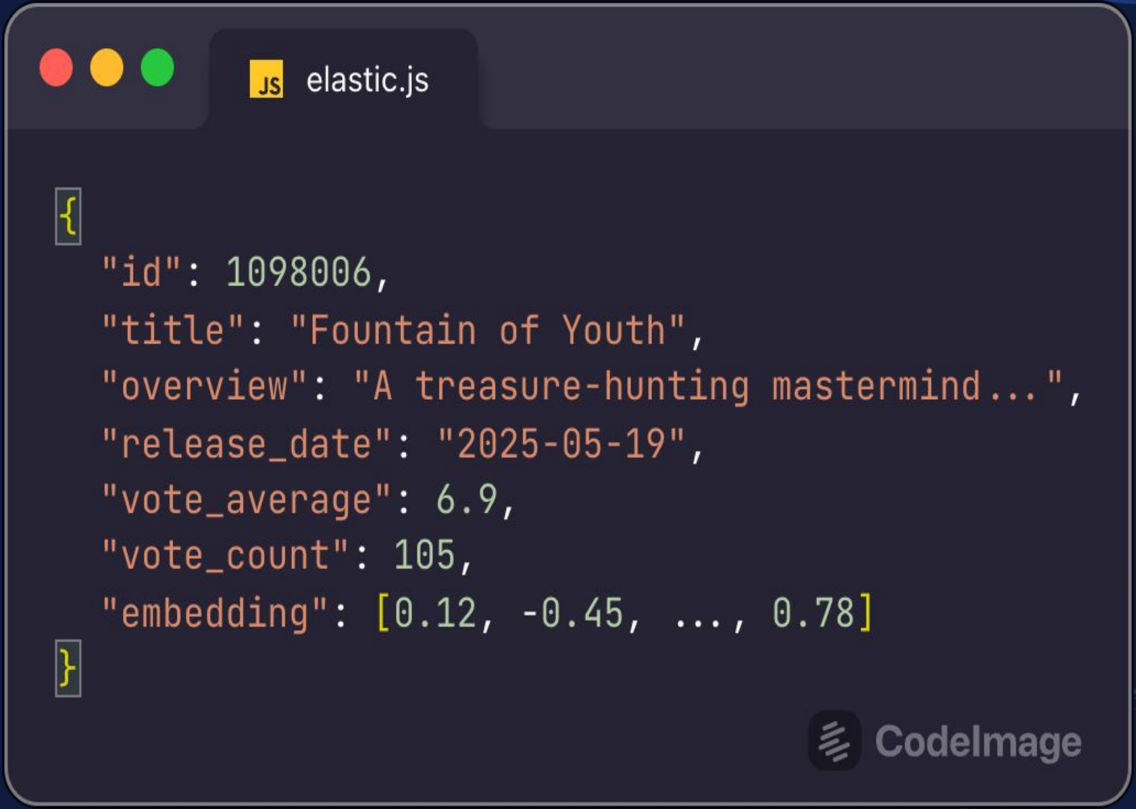
Elasticsearch Nearest Neighbor Search

The `knn` search option accepts a number of parameters that configure the search:

- `field`: the field in the index to search. The field must have a `dense_vector` type.
- `query_vector`: the embedding to search for. This should be an embedding generated from the search text.
- `num_candidates`: the number of candidate documents to consider from each shard. Elasticsearch retrieves this many candidates from each shard, combines them into a single list and then finds the closest "k" to return as results.
- `k`: the number of results to return. This number has a direct effect on performance, so it should be kept as small as possible. The value passed in this option must be less than `num_candidates`.


1.0
1.4
0.5
(1, 1.5)

Dataset Overview



```
{
  "id": 1098006,
  "title": "Fountain of Youth",
  "overview": "A treasure-hunting mastermind...",
  "release_date": "2025-05-19",
  "vote_average": 6.9,
  "vote_count": 105,
  "embedding": [0.12, -0.45, ..., 0.78]
}
```

Final Thoughts

The Future of Search Is Semantic

- Embeddings are the future of search
- Combining LLMs with vector databases unlocks powerful tools
- This project shows how easy it is to build intelligent search today

Questions?

Thank You

More about me

NAVIN BALAJI RK