

NEURAL NETWORKS

CAIRO UNIVERSITY

FACULTY OF ENGINEERING

EECE

Assignment 4: General Information Retrieval (including Q/A) And Benchmarking among LLMs

Submitted to:

DR. Mohsen Rashwan

Name	ID
ريم محمود محمد عزت	9210430
سلمى محمد حامد مصطفى	9210480
يوسف هشام عبدالفتاح محمد ابوزيد	9211451

Table of Contents

1	Problem 1: General Information Retrieval (including Q/A)	2
1.1	Introduction	2
1.2	Methodology	2
1.2.1	Retrieval System	2
1.2.2	RAG System	3
1.3	Results	3
1.3.1	Retrieval System Results	3
1.3.2	Analysis Of Retrieval System Results	12
1.3.3	RAG System Results	13
1.3.4	Analysis Of Retrieval-Augmented Generation (RAG) Results	20
2	Problem 2: Benchmarking Arabic ASR	21
2.1	Checklist Table	21
2.2	Methodology	22
2.2.1	Evaluation	22
2.2.2	Dataset Preparation	22
2.2.3	Used Models	23
2.3	Results	24
2.3.1	Results Table	24
2.3.2	Results Analysis	24

1 Problem 1: General Information Retrieval (including Q/A)

1.1 Introduction

In this problem we explore information Retrieval and Question Answering Using Arabic Language. This problem introduces keyword-based retrieval, semantic search using embeddings, and Retrieval-Augmented Generation (RAG). We have chosen an Arabic book "الف ليلة وليلة" to work with, As it is a book rihv with stories that can be unknown to LLM so I can get better performance with RAG system.

1.2 Methodology

- .1 **Book Selection:** We utilized the The_Arabic_E-Book_Corpus dataset from the datasets library, hosted on Hugging Face. This corpus contains a collection of Arabic e-books in plain text format. We selected `ليلة ألف ليلة` (One Thousand and One Nights), due to its cultural significance and rich narrative content, which provides a suitable context for our task.
- .2 **Text Splitting:** The book's text was split into short paragraphs. Using a Python function to read the text and segment it into sentences, It splits on `([! .]=>?)\s+`. It is splitted into paragraphs of 4-2 sentences, with a maximum of 3 sentences per paragraph.
- .3 **Embedding Generation:** To enable semantic search, we generated embeddings for each paragraph using the `distiluse-base-multilingual-cased-v2` model from the `sentence_transformers` library.

1.2.1 Retrieval System

The retrieval system was designed to process Arabic queries and return relevant paragraphs from the preprocessed text of `ليلة ألف ليلة` using two distinct methods: classical keyword-based search and semantic search.

- .1 **Data Loading:** The system begins by loading the preprocessed paragraphs from `splitted_book.txt`, which contains the text of the book segmented into short paragraphs. The embeddings for these paragraphs, stored in `embeddings.npy`, and the FAISS index, stored in `faiss_index.bin`, are also loaded. The `distiluse-base-multilingual-cased-v2` model from the `sentence_transformers` library is initialized to encode query embeddings for semantic search.
- .2 **Classical Search (TF-IDF):** For the classical search, we employed the TF-IDF (Term Frequency-Inverse Document Frequency) method using the `TfidfVectorizer` from the `scikit-learn` library. The vectorizer was fitted to the paragraphs to create a TF-IDF matrix representing the importance of words in each paragraph. For a given Arabic query, the query is transformed into a TF-IDF vector, and cosine similarity is computed between the query vector and the paragraph matrix. The top 5 paragraphs with the highest similarity scores are retrieved, along with their scores, for display.
- .3 **Semantic Search (Embedding-Based):** The semantic search leverages the precomputed embeddings and the FAISS index. When a query is received, it is encoded into a 512-dimensional embedding using the same `distiluse-base-multilingual-cased-v2` model used for the paragraphs. The FAISS FlatL2 index performs a nearest neighbor search to identify the top 5 paragraphs whose embeddings are closest to the query embedding, based on L2 distance. The retrieved paragraphs and their corresponding distances are returned.

- .4 **User Interface and Result Display:** An interactive command-line interface was implemented in Python to accept Arabic queries from the user. The interface prompts the user to enter a query or type 'quit' to exit. For each valid query, both classical and semantic searches are executed, and the top 5 results from each method are displayed side by side. Each result includes the paragraph text and its associated score (cosine similarity for TF-IDF, L2 distance for semantic search). The output is formatted with clear separators and labels in Arabic, ensuring readability and ease of comparison.

1.2.2 RAG System

The Retrieval-Augmented Generation (RAG) system extends the retrieval interface to answer Arabic questions by combining retrieved context with a large language model (LLM).

- .1 **LLM Setup:** The system utilizes the gpt-4o-mini model API from open AI. Configured with temperature=0.0 for reproducibility and max tokens = .128
- .2 **Query Processing:** The interactive interface, was extended to handle question-answering. Users input Arabic questions, and the system checks for valid input, allowing termination with the command 'quit'. For each question, the system performs both classical (TF-IDF) and semantic (embedding-based) searches to retrieve relevant paragraphs, as described in the retrieval system methodology.
- .3 **Context Retrieval:** For the RAG approach, the system selects the top 3 results from both search methods to serve as context. If The result paragraph is number i in the book paragraphs i-1, i and i+1 are passed to the model as context to ensure better understanding for the content.
- .4 **Answer Generation:** The ask_arabic function processes the question with or without context. Answers are generated for each question:
 - **Baseline (LLM-only):** The question is answered without any retrieved context.
 - **Classical RAG:** The question is answered with the first 3 TF-IDF search results as context.
 - **Semantic RAG:** The question is answered with the first 3 semantic search results as context.
- .5 **Result Display:** The system displays the generated answers.

1.3 Results

1.3.1 Retrieval System Results

Query :1 من هم شهريار وشاهزمان

- .1 **Classical Search Result:** يُحكى أنه في قديم الزمان عاش ملكان عظيمان شقيقان يدعيان شهريار وشاهزمان. كان شهريار هو الأكبر سنًا، وامتد سلطانه إلى أقاصي الأرض. شعر شهريار برغبة في رؤية أخيه، فطلب من كبير وزرائه الذهاب لإحضاره إليه
Score: 0.32989686491101883
- Semantic Search Result:** قصة الملك شهريار وشهرزاد
Distance: 1.081768274307251

2. Classical Search Result: قصة الملك شهريار وشهرزاد

Score: 0.27742265750562584

Semantic Search Result: يُحكى أنه في قديم الزمان عاش ملكان عظيمان شقيقان يُدعيان شهريار وشاهزمان. كان شهريار هو الأكبر سنًا، وامتد سلطانه إلى أقاصي الأرض. شعر شهريار برغبة في رؤية أخيه، فطلب من كبير وزرائه الذهاب لإحضاره إليه.

Distance: 1.3581560850143433

3. Classical Search Result: مرت الأيام، وشاهزمان يزداد نحوًا، واعتقد شهريار أن أخاه مشتاق لوطنه، لكن تبين أن الأمر أسوأ من ذلك بكثير. فبعد عشرة أيام، انهار شاهزمان وأخبر أخاه بالسبب الرئيسي وراء أحرانه. لقد هجرته زوجته الملكة، وهو الآن كسير الفؤاد.

Score: 0.17847705713373826

Semantic Search Result: قصة الصياد والجني

Distance: 1.4182838201522827

4. Classical Search Result: في تلك الليلة، اصطحب الوزير شهرزاد إلى الملك العظيم شهريار، وزوجها له. لكن عندما ذهب الملك شهريار للنوم في تلك الليلة، بكت شهرزاد. وعندما سأها الملك عن سبب بكائها، أجابت: «لديّ أخت، وأود أن أودعها.»

