



Jiffy Internship

Internal review - Team 3

Presented by : Maria Soman
Sanchit Sreekanth
Sreejith S.
Sreevisakh V.

Guided by: Shilpa ma'am
Jiffy.ai

Task 1

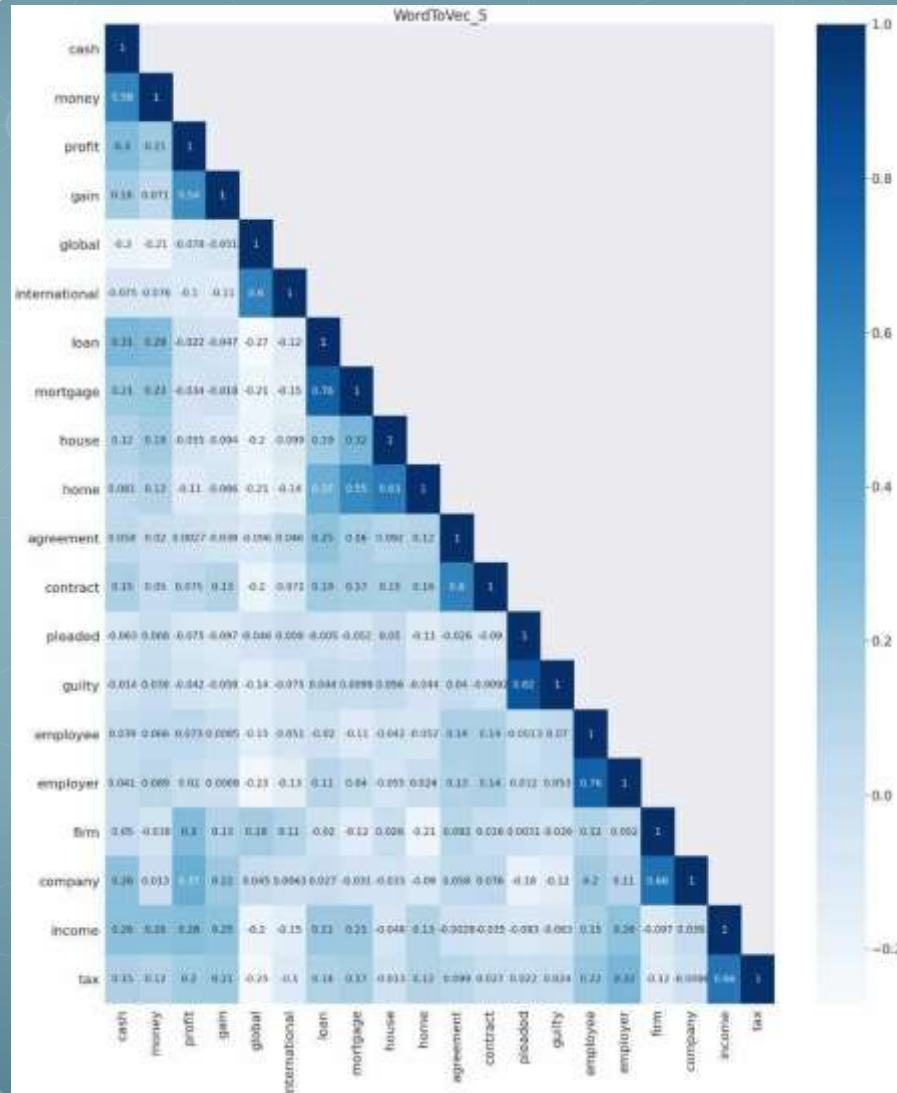
- We developed domain-specific word embeddings for word, phrase and sentence classification tasks in the financial domain.
- Subtasks:
 1. Data Collection- Data was scrapped and collected from financial websites like Investopedia, using BeautifulSoup4 library.
 2. Word Embedding – After organizing and cleaning, the dataset was modelled using Word2Vec/GloVe to vectorize the financial words into a vector space.

