ASME IDETC-CIE 2022
International Design Engineering Technical Conferences & Computers and Information in Engineering Conference

Semantic Search with Sentence-BERT for Design Information Retrieval

Hannah S. Walsh, PhD

Sequoia R. Andrade

Semantic Search with Sentence-BERT for Design Information Retrieval

- **Motivation:**
  - Idea is to leverage knowledge contained in **large amounts of data in natural language format from past projects** to **improve new designs**
  - However, with the **sheer volume** of information available, it can be challenging to identify information relevant to a specific information need
  - Lacking a way to leverage the information available, these repositories may go **underutilized**
  - **Information retrieval (IR)** is the task of obtaining relevant search results from a collection of documents (or other resources)

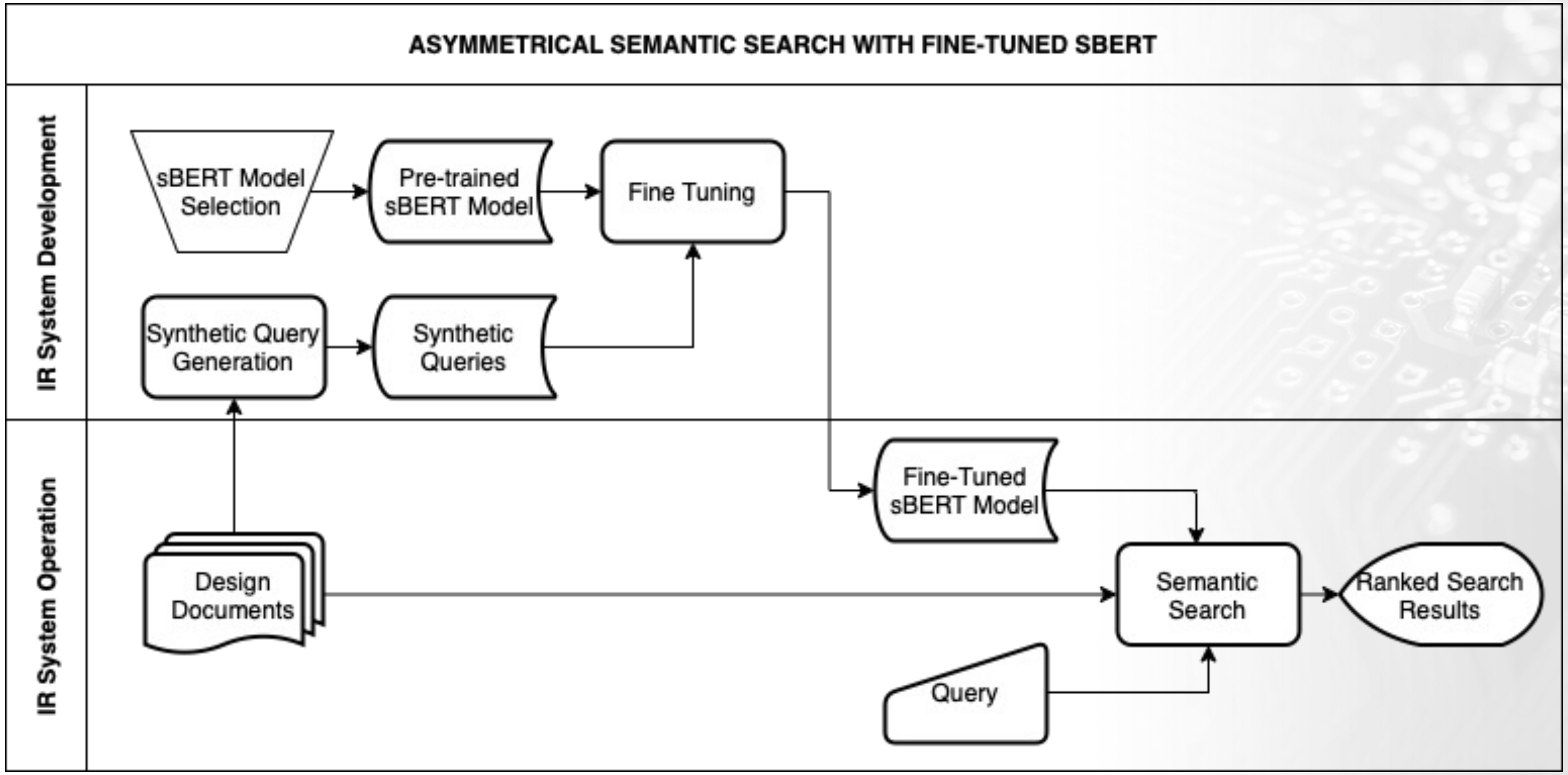
Background: Design Information Retrieval