Score: 0.155031133039674

Semantic Search Result: في تلك الليلة، اصطحب الوزير شهرزاد إلى الملك العظيم شهريار، وزوجها له. لكن عندما ذهب الملك شهريار للنوم في تلك الليلة، بكت شهرزاد. وعندما سأها الملك عن سبب بكائها، أجابت: «لديّ أخت، وأود أن أودعها.»

Distance: 1.437458872795105

5. Classical Search Result: ظل الملك شهريار يعمل طوال اليوم التالي، وعاد إلى المنزل في الليل إلى شهرزاد.

Score: 0.14938705215508133

Semantic Search Result: ظل الملك شهريار يعمل طوال اليوم التالي، وعاد إلى المنزل في الليل إلى شهرزاد.

Distance: 1.4516593217849731

Query :2 من الأكبر سنًا بين شهريار وشاهزمان

1. Classical Search Result: يُحكى أنه في قديم الزمان عاش ملكان عظيمان شقيقان يُدعيان شهريار وشاهزمان. كان شهريار هو الأكبر سنًا، وامتد سلطانه إلى أقاصي الأرض. شعر شهريار برغبة في رؤية أخيه، فطلب من كبير وزرائه الذهاب لإحضاره إليه.

Score: 0.4221575816720085

Semantic Search Result: يُحكى أنه في قديم الزمان عاش ملكان عظيمان شقيقان يُدعيان شهريار وشاهزمان. كان شهريار هو الأكبر سنًا، وامتد سلطانه إلى أقاصي الأرض. شعر شهريار برغبة في رؤية أخيه، فطلب من كبير وزرائه الذهاب لإحضاره إليه.

Distance: 1.0115330219268799

2. Classical Search Result: قصة الملك شهريار وشهرزاد

Score: 0.21388455476367854

Semantic Search Result: قصة الملك شهريار وشهرزاد

Distance: 1.0618455410003662

3. Classical Search Result: مرّت الأيام، وشاهزمان يزداد نحولاً، واعتقد شهربار أن أخاه مشتاق لوطنه، لكن تبين أن الأمر أسوأ من ذلك بكثير. فبعد عشرة أيام، انهار شاهزمان وأخبر أخاه بالسبب الرئيسي وراء أحزانه. لقد هجرته زوجته الملكة، وهو الآن كسير الفؤاد.
Score: 0.1376004622110114
Semantic Search Result: مرّت الأيام، وشاهزمان يزداد نحولاً، واعتقد شهربار أن أخاه مشتاق لوطنه، لكن تبين أن الأمر أسوأ من ذلك بكثير. فبعد عشرة أيام، انهار شاهزمان وأخبر أخاه بالسبب الرئيسي وراء أحزانه. لقد هجرته زوجته الملكة، وهو الآن كسير الفؤاد.
Distance: 1.409240961074829
4. Classical Search Result: في تلك الليلة، اصطحب الوزير شهرزاد إلى الملك العظيم شهربار، وزوجها له. لكن عندما ذهب الملك شهربار للنوم في تلك الليلة، بكت شهرزاد. وعندما سأها الملك عن سبب بكائها، أجابت: «لديّ أخت، وأود أن أودعها»
Score: 0.11952435739329209
Semantic Search Result: عندما حل الليل، قالت الشقيقات الثلاث للحمال: «سيدي، لقد حان الوقت لتنهض. فلتتردي خُفيك وترحل». فرد الحمال: «إن خروج روجي من جسدي أهون عليّ من فراقك. دعونا نواصل ما نحن عليه، ثم نفترق في الصباح»
Distance: 1.463319182395935
5. Classical Search Result: ظل الملك شهربار يعمل طوال اليوم التالي، وعاد إلى المنزل في الليل إلى شهرزاد.
Score: 0.11517294018063412
Semantic Search Result: وعند سماع الوزير ما قالته ابنته، استشاط غضباً وحرّم عليها الزواج من الملك. وأخذها يتجادلان معاً فترة طويلة، لكن شهرزاد لم تكن لتغير رأيها، فأرسلها الوزير لتصبح زوجة الملك، وهو يقول: «أدعو الله ألا يحرميني منك»
Distance: 1.4737191200256348

Query: 3 ما السبب الرئيسي وراء أحزان شاهزمان؟

1. Classical Search Result: مرّت الأيام، وشاهزمان يزداد نحولاً، واعتقد شهربار أن أخاه مشتاق لوطنه، لكن تبين أن الأمر أسوأ من ذلك بكثير. فبعد عشرة أيام، انهار شاهزمان وأخبر أخاه بالسبب الرئيسي وراء أحزانه. لقد هجرته زوجته الملكة، وهو الآن كسير الفؤاد.
Score: 0.26246435251832106
Semantic Search Result: قصة الملك شهربار وشهرزاد
Distance: 1.2975513935089111
2. Classical Search Result: ومع حلول العام الجديد، بدأت الفتيات في البكاء، وأخذن يتعلّقن بي قائلات: «لعلك لا تغادر أبداً». فسألتهن عن سبب بكائهن وظنهن أنني سأغادر. فأجبن: «ستكون أنت نفسك السبب». وأخبرني أنهن أربعون ابنة لأربعين ملكاً، ويجب عليهن ترك القصر مرة كل عام مدة أربعين يوماً. وعند مغادرتهن، لن أصاحبهن.
Score: 0.0795256750164715
Semantic Search Result: فرد عليه التاجر التحية. سأله الرجل العجوز: «لماذا تجلس هنا، يا صاحبي، في هذا المكان الممتلئ بالجن والشياطين؟ فما من شيء في هذا البستان المسكون سوى الحزن والأسى»
Distance: 1.353788137435913

3. Classical Search Result: «شبيك لييك، اطلب ما شئت، فأنا عبد من يرتدي هذا الخاتم في إصبعه، أنفذ له ما يشاء.»
Score: 0.06101734520395203
Semantic Search Result: فرأيت حزن الملك يتحول إلى غضب؛ الأمر الذي أدهشني كثيراً.
Distance: 1.3606433868408203
4. Classical Search Result: عندما سمعتُ ما قالته، اندهشتُ، وشكرتها على إنقاذ حياتي. أخبرتها بكل ما حدث لنا منذ وفاة أبنينا، من البداية إلى النهاية. وعندما سمعت قصتي، غضبت بشدة، وقالت: «سأطير الآن إلى شقيقك، وأغرق سفينتهما.»
Score: 0.043767624266596035
Semantic Search Result: تقدم بعد ذلك الدرويش الأول وقال: يا سيدي، إن سبب فقداني لعيني وحلاقة لحيتي على هذا النحو هو أن والدي كان ملكاً، وكان أخوه ملكاً هو الآخر ولديه ابن وابنة. وبمرور السنوات، اعتدت زيارة عمي من حين لآخر. وثمت بيني وبين ابن عمي صداقة قوية.
Distance: 1.3979297876358032
5. Classical Search Result: قال علاء الدين عندما رأى ما حدث: «أيها الجني، أعد إلي زوجتي وقصري.»
Score: 0.04183775827098927
Semantic Search Result: وعندما رآهم، أخذ يبكي بحرقة؛ فسألته زوجته: «ما بك يا زوجي؟ لماذا تشعر بالحزن وجميعنا سعداء ونحتفل بعودتك إلينا؟» فأجاب: «ولماذا لا أحزن وما تبقى لي من العمر سوى سنة واحدة فقط؟» ثم أخبرها بكل شيء حدث مع الجني. عندما سمعت أسرته ما قاله، أخذوا يبكون، وكان يوماً حزيناً حيث تجمع جميع الأطفال حول والدهم.
Distance: 1.3984439373016357