Results

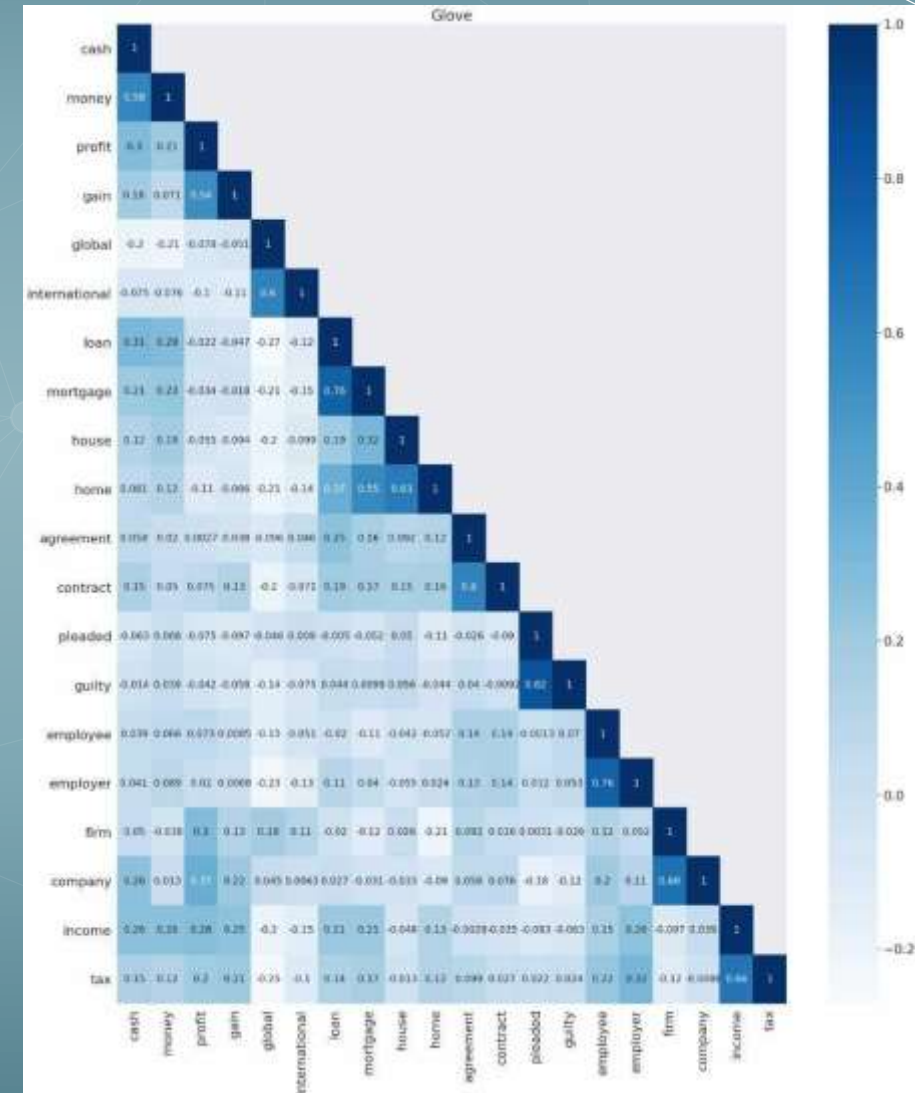
- We have made our comparison with two word2vec and two GloVe models.
- As inferred from the table below, the GloVe model gave better results.

word1	word2	word2vec_5	word2vec_3	glove	glove1
cash	money	0.57820755	0.5993287	0.6222339	0.64259094
profit	gain	0.5430716	0.54125273	0.6959306	0.6752343
global	international	0.5998398	0.5401957	0.68960184	0.6652394
house	home	0.6349237	0.59740716	0.7524492	0.7697871
agreement	contract	0.59714466	0.5767704	0.7604899	0.7436485
pleaded	guilty	0.82068413	0.80703974	0.81470835	0.821982
employee	employer	0.758748	0.74721515	0.88157576	0.86155283
firm	company	0.67777026	0.7051313	0.7902877	0.8086983
income	tax	0.6581138	0.64398324	0.7765005	0.79058427

Comparison of the models



Word2Vec



GloVe

Task 2

- We developed a helper module for financial application development (using FinancePy and yfinance packages in Python).
- If the user/developer types in a word or phrase, the helper module lists the relevant functions and classes related to that word or phrase in the financial package.
- For finding similarity, the custom GloVe embeddings which we created are used.

The background is a solid blue color with a subtle pattern of light blue dots connected by thin white lines, creating a network or molecular structure. In the center, the words "THANK YOU" are written in a bold, white, serif font. The text is slightly shadowed, giving it a three-dimensional appearance as if it's floating within the network. The overall aesthetic is clean, modern, and professional.

THANK YOU