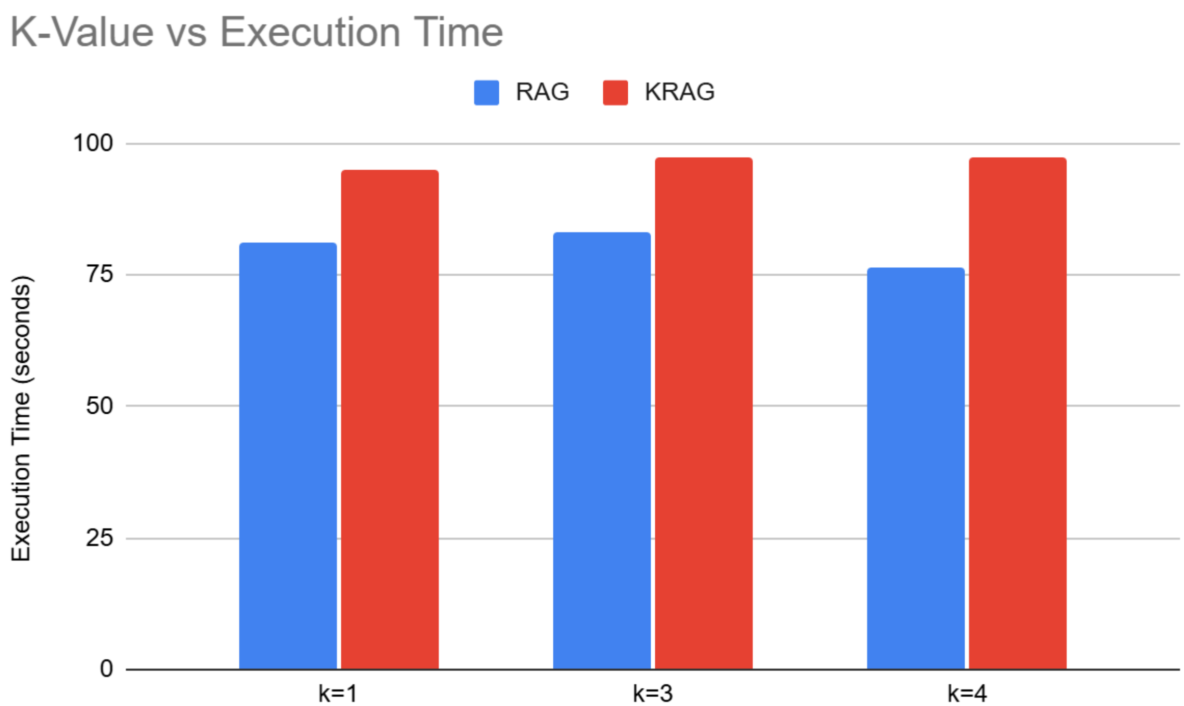
Exam 2

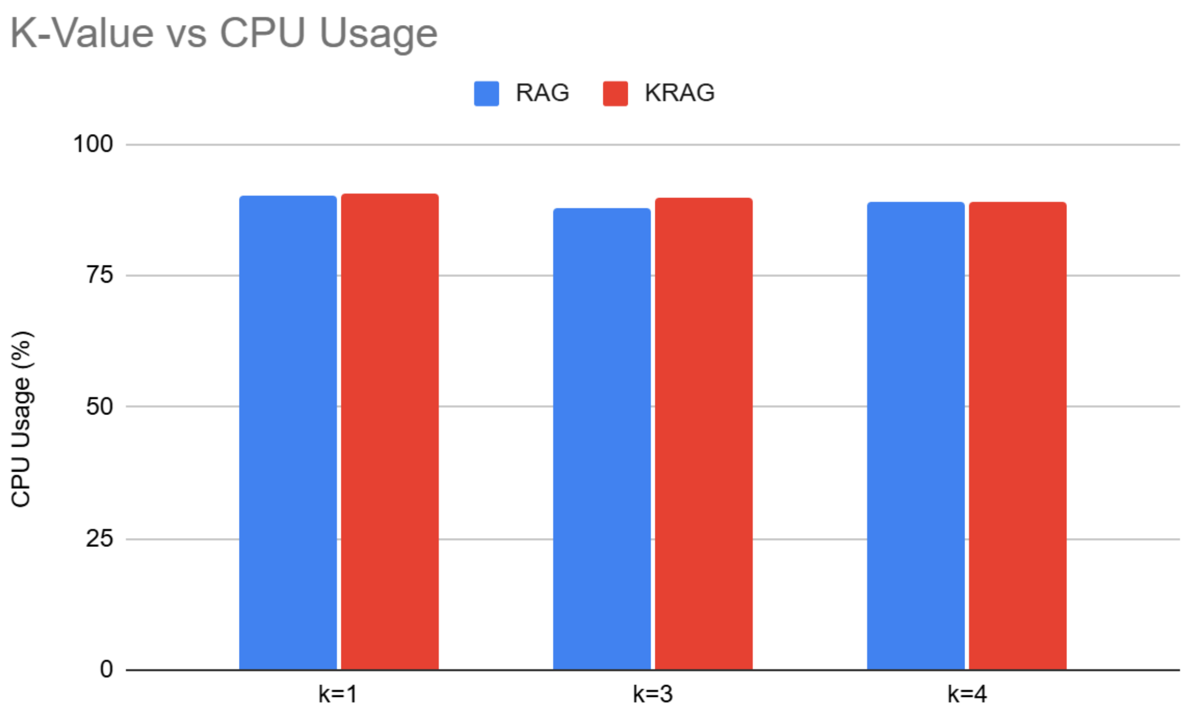
CS5393-004
Gregory Weinrich
Repository: https://github.com/gweinrich/python_exam_3

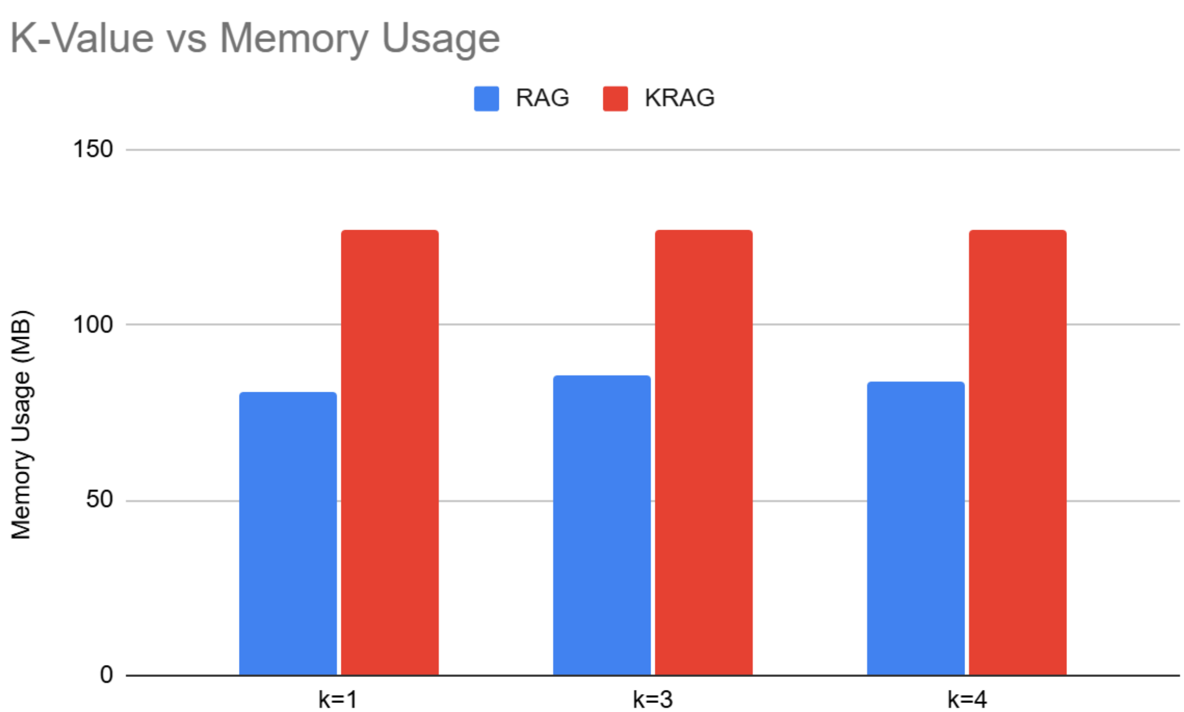
For this project, I decided to compare RAG and KRAG systems. A KRAG system is unique, as it requires several steps in order to create the necessary knowledge graph. Named entity recognition (NER) is a way to train a knowledge graph. The training data must be annotated in order to identify entities in the text and assign them categories. This way, the knowledge graph can build relationships accordingly, grouping together entities based on category. Relationship extraction involves then identifying named entities within the text and adding a textual description to define their relationship. Given two entities found within the same text of a JSON, it can be assumed they are related and a predicted relationship will be chosen from the set of identified relationship types found in the document. Finally, the construction of the knowledge graph involves taking these named entities and representing them as nodes. Then, any relationships previously identified between them will lead to an edge connecting the nodes, with a text descriptor specifying the relationship type.