Query :4 لماذا غضب شهريار غضباً شديداً؟

1. Classical Search Result: وبعد سماع ما رواه أخوه، غضب شهريار غضباً شديداً حتى إنه فقد صوابه، وقرر أنه لا يمكن الثقة في أي امرأة قط، وخطط لأمر مريع. استدعى الملك وزيره، وطلب منه أن يبحث له عن زوجة. وكان يعتزم الزواج منها يوماً واحداً ثم قتلها، على أن يستمر في ذلك حتى يقضي على جميع النساء في مملكته.
Score: 0.3034526751754444
Semantic Search Result: فرأيت حزن الملك يتحول إلى غضب؛ الأمر الذي أدهشني كثيراً.
Distance: 1.0196154117584229
2. Classical Search Result: فرأيت حزن الملك يتحول إلى غضب؛ الأمر الذي أدهشني كثيراً.
Score: 0.25120586632808317
Semantic Search Result: قصة الملك شهريار وشهرزاد
Distance: 1.2158721685409546
3. Classical Search Result: لكنني كنت غاضباً غضباً جعلني لا أستمع إليها. وركلت الباب بقدمي.
Score: 0.2183099706829414
Semantic Search Result: فرد عليه التاجر التحية. سأله الرجل العجوز: «لماذا تجلس هنا، يا صاحبي، في هذا المكان الممتلئ بالجن والشياطين؟ فما من شيء في هذا البستان المسكون سوى الحزن والأسى.»
Distance: 1.2283470630645752

4. Classical Search Result: قصة الملك شهريار وشهرزاد

Score: 0.20473163951202672

Semantic Search Result: وبعد سماع ما رواه أخوه، غضب شهريار غضباً شديداً حتى إنه فقد صوابه، وقرر أنه لا يمكن الثقة في أي امرأة قط، وخطط لأمر مريع. استدعى الملك وزيره، وطلب منه أن يبحث له عن زوجة. وكان يعتزم الزواج منها يوماً واحداً ثم قتلها، على أن يستمر في ذلك حتى يقضي على جميع النساء في مملكته.

Distance: 1.2759933471679688

5. Classical Search Result: يُحكى أنه في قديم الزمان عاش ملكان عظيمان شقيقان يُدعيان شهريار وشاهزمان. كان شهريار هو الأكبر سنّاً، وامتد سلطانه إلى أقاصي الأرض. شعر شهريار برغبة في رؤية أخيه، فطلب من كبير وزرائه الذهاب لإحضاره إليه.

Score: 0.1577834085076877

Semantic Search Result: وعندما رآهم، أخذ يبكي بحرقة؛ فسألته زوجته: «ما بك يا زوجي؟ لماذا تشعر بالحزن وجميعنا سعداء ونحتفل بعودتك إلينا؟» فأجاب: «ولماذا لا أحزن وما تبقى لي من العمر سوى سنة واحدة فقط؟» ثم أخبرها بكل شيء حدث مع الجني. عندما سمعت أسرته ما قاله، أخذوا يبكون، وكان يوماً حزيناً حيث تجمع جميع الأطفال حول والدهم.

Distance: 1.279199242591858

Query: 5: لماذا قرر شهريار أنه لا يمكن الثقة في أي امرأة قط؟

1. Classical Search Result: وبعد سماع ما رواه أخوه، غضب شهريار غضباً شديداً حتى إنه فقد صوابه، وقرر أنه لا يمكن الثقة في أي امرأة قط، وخطط لأمر مريع. استدعى الملك وزيره، وطلب منه أن يبحث له عن زوجة. وكان يعتزم الزواج منها يوماً واحداً ثم قتلها، على أن يستمر في ذلك حتى يقضي على جميع النساء في مملكته.

Score: 0.37301669658764786

Semantic Search Result: وبعد سماع ما رواه أخوه، غضب شهريار غضباً شديداً حتى إنه فقد صوابه، وقرر أنه لا يمكن الثقة في أي امرأة قط، وخطط لأمر مريع. استدعى الملك وزيره، وطلب منه أن يبحث له عن زوجة. وكان يعتزم الزواج منها يوماً واحداً ثم قتلها، على أن يستمر في ذلك حتى يقضي على جميع النساء في مملكته.

Distance: 0.8278130888938904

2. Classical Search Result: يُحكى أنه في قديم الزمان عاش ملكان عظيمان شقيقان يُدعيان شهريار وشاهزمان. كان شهريار هو الأكبر سنّاً، وامتد سلطانه إلى أقاصي الأرض. شعر شهريار برغبة في رؤية أخيه، فطلب من كبير وزرائه الذهاب لإحضاره إليه.

Score: 0.1869212300947148

Semantic Search Result: وعند سماع الوزير ما قالته ابنته، استشاط غضباً وحرّم عليها الزواج من الملك. وأخذاً يتجادلان معاً فترة طويلة، لكن شهرزاد لم تكن لتغير رأيها، فأرسلها الوزير لتصبح زوجة الملك، وهو يقول: «أدعو الله ألا يحرمني منك.»

Distance: 1.245424747467041

3. Classical Search Result: قصة الملك شهريار وشهرزاد

Score: 0.17916329703003445

Semantic Search Result: وبعد ذلك، لم يشك الحمال حاله في الدنيا ثانية أبداً.

Distance: 1.2796372175216675

4. Classical Search Result: واستدار الخليفة إلى الدراويش، وقال: «أنتم من أفراد المنزل، ربما يمكنكم تفسير هذه المشاهد الغريبة لنا.» فأجابوا: «نحن لا نعلم شيئاً، ولم نر هذا المكان قط قبل الليلة. هذا الرجل الموجود بجانبك يمكن أن يجيبك.»

Score: 0.16010074395439988

Semantic Search Result: فُتحت الستارة، وظهرت من خلفها امرأة أخرى مبهرة وجهها يفوق الشمس إشراقاً. سارت نحوهم وقالت: «لماذا تقفان هكذا؟ لتخففا الحمل عن هذا الرجل المسكين.»

Distance: 1.2847260236740112

5. Classical Search Result: فُتحت الستارة، وظهرت من خلفها امرأة أخرى مبهرة وجهها يفوق الشمس إشراقاً. سارت نحوهم وقالت: «لماذا تقفان هكذا؟ لتخففا الحمل عن هذا الرجل المسكين.»

Score: 0.1277367628218387

Semantic Search Result: وفي أحد الأيام، والقمر لا يزال في السماء، والنهار لم يطلع بعد، خرج الصياد متجاوزاً أطراف المدينة وصولاً إلى شاطئ البحر، ثم خاض في الماء، وألقى بشبكته، وانتظر حتى غاصت. وجمع الحبل، وبدأ في سحب الشبكة، وشعر أنها تزداد ثقلاً حتى أصبح غير قادر على سحبها أكثر من ذلك. نخلع ملابسه وغطس في الماء، سابحاً تحت الشبكة، وأخذ يهزها ويجرها حتى جلبها أخيراً إلى الشاطئ.

Distance: 1.3000266551971436

Query 6: ما الأمر المريع الذي خطط له شهربار؟

1. Classical Search Result: قصة الملك شهربار وشهرزاد

Score: 0.3334224283628586

Semantic Search Result: قصة الملك شهربار وشهرزاد

Distance: 1.2747981548309326

2. Classical Search Result: يُحكى أنه في قديم الزمان عاش ملكان عظيمان شقيقان يُدعيان شهربار وشاهزمان. كان شهربار هو الأكبر سنّاً، وامتد سلطانه إلى أقاصي الأرض. شعر شهربار برغبة في رؤية أخيه، فطلب من كبير وزرائه الذهاب لإحضاره إليه.

Score: 0.2569633464831983

Semantic Search Result: فأجاب الجني: «تمن أمنية!» سعد الصياد ثانياً، وقال: «ماذا أتمنى؟» فأجاب الجني: «أخبرني كيف تود أن تموت.»

