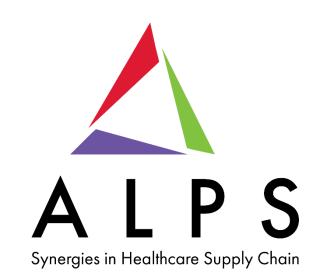


Application of Semantic Textual Similarity in Contract Review

AUTHORS:

- Ge Zhuo Ran, Financial Analyst
- Wang Min Hao, Assistant Director, Finance



Introduction

ALPS handles a large number of vendor contracts. All of the contracts contain "Master Agreement" as the key legal documentation.

The master agreement is more than 20 pages long, with more than 40 clauses. It would take a legally trained person a few hours to complete a review.

We developed an application with artificial intelligence and natural language processing (NLP) techniques to improve the process.

The application can intelligently find out terms and clauses that are "out of specification" from standard contract templates within minutes.

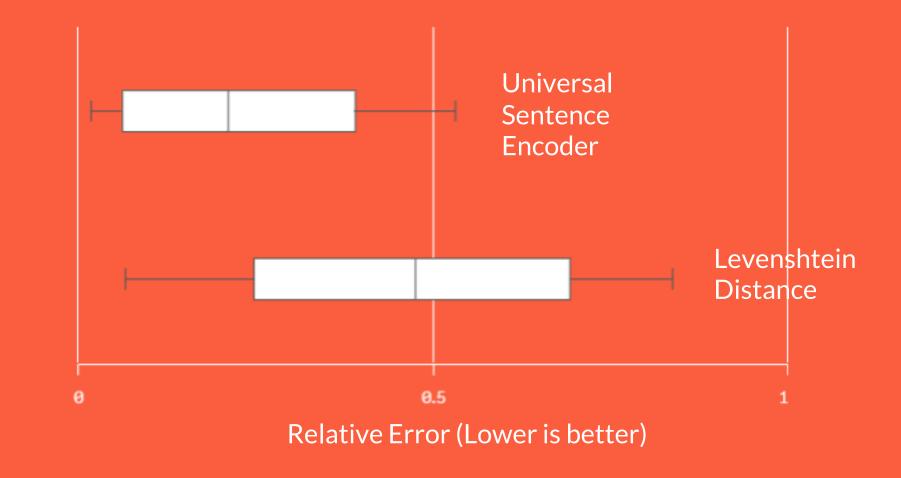
Methodology

OCR Split Encode Compare Score

- 1) Use OCR to extract text from signed contracts pdf file.
- 2) Split the contracts into clauses and then sentences.
- 3) Encode sentence into high dimensional vectors with a deep learning model (Google Universal Sentence Encoder). The model is trained to abstracts semantic meanings from texts.
- 4) For each clause, perform a two-way semantic similarity comparison between the contract and a standard template at sentence level using the high dimensional vectors.
- 5) Score each sentence in the contract. Aggregate the score for each clause and the entire contract.

Results

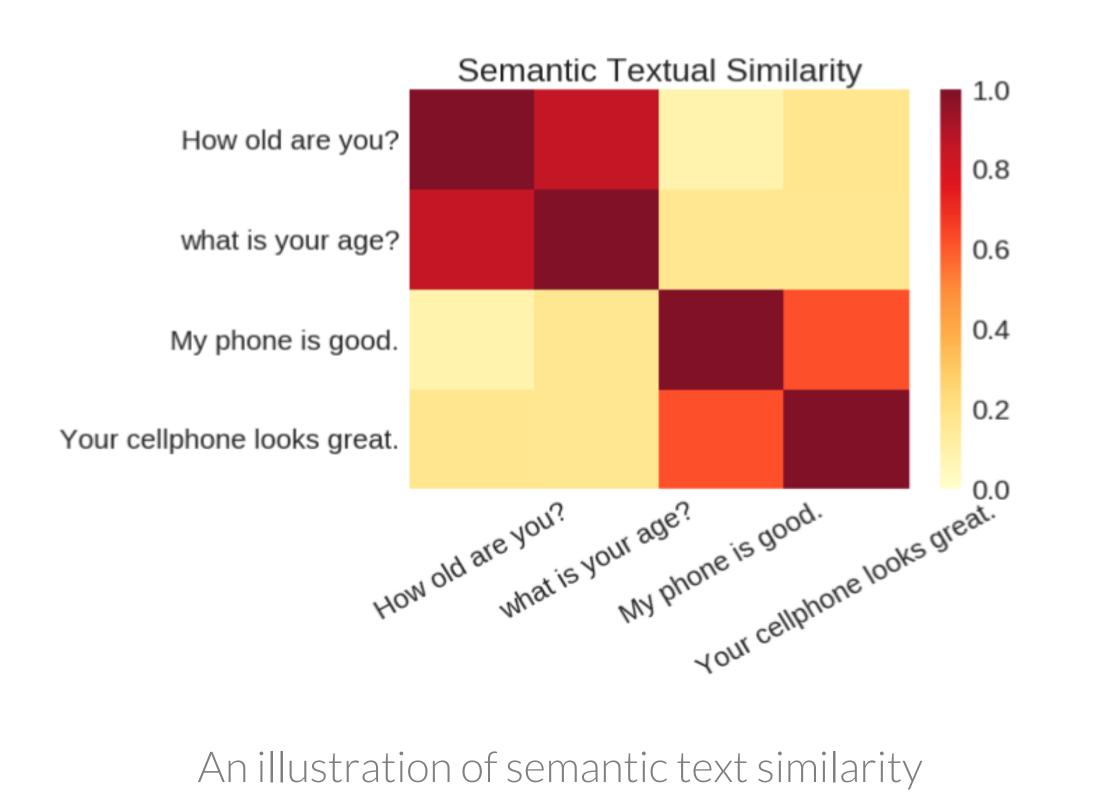
1) Evaluated with 600 sentence pairs from ALPS contracts, the semantic text similarity is near 2x more effective, as compared to the non-deep learning text similarity method.