- Possible IR methods include: **indexing**[1], **hierarchical thesauri**[2], and **ontologies**[3]
- **Semantic network**[4] methods are popular in design information retrieval, including the SemanticOrganizer[5] at NASA
- New and emerging methods in artificial intelligence include **Bidirectional Encoder Representations from Transformers (BERT)**[6] models, which have recently revolutionized natural language processing
- **Sentence-BERT, or sBERT**[7], models are particularly suited to information retrieval tasks
- **Fine-tuning** produces sBERT models that are specifically tuned for design information retrieval tasks

Semantic Search with Sentence-BERT for Design Information Retrieval

In this research, we fine-tune an sBERT model for design information retrieval. Our use case considers a NASA project manager capturing lessons learned relevant to a current project from NASA's Lessons Learned Information System (LLIS). We find the fine-tuned sBERT model performs well using standard information retrieval metrics, and that it outperforms the pretrained (not fine-tuned) model.

Semantic Search with Sentence-BERT for Design Information Retrieval: Overview



NASA Lessons Learned Information System (LLIS)

- https://llis.nasa.gov
- Publicly available database of lessons learned at NASA
- 2096 lessons at time of writing
- For internal users, the database can be downloaded as a spreadsheet
- At right: Lesson 630

Subject

Space Flight Thermal Test Considerations

Abstract

None

Driving Event

When performing ground based functional testing of space flight hardware, it is difficult to perform accurate thermal testing. This is because natural hot air convection is present during ground based testing, and it is not a factor in the microgravity environment of space flight.

There have been several instances where this factor has been forgotten and where the resultant thermal testing was not representative of the in-flight conditions.

Lesson(s) Learned

When performing ground based thermal testing of space flight hardware, care has to be taken to remove the effect of natural hot air convection.

One of the most practical ways to accomplish this is to perform the thermal tests in an oven or thermal chamber where the overall environmental temperature can be raised to be the maximum temperature allowable. In this way, the effect (both heat load and performance) of the hardware being tested can be identified and evaluated.

Recommendation(s)

See lesson(s) learned.

Use Case and Query Formulation

- Our use case is a NASA project manager capturing lessons learned relevant to a current project from NASA's Lessons Learned Information System (LLIS).
- We choose three information needs, inspired by real requirements from NASA's Human Landing System (HLS) project:
  - Cyber Security: "cyber security data and systems"
  - Fault Tolerance: "fault tolerance methods"
  - Fault Isolation: "fault isolation methods"

Search Formulation

- Symmetric vs. **asymmetric**?
- **Passages** vs. documents?

Sentence-BERT Model Selection

- Pretrained sBERT models are available on Huggingface: https://huggingface.co
- Symmetric vs. **asymmetric**?
- Tuned with dot product, cosine similarity, or **normalized**?
- **English Language** vs. other?
- Performance?