Distance: 1.3963439464569092

3. Classical Search Result: في تلك الليلة، اصطحب الوزير شهرزاد إلى الملك العظيم شهربار، وزوجها له. لكن عندما ذهب الملك شهربار للنوم في تلك الليلة، بكت شهرزاد. وعندما سألها الملك عن سبب بكائها، أجابت: «لديّ أخت، وأود أن أودعها.»

Score: 0.2416717647372499

Semantic Search Result: ظل الملك شهربار يعمل طوال اليوم التالي، وعاد إلى المنزل في الليل إلى شهرزاد.

Distance: 1.411792278289795

4. Classical Search Result: ظل الملك شهربار يعمل طوال اليوم التالي، وعاد إلى المنزل في الليل إلى شهرزاد.

Score: 0.17954190960234082

Semantic Search Result: اصطحب الوزير شهرزاد إلى الملك العظيم شهریار، وزوجها له. لكن عندما ذهب الملك شهریار للنوم في تلك الليلة، بكت شهرزاد. وعندما سألها الملك عن سبب بكائها، أجابت: «لديّ أخت، وأود أن أودعها.»

Distance: 1.495529294013977

5. Classical Search Result: قال الجني: «شبيك لييك، اطلب ما شئت، فأنا عبد من يرتدي هذا الخاتم في إصبعه، أنفذ له ما يشاء.»

Score: 0.1738954365530399

Semantic Search Result: وأخبرته أن الساحر حاول أن يكسب حباً بتقديم الجواهر والذهب لها. استمع علاء الدين إلى كل ما قالته، وفكر بامعان حتى توصل إلى خطة. وقبل زوجته، وأخبرها بما سيفعله.

Distance: 1.5034421682357788

Query: لماذا قالت شهرزاد أود أن تزوجني الملك شهریار؟ 7:

1. Classical Search Result: طالعت الابنة الكبرى كتباً كثيرة واتسمت باتساع ثقافتها. وعندما سمعت بخطط الملك البغيض، قالت لوالدها: «أود أن تزوجني الملك شهریار، حتى أتمكن من إنقاذ شعبنا أو الموت مثل باقي الفتيات.»

Score: 0.3376613456545272

Semantic Search Result: وعند سماع الوزير ما قالته ابنته، استشاط غضباً وحرّم عليها الزواج من الملك. وأخذ يتجادل معاً فترة طويلة، لكن شهرزاد لم تكن لتغير رأيها، فأرسلها الوزير لتصبح زوجة الملك، وهو يقول: «أدعو الله ألا يحرمني منك.»

Distance: 0.9920270442962646

2. Classical Search Result: قصة الملك شهریار وشهرزاد

Score: 0.28310697532991513

Semantic Search Result: في تلك الليلة، اصطحب الوزير شهرزاد إلى الملك العظيم شهریار، وزوجها له. لكن عندما ذهب الملك شهریار للنوم في تلك الليلة، بكت شهرزاد. وعندما سألها الملك عن سبب بكائها، أجابت: «لديّ أخت، وأود أن أودعها.»

Distance: 0.9946215152740479

3. Classical Search Result: في تلك الليلة، اصطحب الوزير شهرزاد إلى الملك العظيم شهریار، وزوجها له. لكن عندما ذهب الملك شهریار للنوم في تلك الليلة، بكت شهرزاد. وعندما سألها الملك عن سبب بكائها، أجابت: «لديّ أخت، وأود أن أودعها.»

Score: 0.23680276185701477

Semantic Search Result: بلغني — أيها الملك — أن علاء الدين وقع في حب ابنة الملك.

Distance: 0.9974219799041748

4. Classical Search Result: ظل الملك شهریار يعمل طوال اليوم التالي، وعاد إلى المنزل في الليل إلى شهرزاد.

Score: 0.19575134996534033

Semantic Search Result: قصة الملك شهریار وشهرزاد

Distance: 1.0369329452514648

5. Classical Search Result: «رجاء يا أختاه، إذا لم يكن النوم يغالبك، فاروي لنا واحدة من قصصك الممتعة.» وأضاف الملك: «لتكن نهاية قصة التاجر، فأنا أود سماعها.» فأجابت شهرزاد: «كما تشاءان.»

Score: 0.1908122663316863

Semantic Search Result: عبس الصياد عند سماعه ذلك، وقال: «لماذا ترغب في قتلي وأنا أعدتك إلى هذا العالم؟»

Distance: 1.0437206029891968

Query :8 من دینارزاد؟

1. Classical Search Result: «عظيم.» فأجابت دینارزاد:

Score: 0.38851371050402356

Semantic Search Result: قصة علاء الدين

Distance: 1.3227245807647705

2. Classical Search Result: «أختاه، يا لها من قصة غريبة ومذهلة!» فردت شهرزاد: «هذا لا يقارن بما سأرويها لكما غداً.»

Score: 0.2270181177871458

Semantic Search Result: «عظيم.» فأجابت دینارزاد:

Distance: 1.3507766723632812

3. Classical Search Result: «يا لها من قصة غريبة وجميلة يا أختاه!» فردت شهرزاد: «هذا لا يقارن بما سأرويها لكما غداً.»

Score: 0.21716601962350177

Semantic Search Result: من يقايض مصباحاً قديماً بواحد من هذه المصباح الجديدة؟»

Distance: 1.3800948858261108

4. Classical Search Result: «رجاء يا أختاه، قالت أختها دینارزاد: «رجاء يا أختاه، اروي لنا المزيد من قصصك الممتعة.» وأضاف الملك: «لتكن ما تبقى من قصة الخمال.» فأجابت شهرزاد: «على الرحب والسعة!»

Score: 0.21239747581780455

Semantic Search Result: مصباح! من يقايض مصباحاً قديماً بواحد من هذه المصباح الجديدة؟»

Distance: 1.4255802631378174

5. Classical Search Result: «رجاء يا أختاه، قالت أختها دینارزاد: «رجاء يا أختاه، اروي لنا المزيد من قصصك الممتعة.» وأضاف الملك: «لتكن أكثر إثارة من القصة السابقة.» فأجابت شهرزاد: «كما تشاءان!»

Score: 0.20743573756800032

Semantic Search Result: بلغني — أيها الملك السعيد — أن الدرويش الثاني قال: عندما رأيتي الفتاة، سألتني: «هل أنت إنس أم جن؟» فأجبتها: «إنس.» وسألت: «ما الذي أتى بك إلى هنا؟ لقد عشت في هذا المكان طيلة خمسة وعشرين عاماً دون أن أرى أحداً على الإطلاق.»

