

Case Study

Data Science



Case Study: Information Retrieval Engine for Trading Strategy Creation

Problem Statement

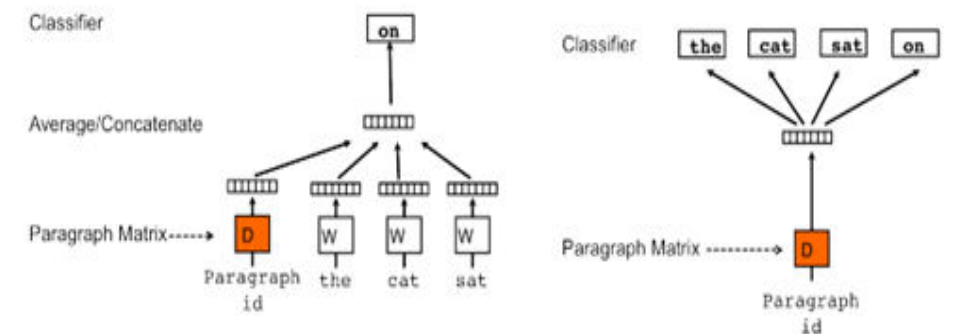
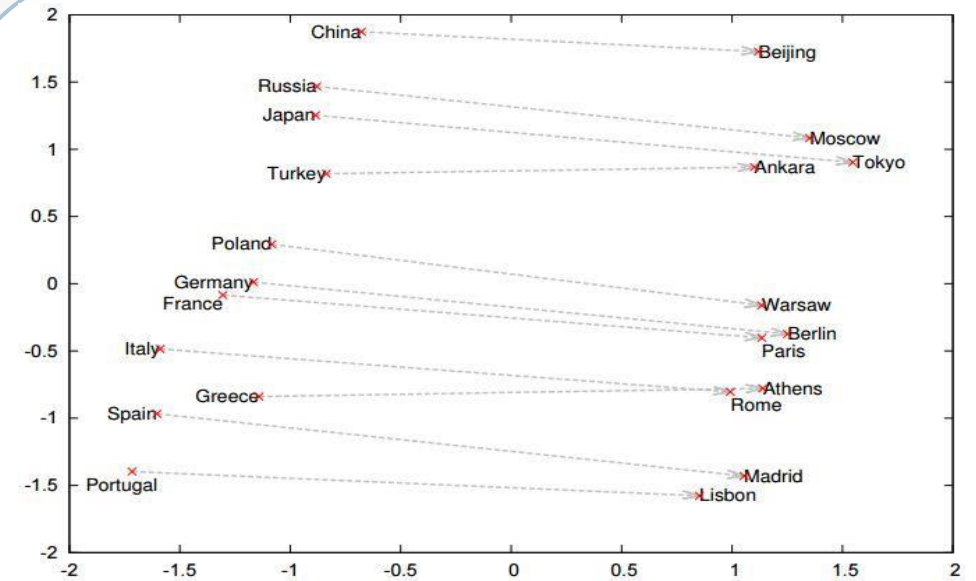
- ✓ A query based system was needed to extract and compare information from financial document corpus consisting of 10K and 10Q filings.
- ✓ This was then used for the creation of trading strategies.

Deliverables identified

- ✓ **Word Vectors**
 - The word embeddings capture meaningful syntactic and semantic regularities.
 - The regularities are observed as constant vector offsets between pair of words sharing a relationship
- ✓ **Document Vectors**
 - Document vectors are extension of word vector.
 - A trained doc2vec model can be used to score new documents.
 - The cosine of angle between two document vectors obtained from Doc2Vec model gives the similarity score
- ✓ **Information Retrieval Engine**
 - Each of the document in the financial corpus is scored.
 - The queries are converted into vectors
 - Using document similarity information retrieval engine is made.

Impact / Result

- ✓ The retrieval system allowed all users to interact with data using user-friendly commands and handled all pre-processing, training and inference underneath.
- ✓ The trading strategy created using the document similarity index gave an alpha of 30 basis points.



“We Just Do Common Things Uncommonly Well”



USA OFFICE

1067 N Mason Rd., Ste 2, Saint Louis,
Missouri, 63141

Phone: +1 314-548-2200

INDIA OFFICE

Ground and 4th Floor, Plot Number 404 – 405,
Udyog Vihar Phase 3,Gurgaon, Haryana 122016

Phone: +91-880-099-8042

GLOBAL CLIENT SUPPORT

Phone: +91 124 626 6300

Email: info@geminisolutions.in