Fine-Tuning

- Using LLIS documents to fine-tune the model
- For **asymmetric** search, we need query-result pairs, where a query is the short text a user searches and the result is the document the IR system returns
- LLIS documents are not natively in query-result pairs, but we can instead generate **synthetic queries** using **docTTTTTquery**[8]
- 31,530 query-result pairs are generated (156 minutes) and used to fine-tune the model (31 minutes) given moderate GPU

Synthetic Query	Lesson Excerpt
What can you test for HPH	Trace contaminants in high-purity hydrazine (HPH) propellant impact a wide variety of commercial, Department of Defense (DoD), and NASA missions. Depending on thruster design, contaminants must be kept at extremely low levels and are verified as such by routine analysis...
What would happen if the propulsion subsystem fail	Propulsion subsystem check valves on the Juno spacecraft malfunctioned during preparations for a bi-propellant main engine orbital maneuver. Although the failure mechanism had no major impact on the Juno mission, it poses a risk that an engine may operate outside of its qualified mixture ratio, which could lead to mission loss...
Why did my VFM go wrong during welding	A failure occurred during the first attempt at welding of the Europa Clipper Venturi Flow Meter (VFM) flight units. During the first pass, excessive heat input to the welding area caused the weld root reinforcement material to melt. This left a divot on the top surface and an obstruction in the internal flow passage of the VFM...

Semantic Search

- Fine-tuned model is used to generate embeddings for the LLIS corpus (30 seconds) – these embeddings are **stored** and referenced for each search
- Embeddings are also calculated for each query (~0.086 seconds)
- The search itself takes 0.014 seconds, for a total time of **0.1 seconds per search** in our setup

Query 1: Cyber security measures for data and systems

Fine-Tuned Model

Rank	Score	Lesson	Lesson Title	Passage Excerpt
1	0.517	1250	Network Security/ Reduction of Vulnerabilities/ Penetration Exercises	The terrorist attacks on September 11 emphasized the need for increased security of all national assets including NASA's computer systems...
2	0.513	1175	Computer Hardware-Software/ System Security/Personnel Awareness and Training	16a. Complete and maintain security plans for all appropriate computer systems and ensure that the computer security program is sustaining...
3	0.469	1250	Network Security/Reduction of Vulnerabilities/Penetration Exercises	Accelerate the schedule of penetration exercises to gain greater insights into computer security vulnerabilities...

IR Performance Metrics

- Precision and recall for IR:
  - RET = documents retrieved by the IR system
  - REL = documents relevant to the query (assessed by human)