Distance: 1.5217704772949219

Query :9 لماذا قرر شهریار ان يقتلها في اليوم التالي

1. Classical Search Result: ظل الملك شهریار يعمل طوال اليوم التالي، وعاد إلى المنزل في الليل إلى شهرزاد.
Score: 0.40093955995304476
Semantic Search Result: وبعد سماع ما رواه أخوه، غضب شهریار غضباً شديداً حتى إنه فقد صوابه، وقرر أنه لا يمكن الثقة في أي امرأة قط، وخطط لأمر مريع. استدعى الملك وزيره، وطلب منه أن يبحث له عن زوجة. وكان يعتزم الزواج منها يوماً واحداً ثم قتلها، على أن يستمر في ذلك حتى يقضي على جميع النساء في مملكته.
Distance: 1.060058355331421
2. Classical Search Result: بلغني — أيها الملك السعيد — أن الحمال عاد إلى المنزل في اليوم التالي.
Score: 0.30019655150807434
Semantic Search Result: عبس الصياد عند سماعه ذلك، وقال: «لماذا ترغب في قتلي وأنا أعدتك إلى هذا العالم؟»
Distance: 1.1691937446594238
3. Classical Search Result: قصة الملك شهریار وشهرزاد
Score: 0.29607745257341556
Semantic Search Result: لكن الآن يجب أن أقتل شقيقك.
Distance: 1.1793206930160522
4. Classical Search Result: يُحكى أنه في قديم الزمان عاش ملكان عظيمان شقيقان يُدعيان شهریار وشاهزمان. كان شهریار هو الأكبر سنّاً، وامتد سلطانه إلى أقاصي الأرض. شعر شهریار برغبة في رؤية أخيه، فطلب من كبير وزرائه الذهاب لإحضاره إليه.
Score: 0.2511222351580609
Semantic Search Result: ظل الملك شهریار يعمل طوال اليوم التالي، وعاد إلى المنزل في الليل إلى شهرزاد.
Distance: 1.1847589015960693
5. Classical Search Result: ودعاه التاجر البحار للعودة في اليوم التالي وتناول العشاء معه مرة أخرى.
Score: 0.24524147953659478
Semantic Search Result: ودعاه التاجر البحار للعودة في اليوم التالي وتناول العشاء معه مرة أخرى.
Distance: 1.2514185905456543

Query 10: ماذا طلبت السيدة التي ترتدي ملابس بالية من الرجل العجوز

1. Classical Search Result: وعند وصولنا إلى شاطئ البحر، قابلت سيدة ترتدي ملابس بالية قبلت يداي، وقالت: «سأسدي لك الخدمة دون مكافأة.»
«سيدة، لتسد إليّ خدمة، وأعتقد أنني سأتمكن من مكافأتك عليها.» أجبتها: «سأسدي لك الخدمة دون مكافأة.»
فقلت: «سيدة، لتزوجني وتأخذني معك على متن هذه السفينة.»
Score: 0.1953447496176008
Semantic Search Result: فقابل امرأة عجوز، ودفع لها مالاً مقابل أن تقايض ملابسها بملابسه. وذهب علاء الدين مرتدياً هذه الملابس إلى أحد المتاجر في المدينة، وابتاع سمّاً قوي المفعول، ثم عاد إلى القصر. وصل الساحر وهو يتنّى أن يقنع الأميرة بحبه لها.
Distance: 1.2328011989593506

2. Classical Search Result: وظلت هناك مدة أسبوع حتى نظر الملك في أمرها، وطلب من وزيره إحضارها إليه. أخبرت السيدة العجوز الملك بشأن حب ابنها لابنته، ورغبته في الزواج منها. ضحك الملك بصوت مرتفع مما سمعه، وسأل السيدة عما جلبته له.
Score: 0.19382395122296742
Semantic Search Result: فرد عليه التاجر التحية. سأله الرجل العجوز: «لماذا تجلس هنا، يا صاحبي، في هذا المكان الممتلئ بالجن والشياطين؟ فما من شيء في هذا البستان المسكون سوى الحزن والأسى»
Distance: 1.2874681949615479
3. Classical Search Result: رد الجني: «والله، إنها لقصة غريبة ومدهشة، كما وعدت. لقد فزت، وسأعق الرجل». ثم أطلق الجني سراح الرجل، ورحل. شكر التاجر الرجل العجوز الذي هنأه بدوره على استرداد حريته.
Score: 0.15547409034430923
Semantic Search Result: من يقايض مصباحاً قديماً بواحد من هذه المصابيح الجديدة؟
Distance: 1.2879889011383057
4. Classical Search Result: وفي أحد الأيام، عثرنا على بعض العنب، الأمر الذي أسعد الرجل العجوز للغاية. أخبرته أن بإمكانني تحويل العنب إلى شراب له، وابتهج لهذه الفكرة. عندما أصبح الخليط جاهزاً، شربه الرجل العجوز وكأنه ماء.
Score: 0.15226016364450282
Semantic Search Result: سمعت إحدى خادومات القصر نداء الساحر، وتذكرت أنها رأت مصباحاً قديماً كان علاء الدين يحتفظ به على أحد الأرفف في غرفة نومه. فذهبت لإحضاره، وقايضته مع الغريب، وهي تفكر في مدى حكمتها ومدى سذاجة الرجل العجوز. وابتهج الساحر عندما رأى المصباح بين يديه، وفر من المدينة حتى وصل إلى الصحراء المحيطة بها بعد أن تأكد أنه ما من أحد يتبعه.
Distance: 1.2996954917907715
5. Classical Search Result: أخبره التاجر بكل ما حدث مع الجني، فاندھش الرجل العجوز من القصة، وقال: «إن وعدك بالرجوع إلى هنا من أفعال الشجعان»
Score: 0.13579565563734247
Semantic Search Result: وظلت هناك مدة أسبوع حتى نظر الملك في أمرها، وطلب من وزيره إحضارها إليه. أخبرت السيدة العجوز الملك بشأن حب ابنها لابنته، ورغبته في الزواج منها. ضحك الملك بصوت مرتفع مما سمعه، وسأل السيدة عما جلبته له.
Distance: 1.3210644721984863

1.3.2 Analysis Of Retrieval System Results

The retrieval system results for the Arabic queries based on *One Thousand and One Nights* demonstrate a clear distinction between **classical (TF-IDF)** and **semantic (embedding-based)** search methods:

- **Classical Search (TF-IDF):** This method consistently retrieves relevant paragraphs with higher precision for queries directly referencing key terms (e.g., Query 2: من الأكبر سنًا بين شهباز وشاهزمان yields a top result with a score of 0.422 accurately identifying Shahryar as the elder). However, it struggles with queries requiring deeper contextual understanding (e.g., Query 6: ما الأمر المريع returns less relevant results like the general story of Shahryar and Shahrazad, score 0.333). This highlights TF-IDF's reliance on keyword matching, which can miss nuanced meanings.

- **Semantic Search (Embedding-Based):** Semantic search performs better for queries needing contextual inference (e.g., Query :5 لماذا قرر شهريار أنه لا يمكن الثقة في أي امرأة قط . However, it occasionally retrieves irrelevant results for simpler queries (e.g., Query :3 وراء أحزان شاهزمان returns a generic story, distance .1.298 This suggests embeddings capture broader semantic relationships but may overgeneralize for precise queries.
- **Comparison:** Classical search excels in keyword-driven queries, while semantic search is superior for abstract or context-heavy queries. The side-by-side display of results enhances user comparison, but the varying relevance (e.g., Query :8 من دینارزاد has classical search correctly identifying her as Shahrazad's sister, score ,0.389 while semantic search retrieves an unrelated story, distance 1.323 indicates room for hybrid approaches to balance precision and context.