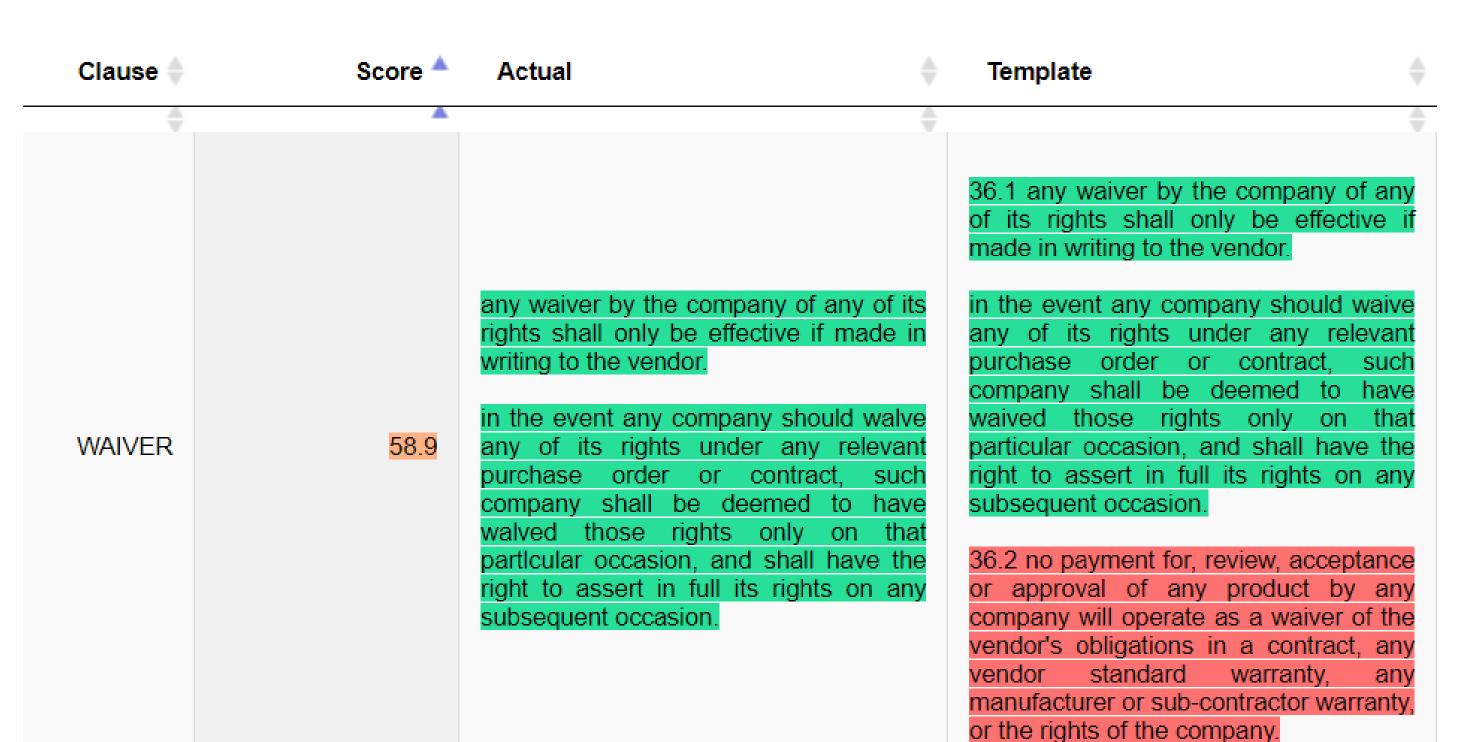


2) We tested the application with 22 actual contracts. The analysis finished in 30 mins and detected significant out of specification clauses in 2 out of the 22 contracts.

Conclusion

The application demonstrated the efficiency and improvement NLP techniques could bring to niche areas, such as the contract review process in supply chain management. Inspired by this application, other NLP applications such as text summarization and question and answering can be explored in future.





A demonstration showing the application detects missing term (red colour) in actual contract's "Waiver" clause, as compared to the template.