RAG and KRAG operate similarly, as they both generate responses based on relevant data retrieved from a given context. However, RAG uses vector embeddings to identify words and phrases with similar meanings. It can compare the text found in the context with the prompt and find data with similar meaning according to the vector embedding. However, KRAG uses a knowledge graph. This graph maps the relationships between words and phrases. This allows for descriptors to be included when defining the relationships, thereby allowing for greater precision in data retrieval. The graph is structured such that words and phrases are nodes and the relationships connecting them are edges. As opposed to a vector embedding, a KRAG system allows for explanation behind connections, theoretically providing more relevant responses than a mere similarity check could.

My experimentation focused on the number of documents referenced. For both the RAG and KRAG tests, I used the same prompt, which is as follows: “What engineering experience is described in the documents?" Between each trial, I altered the value of k, a parameter for the similarity_search() function. This function selects k documents based on their relevance to the prompt. I tested k equaling 1, 3, and 4. I originally intended to test with k equaling 5 as well, however ChatGPT would reach its token limit with that many documents. For each trial, I recorded the response, the execution time, the CPU usage, and the memory usage. When looking at the trends in the data, as depicted in the charts below, the execution time and CPU usage of both RAG and KRAG were comparable, with KRAG being slightly greater in both metrics. However, KRAG used a significantly larger amount of memory than RAG. This is likely because of the complex knowledge that must be produced. For both systems, their answers became more informative as k increased. Notably, a lower k value led to more precise references, as the models would have fewer documents as context. However, a greater k value led to responses that included a wider variety of information that was more reflective of the whole corpus. Regarding the quality of the responses, KRAG tended to have higher quality responses, as it was able to itemize each skill found. It was also able to identify and describe a wider variety of skills, such as those that weren't necessarily inherently technical.







One of the biggest challenges I faced was finding properly formatted data for training. The original training data I used did not use entities, but rather attributes. This made building the

NER graph difficult, even after attempting to add code to reformat the data. Eventually, I found a new dataset that correctly utilized entities. Both of these training data sets were based on resumes, allowing me to use the same corpus of engineering resumes for testing.

The most obvious use for this technology would be to automatically filter and assess resumes. Given that this example is trained on and uses resumes, a company could implement a KRAG system to automatically find candidates that fit certain criteria. This could be improved upon with larger training datasets that feature skills and requirements more relevant to the company or a given position.

Output:
Question: “What engineering experience is described in the documents?"