1.3.3 RAG System Results

Query :1 من هم شهريار وشاهزمان؟

- Answer Without Context: شهريار وشاهزمان هما شخصيتان رئيسيتان في مجموعة من القصص المعروفة باسم "ألف ليلة وليلة". شهريار هو ملك يكتشف خيانة زوجته، مما يدفعه إلى اتخاذ قرار بالزواج من امرأة جديدة كل ليلة ثم إعدامها في الصباح التالي. هذا السلوك يستمر حتى يتزوج من شهرزاد. شاهزمان هو شقيق شهريار، ويظهر في بعض القصص كملك آخر. يشارك في الأحداث التي تتعلق بشهريار ويعكس بعض التوترات العائلية بينهما.
- Classical Search Answers:
 1. شهريار وشاهزمان هما ملكان عظيمان شقيقان. شهريار هو الأكبر سنًا ويمتد سلطانه إلى أقاصي الأرض، بينما score = 0.3299, شاهزمان هو الأخ الأصغر الذي يعاني من الحزن بسبب هجر زوجته الملكة له.
 2. شهريار وشاهزمان هما ملكان عظيمان شقيقان. شهريار هو الأكبر سنًا، ويمتد سلطانه إلى أقاصي الأرض، بينما score = 0.2774, شاهزمان هو الأخ الأصغر.
 3. شهريار وشاهزمان هما ملكان عظيمان شقيقان. شهريار هو الأكبر سنًا ويمتد سلطانه إلى أقاصي الأرض، بينما score = 0.1785, شاهزمان هو الأخ الأصغر الذي يعاني من حزن عميق بسبب هجر زوجته الملكة.
- Semantic Search Answers:
 1. شهريار وشاهزمان هما ملكان عظيمان شقيقان. شهريار هو الأكبر سنًا، ويمتد سلطانه إلى أقاصي الأرض، بينما distance = 1.1678, شاهزمان هو الأخ الأصغر.
 2. شهريار هو الملك الذي يروي له شهرزاد القصص في كل ليلة، بينما شاهزمان هو أخو شهريار. في القصة، يعاني distance = 1.4751, كلاهما من خيانة زوجتهما، مما يؤدي إلى تصرفات شهريار القاسية تجاه النساء.
 3. شهريار وشاهزمان هما شخصيتان من حكايات "ألف ليلة وليلة". شهريار هو الملك الذي قرر الزواج من النساء ثم قتلهن بعد ليلة واحدة، بينما شاهزمان هو شقيقه الذي يشاركه في هذه القصة. تتناول الحكايات الصراعات distance = 1.4866, والمغامرات التي يواجهها الأخوان، وتظهر كيف تؤثر قصص النساء على مصيرهم.

Query :2 من الأكبر سنًا بين شهريار وشاهزمان؟

- Answer Without Context: شهريار هو الأكبر سنًا بين شهريار وشاهزمان.

- Classical Search Answers:

1. الأكبر سنًا بين شهريار وشاهزمان هو شهريار. score = 0.4222
2. الأكبر سنًا بين شهريار وشاهزمان هو . score = 0.2139
3. الأكبر سنًا بين شهريار وشاهزمان هو شهريار. score = 0.1376

- Semantic Search Answers:

1. الأكبر سنًا بين شهريار وشاهزمان هو شهريار. distance = 1.1032
2. الأكبر سنًا بين شهريار وشاهزمان هو شهريار. distance = 1.1401
3. شهريار هو الأكبر سنًا بين شهريار وشاهزمان. distance = 1.1832

Query :3 ما السبب الرئيسي وراء أحزان شاهزمان؟

- Answer Without Context: السبب الرئيسي وراء أحزان شاهزمان هو فقدانه لزوجته، حيث يعاني من الألم والحزن العميق بسبب فقدان الحب الذي كان يجمعهما. هذا فقدان يؤثر على حياته بشكل كبير، مما يجعله يعيش في حالة من الكآبة والحزن.

- Classical Search Answers:

1. السبب الرئيسي وراء أحزان شاهزمان هو أنه هجرته زوجته الملكة، مما جعله يشعر بالكسرة والحزن الشديد. score = 0.2625
2. السبب الرئيسي وراء أحزان شاهزمان هو أن الفتيات الأربعين، اللواتي هن بنات لأربعين ملكًا، سيضطرون لمغادرة القصر لمدة أربعين يومًا، وعندما يغادرن، لن يتمكن من اصطحابه معهن. هذا الفراق يسبب لمن الحزن، حيث score = 0.0795، يشعرون بأنهن سيفقدن رفقته ودعمه، ويعتقدن أن مغادرتهم ستؤدي إلى عدم رؤيته مرة أخرى.
3. السبب الرئيسي وراء أحزان شاهزمان هو فقدانه لابنه علاء الدين، حيث كان يشعر بالقلق والقلق بسبب عدم معرفته بمكانه أو حالته، مما جعله يعيش في حالة من الحزن والاضطراب. score = 0.0610

- Semantic Search Answers:

1. السبب الرئيسي وراء أحزان شاهزمان هو شعوره بالوحدة والفراغ العاطفي، حيث كان يعيش بعيدًا عن أخيه شهريار، وكان يفتقده بشدة. كما أن الأحداث التي مر بها شهريار، والتي تتعلق بخيانة زوجته، أثرت أيضًا على distance = 1.2976، مشاعر شاهزمان وجعلته يشعر بالحزن والقلق.
2. السبب الرئيسي وراء أحزان شاهزمان هو فقدانه لأبنائه وبناته وزوجته، حيث كانوا يكون ويعبرون عن حزنهم بسبب مغادرته. كما أن مشيئته للعودة إلى البستان لملاقاة الجني تحمل في طياتها مشاعر القلق والخوف من المجهول، distance = 1.3538، مما يزيد من حزنه.
3. السبب الرئيسي وراء أحزان شاهزمان هو فقدان ابنه، حيث أنه مفقود ولا يعلم ما حدث له، مما يسبب له حزنًا شديدًا. distance = 1.3606

Query :4 لماذا غضب شهريار غضبًا شديدًا؟

- Answer Without Context: غضب شهريار غضباً شديداً بسبب خيانة زوجته له، حيث اكتشف أنها كانت تخونه مع رجل آخر. هذا الأمر أثر عليه بشكل عميق وجعله يفقد الثقة في النساء بشكل عام، مما دفعه إلى اتخاذ قرار بالزواج من امرأة جديدة كل ليلة ثم إعدامها في الصباح التالي، كنوع من الانتقام من النساء.

- Classical Search Answers:

1. غضب شهريار غضباً شديداً لأنه اكتشف أن أخاه شاهزمان قد هجرته زوجته الملكة، مما جعله يشعر بعدم الثقة في النساء بشكل عام. هذا الأمر أثار مشاعر الغضب والخيبة لديه، وقرر أنه لا يمكن الوثوق بأي امرأة، مما دفعه إلى score = 0.3035, التخطيط لقتل جميع النساء في مملكته.
2. غضب شهريار غضباً شديداً بسبب الحزن الذي شعر به عندما سمع عن معاناة عمه وفقدان ابنه. هذا الحزن تحول إلى غضب، مما يدل على عمق مشاعره وتأثره بالموقف. قد يكون أيضاً مرتبطاً بشعوره بالعجز عن مساعدته أو score = 0.2512, بسبب تذكيره بخيانتته الشخصية ومعاناته الخاصة.
3. غضب شهريار غضباً شديداً بسبب مشاعر الإحباط والاستياء التي كانت تتراكم في داخله، مما دفعه إلى التفكير في استخدام القوة لتدمير الجني والنقش المحفور. يبدو أن هناك موقفاً أو حدثاً معيناً أثار هذا الغضب، لكنه لم يكن score = 0.2183, قادراً على الاستماع إلى تحذيرات شهرزاد أو نصائحها بسبب شدة انفعاله.

- Semantic Search Answers:

1. غضب شهريار غضباً شديداً بسبب الحزن الذي شعر به عندما سمع عن معاناة عمه وفقدان ابنه. هذا الحزن تحول إلى غضب، مما يدل على عمق مشاعره وتأثره بالموقف. قد يكون أيضاً مرتبطاً بشعوره بالعجز عن مساعدته أو distance = 1.0196, بسبب تذكيره بخيانتته الشخصية، مما زاد من مشاعر الغضب والقلق لديه.
2. غضب شهريار غضباً شديداً لأنه اكتشف خيانة زوجته له، مما أدى إلى فقدان ثقته في النساء بشكل عام. هذا الغضب دفعه إلى اتخاذ قرار قاسي بالزواج من امرأة جديدة كل ليلة ثم إعدامها في الصباح التالي، ليضمن عدم distance = 1.2159, تعرضه للخيانة مرة أخرى.
3. غضب شهريار غضباً شديداً لأنه اكتشف خيانة زوجته له، مما أدى إلى فقدان ثقته في النساء بشكل عام. هذا الغضب دفعه إلى اتخاذ قرار بقتل كل امرأة يتزوجها بعد ذلك، حيث قرر أن يتزوج كل ليلة امرأة جديدة ثم distance = 1.2283, يقتلها في الصباح التالي، ليضمن عدم تعرضه للخيانة مرة أخرى.