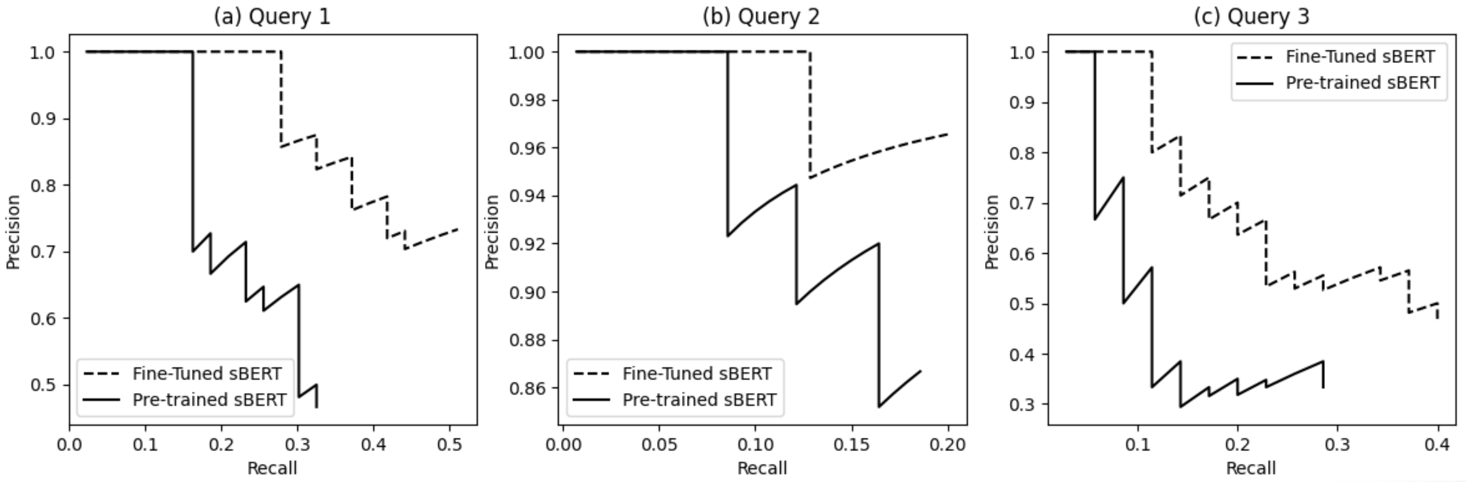
$$P = \frac{RET \cap REL}{RET}$$

$$R = \frac{RET \cap REL}{REL}$$

- Ranked retrieval metrics:
  - Precision at *k*: P at *k*=10, 20, 30,…
  - Mean average precision (MAP): requires complete list of REL
    - n_q: number of queries tested
    - $n_{d,j}$: number of relevant documents for query *j*
    - P_i: precision at *i*

$$MAP = \frac{1}{n_q} \sum_{j=1}^{n_q} \frac{1}{n_{d,j}} \sum_{i=1}^{n_{d_j}} P_i$$

Results



IR Method	MAP
Pre-trained sBERT	0.648
Fine-tuned sBERT	0.807

Query 1: cyber security of data and systems			
	Precision at:		
IR Method	$k = 10$	$k = 20$	$k = 30$
Pre-trained sBERT	0.700	0.650	0.466
Fine-tuned sBERT	1.000	0.800	0.733
Query 2: fault tolerance methods			
	Precision at:		
IR Method	$k = 10$	$k = 20$	$k = 30$
Pre-trained sBERT	1.000	0.900	0.866
Fine-tuned sBERT	1.000	0.950	0.965
Query 3: fault isolation methods			
	Precision at:		
IR Method	$k = 10$	$k = 20$	$k = 30$
Pre-trained sBERT	0.400	0.350	0.333
Fine-tuned sBERT	0.700	0.550	0.466

Wrap Up & Future Directions

- Devised an IR system using a fine-tuned sBERT model that performs well using standard ranked results IR performance metrics and retrieves results in 0.1 seconds
- Fine-tuning significantly outperforms the baseline model and is possible in a few hours given modest GPU
- The completed model is usable for this use case and similar, or it can be further fine-tuned for other design information retrieval tasks

ASME
IDETC-CIE 2022
International Design Engineering Technical Conferences & Computers and Information in Engineering Conference

Thank You!

hannah.s.walsh@nasa.gov

1. Ruecker, L., and Seering, W. P., 1992. "User-oriented intelligent indexing and retrieval in a design documentation system". Vol. ASME 1992 6th Annual Database Symposium: Engineering Data Management — Key to Integrated Product Development of International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, pp. 21–28.
2. Wood, W. H., Yang, M. C., Cutkosky, M. R., and Agogino, A. M., 1998. "Design information retrieval: Improving access to the informal side of design". Vol. Volume 3: 10th International Conference on Design Theory and Methodology of International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. V003T03A001.
3. Li, L., Qin, F., Gao, S., and Qin, X., 2014. "Ontology-based design rationale retrieval supporting natural language query". Vol. Volume 1B: 34th Computers and Information in Engineering Conference of International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. V01BT02A017.
4. Han, J., Sarica, S., Shi, F., and Luo, J., 2020. Semantic networks for engineering design: a survey. arXiv. https://arxiv.org/abs/2012.07060.
5. Keller, R., Berrios, D., Carvalho, R., Hall, D., Rich, S., Sturken, I., Swanson, K., and Wolfe, S., 2004. "SemanticOrganizer: A customizable semantic repository for distributed NASA project teams". In The Semantic Web - ISWC 2004: Third International Semantic Web Conference, Vol. 3298, pp. 767–781.
6. Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K., 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv. https://arxiv.org/abs/1810.04805.
7. Reimers, N., and Gurevych, I., 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. arXiv. https://arxiv.org/abs/1908.10084.
8. Nogueira, R., Lin, J., and Epistemic, A., 2019. "From doc2query to docTTTTTquery". Online preprint