RAG k = 1:
The documents describe experience in site engineering, specifically in the assembly and quality control of HP ProLiant servers and 3PAR storage systems. The individual also has experience in data entry, records management, and technical support. They have worked with various types of equipment and have strong skills in troubleshooting and problem-solving. They also have a Bachelor of Science in Military Science Engineering.
Execution Time: 81.3455 seconds
CPU Usage: 90.10%
Memory Usage: 81.0234 MB

RAG k = 3:
The documents describe experience in various engineering fields, including site engineering, manufacturing quality, controls and instrumentation, equipment maintenance, systems engineering, failure analysis, and software systems engineering. The experience includes tasks such as assembly, troubleshooting, problem-solving, data entry, and coordinating with teams. The documents also mention experience with specific tools and software, such as Microsoft Office, Auto CAD, and Oracle.
Execution Time: 83.2700 seconds
CPU Usage: 88.00%
Memory Usage: 85.6523 MB

RAG k = 4:
The documents describe experience in various engineering roles, including site engineering, technical skills in manufacturing and quality control, project management, systems engineering, and process design. The candidate also has experience in areas such as testing, evaluation, and analysis, as well as cost reduction and avoidance. They have worked with various software and tools, including Microsoft Office, Auto CAD, and statistical process control. The candidate has also been involved in projects related to natural and essential oils, extraction of nutraceuticals, and isolation of chemicals. They have experience in both research and industrial settings, and have worked on teams and independently.
Execution Time: 76.6080 seconds
CPU Usage: 89.20%
Memory Usage: 83.5859 MB

KRAG k = 1:
Result: The documents describe experience in assembly, mechanical inspection, troubleshooting, and problem-solving in the fields of manufacturing, controls and

instrumentation, and equipment maintenance. The individual also has experience with HP and 3PAR servers, data entry, and records management. They have knowledge of Microsoft Office and mass storage units.
Execution Time: 95.0698 seconds
CPU Usage: 90.70%
Memory Usage: 127.2227 MB

KRAG k = 3:
Result:
1. 10+ years total information technology experience both locally and internationally with major information Technology and energy companies.
2. Experience with report all non-conforming parts, materials or drawings to supervisor.
3. Resolve assembly problems, troubleshoot malfunctions and defects of final assembly.
4. Build a defect-free servers unit and or sub-assembly.
5. Complete required in-process inspection, repair any defects found.
6. Complete final inspection documentation and test of equipment to insure highest quality for delivery to the end user.
7. Follow all procedures and processes provided by engineering in accordance to assembly of unit.
8. Document quality information as required to support quality process.
9. Identify any defects found are documented and corrected prior to delivery of unit.
10. Build Quality into all products without exception.
11. Daily/weekly production schedule as determined by scheduler.
12. Ability to improvise when needed.
13. Ability to work with basic math and algebra.
14. Ability to define problems collects data, establish facts, and draw valid conclusions.
15. Attention to detail and ability to complete essential duties and responsibilities in an efficient and timely manner.
16. Positive attitude with management and co-workers.
17. Ability to work under stress.
18. Effective communication and interpersonal skills.
19.
Execution Time: 97.1831 seconds
CPU Usage: 89.90%
Memory Usage: 127.2266 MB

KRAG k = 4:
Result:
1. The candidate has 10+ years of experience in information technology and energy companies.

2. They have experience with report all non-conforming parts, materials or drawings to supervisor.
3. They have experience with assembly and mechanical work, troubleshooting malfunctions and defects.
4. The candidate has experience with building defect-free servers and completing required inspections and tests.
5. They have experience with identifying and correcting defects before delivery.
6. The candidate has experience with daily/weekly production schedules and the ability to improvise when needed.
7. They have experience with basic math and algebra.
8. The candidate has experience with manufacturing quality, controls and instrumentation, equipment maintenance, and mechanical inspection tools.
9. They have experience with problem solving, judgment, and job knowledge.
10. The candidate has experience with Microsoft Office (Word, Excel, Outlook, Power Point), inventory control, and data entry.
11. They have experience with mass storage methods and training.
12. The candidate has experience with coordinating with test and build teams.
13. They have experience with HP server repair and testing, including ProLiant Blade, Tower, and Rack servers.
14. The candidate has experience with 3PAR rack and server assembly, including F200, F400, T400,

Execution Time: 97.1897 seconds

CPU Usage: 89.10%

Memory Usage: 127.1094 MB