Query 5: لماذا قرر شهريار أنه لا يمكن الثقة في أي امرأة قط؟

- Answer Without Context: قرر شهريار أنه لا يمكن الثقة في أي امرأة بسبب خيانتته من قبل زوجته، التي كانت تخونه مع رجل آخر. هذه الخيانة أثرت عليه بشكل عميق، مما دفعه إلى اتخاذ قرار قاسي بأن يتزوج كل ليلة امرأة جديدة ويقوم بقتلها في الصباح التالي، حتى لا يتيح لنفسه الفرصة للثقة في أي امرأة مرة أخرى. هذا القرار يعكس شعوره بالخذلان والمرارة تجاه النساء بشكل عام.

- Classical Search Answers:

1. قرر شهريار أنه لا يمكن الثقة في أي امرأة قط بسبب خيانة زوجته الملكة له، مما أدى إلى شعوره بالخذلان والألم. بعد أن علم بأن أخاه شاهزمان قد هجرته زوجته، زاد ذلك من غضبه وشكوكه تجاه النساء بشكل عام، مما دفعه score = 0.3730, إلى اتخاذ قرار مروع بالانتقام من جميع النساء في مملكته.

- 2. قرر شهريار أنه لا يمكن الثقة في أي امرأة قط بسبب خيانة زوجة أخيه شاهزمان له، مما جعله يشعر بخيبة أمل عميقة تجاه النساء. هذه الخيانة أثرت عليه بشكل كبير، وجعلته يعتقد أن جميع النساء غير مخلصات، مما دفعه إلى $score = 0.1869$, اتخاذ قرار قاسي بعدم الثقة في أي امرأة.
- 3. قرر شهريار أنه لا يمكن الثقة في أي امرأة قط بعد أن اكتشف خيانة زوجته له. فقد كانت لديه تجربة مؤلمة عندما اكتشف أن زوجته كانت تخونه مع رجل آخر، مما أدى إلى شعوره بالخيانة والخذلان. نتيجة لذلك، قرر أن يتخذ قراراً قاسياً بعدم الثقة في النساء، وبدأ في الزواج من امرأة جديدة كل ليلة ثم يقتلها في الصباح التالي، $score = 0.1792$, ليضمن عدم تعرضه للخيانة مرة أخرى.
- Semantic Search Answers:
 - 1. قرر شهريار أنه لا يمكن الثقة في أي امرأة قط بسبب خيانة زوجة أخيه شاهزمان له، حيث هجرته وتركته كسير الفؤاد. هذا الأمر أثار غضبه الشديد وجعله يشعر بالخيانة من النساء بشكل عام، مما دفعه إلى اتخاذ قرار مروع $distance = 0.8278$, بالانتقام من جميع النساء في مملكته.
 - 2. قرر شهريار أنه لا يمكن الثقة في أي امرأة قط بسبب خيائته من قبل زوجته السابقة، التي خانته مع رجل آخر. هذه الخيانة أثرت عليه بشكل عميق وجعلته يشعر بالخذلان، مما دفعه إلى اتخاذ قرار قاسي بأن يتزوج كل ليلة امرأة جديدة ويقوم بقتلها في الصباح التالي، حتى لا يتعرض للخيانة مرة أخرى. $distance = 1.2454$
 - 3. قرر شهريار أنه لا يمكن الثقة في أي امرأة قط بسبب خيائتها له، حيث اكتشف أن زوجته كانت تخونه مع رجل آخر. هذه الخيانة أثرت عليه بشكل عميق وجعلته يشعر بالخذلان، مما دفعه إلى اتخاذ قرار قاسي بعدم الثقة في النساء بشكل عام، وبدأ في الزواج من نساء ثم قتلهن بعد ليلة واحدة، اعتقاداً منه أن هذا هو السبيل لحماية نفسه $distance = 1.2796$, من الألم والخيانة.

ما الأمر المريع الذي خطط له شهريار؟ 6: Query

- Answer Without Context: الأمر المريع الذي خطط له شهريار هو أن يتزوج كل ليلة امرأة جديدة ثم يقتلها في الصباح التالي، وذلك بعد أن اكتشف خيائتها له. هذا التصرف كان نتيجة لخيانة أمه في النساء، وقد استمر في هذا الفعل حتى قابل شهرزاد، التي استخدمت ذكاءها وحيلتها لإنقاذ نفسها وإيقاف هذا السلوك.
- Classical Search Answers:
 - 1. الأمر المريع الذي خطط له شهريار هو أنه بعد أن اكتشف خيانة زوجته، قرر أن يتزوج كل ليلة من امرأة جديدة ثم يقتلها في الصباح التالي، وذلك كنوع من الانتقام من النساء. هذا القرار أدى إلى انتشار الخوف والرعب في مملكته. $score = 0.3334$
 - 2. الأمر المريع الذي خطط له شهريار هو الانتقام من النساء بسبب خيانة زوجته له. بعد أن اكتشف خيانة زوجته، قرر أن يتزوج كل ليلة من امرأة جديدة ثم يقتلها في الصباح التالي، ليضمن ألا تتكرر خيانة النساء له. هذا القرار $score = 0.2570$, أدى إلى سلسلة من الأحداث المأساوية التي أثرت على المملكة بأكملها.
 - 3. الأمر المريع الذي خطط له شهريار هو أنه كان ينوي قتل شهرزاد بعد انتهاء ليلة زفافهما، كما كان يفعل مع كل زوجة له، حيث كان يعتقد أن ذلك هو السبيل للتخلص من خيانة النساء. ولكن شهرزاد كانت تخطط لإنقاذ نفسها من هذا المصير من خلال سرد القصص المثيرة التي تجعل الملك يرغب في الاستماع إليها وعدم قتلها. $score = 0.2417$
- Semantic Search Answers:

1. الأمر المريع الذي خطط له شهريار هو أنه بعد أن اكتشف خيانة زوجته له، قرر أن يتزوج كل ليلة من امرأة جديدة ثم يقتلها في الصباح التالي، وذلك كنوع من الانتقام من النساء. هذا القرار أدى إلى انتشار الخوف والرعب له. distance = 1.2748
2. الأمر المريع الذي خطط له شهريار هو أن يقوم بقتل كل امرأة يتزوجها بعد ليلة واحدة، وذلك بسبب خيانتها له. فقد قرر أن ينتقم من النساء بشكل عام بعد أن اكتشف خيانة زوجته، فبدأ في الزواج من نساء ثم يقتلن في الصباح التالي. distance = 1.3963
3. الأمر المريع الذي خطط له شهريار هو أنه كان يعتقد أن لديه فقط سنة واحدة متبقية من عمره، وذلك بسبب ما حدث له مع الجني. هذا الأمر جعله يشعر بالحزن الشديد، وأثر على عائلته أيضاً، حيث تجمعوا حوله وهم سيكون distance = 1.4118

لماذا قالت شهرزاد أود أن تزوجني الملك شهريار؟ 7: Query

- Answer Without Context: شهرزاد لم تقل مباشرة أنها ترغب في الزواج من الملك شهريار، بل كانت تخطط لاستخدام ذكائها وحيلتها لإنقاذ نفسها من مصير الموت. كانت تعرف أن الملك شهريار يقتل زوجاته بعد ليلة واحدة، لذا قررت أن تروي له قصصاً مشوقة كل ليلة، مما يجعله يرغب في الإبقاء عليها حية لسماع المزيد من القصص. من خلال هذه الحيلة، تمكنت من كسب وقته وثقته، وفي النهاية، تغيرت مشاعر الملك نحوها.
- Classical Search Answers:
 1. قالت شهرزاد "أود أن تزوجني الملك شهريار" لأنها أرادت أن تنقذ شعبها من مصير مروع، حيث كان الملك يخطط لقتل جميع النساء في مملكته. كانت تأمل أن تتمكن من إيقاف هذا المخطط البغيض، أو على الأقل أن تموت مثل باقي الفتيات إذا لم تتمكن من ذلك. score = 0.3377
 2. قالت شهرزاد إنها ترغب في الزواج من الملك شهريار لأنها كانت تعرف أنه ملك قوي وعظيم، ولديها خطة لإنقاذ نفسها والنساء من مصيرهن المحتوم. كانت تدرك أن الملك كان يتزوج كل ليلة بامرأة جديدة ثم يقتلها في الصباح، لذا أرادت أن تكون هي من يروي له القصص ويشغله عن هذا السلوك، مما يمنحها الفرصة للبقاء على قيد الحياة. score = 0.2831
 3. شهرزاد لم تقل "أود أن تزوجني الملك شهريار"، بل كانت تبكي بسبب رغبتها في وداع أختها. كان بكائها يعبر عن مشاعر الحزن والفراق، وليس عن رغبتها في الزواج. هي استخدمت هذه الحيلة لتأجيل مصيرها، حيث أرادت أن تروي قصة لأختها قبل أن تواجه مصيرها المحتوم. score = 0.2368
- Semantic Search Answers:
 1. قالت شهرزاد "أود أن تزوجني الملك شهريار" لأنها كانت ترغب في إنقاذ شعبها من مصيرهم المأساوي، حيث كان الملك يقتل الفتيات بعد الزواج منهن. كانت تأمل أن تتمكن من تغيير هذا الوضع أو أن تموت مثل باقي الفتيات، مما يعكس شجاعتها وإرادتها القوية في مواجهة الظلم. distance = 0.9920
 2. شهرزاد لم تقل "أود أن تزوجني الملك شهريار"، بل كانت تبكي بسبب رغبتها في وداع أختها. كان بكائها يعبر عن مشاعر الحزن والفراق، وليس عن رغبتها في الزواج. هي استخدمت هذه الحيلة لتأجيل مصيرها، حيث كانت تخطط لسرد القصص للملك شهريار في الليالي القادمة، مما يمنحها فرصة للبقاء على قيد الحياة. distance = 0.9946
 3. قالت شهرزاد أود أن تزوجني الملك شهريار لأنها كانت تأمل في أن تتمكن من إنقاذ نفسها من مصير النساء اللواتي سبقن، حيث كان الملك شهريار يتزوج كل ليلة امرأة جديدة ثم يقتلها في الصباح. كانت شهرزاد ذكية وملمة

بالقصص، واعتقدت أن بإمكانها استخدام حكاياتها لإبقاء الملك مهتماً بها، مما يمنحها فرصة للبقاء على قيد الحياة.
distance = 0.9974

Query :8 من دینارزاد؟

- Answer Without Context: دینارزاد هي شخصية خيالية من حكايات "ألف ليلة وليلة". هي الأخت الصغرى لشهريار، الملك الذي يروي القصص. تُعرف دینارزاد بأنها الذكية والمبدعة، حيث كانت تروي القصص لشقيقها الملك كل ليلة، مما ساعد على إنقاذ حياتها وحياة النساء الأخريات من مصيرهن المحتوم. قصصها كانت مليئة بالحكمة والمغامرات، مما جعلها شخصية محورية في السرد القصصي.
- Classical Search Answers:
 1. دینارزاد هي أخت شهرياد، الشخصية الرئيسية في القصة. score = 0.3885
 2. دینارزاد هي الأخت الصغرى لشهرياد، وهي الشخصية التي تروي القصص في "ألف ليلة وليلة". تقوم دینارزاد بدور المحفز لقصص شهرياد، حيث تطلب منها أن تروي لها قصة كل ليلة، مما يساعد على تأجيل مصير شهرياد. score = 0.2270
 3. دینارزاد هي إحدى الشقيقتين في القصة، وهي شخصية محورية تُعرف بكونها تحب سرد القصص. في هذا السياق، تظهر كأخت لشهرياد، التي تُعتبر الراوية الرئيسية في الحكايات. دینارزاد تلعب دوراً مهماً في تحفيز شهرياد على سرد القصص من خلال تعبيرها عن إعجابها بالقصص التي تُروى. score = 0.2172
- Semantic Search Answers:
 1. دینارزاد هي شخصية رئيسية في قصة "علاء الدين" وهي الأخت الصغرى لشهريار. تُعرف بأنها الذكية والحكيمة، وتلعب دوراً مهماً في سرد القصص، حيث تساعد في تخفيف حزن الملك شهريار من خلال حكاياتها المثيرة. distance = 1.3227
 2. دینارزاد هي أخت شهرياد، الشخصية الرئيسية في القصة. distance = 1.3508
 3. دینارزاد هي شخصية خيالية من حكايات "ألف ليلة وليلة"، وهي الأخت الصغرى لشهريار. تُعرف بأنها الراوية التي تروي القصص لشهريار في كل ليلة، مما يمنع الملك من قتلها ويجعله يستمع إلى حكاياتها الشيقة. distance = 1.3801

Query :9 لماذا قرر شهريار ان يقتلها في اليوم التالي؟

- Answer Without Context: قرر شهريار أن يقتل شهرياد في اليوم التالي لأنه كان قد فقد الثقة في النساء بعد خيانتها له، واعتقد أن قتلها هو الحل للتخلص من الألم الذي يشعر به. كان يعتقد أن كل النساء غير مخلصات، ولذلك قرر أن ينهي حياتها كنوع من الانتقام. لكن شهرياد استخدمت ذكائها وحيلتها لتأجيل مصيرها من خلال سرد القصص المثيرة التي جعلته يتشوق لسماع المزيد، مما أنقذ حياتها في النهاية.
- Classical Search Answers:
 1. قرر شهريار أن يقتل شهرياد في اليوم التالي لأنه كان يشعر بالحزن واليأس بعد أن علم أن عمره المتبقي هو سنة واحدة فقط، وكان يعتقد أن مصيره المحتوم هو الموت. في ظل هذا الشعور، قد يكون قد رأى في شهرياد رمزاً للتهديد أو الخطر، مما دفعه إلى اتخاذ قرار القتل كوسيلة للتخلص من هذا الخوف. score = 0.4009
 2. قرر شهريار أن يقتلها في اليوم التالي بسبب خيانتها له، حيث اكتشف أنها كانت تخونه مع رجل آخر. هذا الأمر أثار غضبه ورغبته في الانتقام، مما دفعه إلى اتخاذ قرار بقتلها. score = 0.3002

- 3. قرر شهريار أن يقتل شهرزاد في اليوم التالي لأنه كان قد فقد الثقة في النساء بعد أن خانته زوجته الأولى. وقد قرر أن يتزوج كل ليلة امرأة جديدة ثم يقتلها في الصباح التالي حتى لا تتاح لها الفرصة لخيانته كما حدث في الماضي. ولكن شهرزاد، بذكائها وحيلتها، استطاعت أن تؤجل مصيرها من خلال سرد القصص المثيرة التي تأسر انتباهه. score = 0.2961

- Semantic Search Answers:

- 1. قرر شهريار أن يقتل زوجته في اليوم التالي لأنه فقد الثقة في النساء بعد أن هجرته زوجته الملكة، واعتقد أن جميع النساء غير مخلصات. كان يخطط للزواج من امرأة واحدة فقط ليوم واحد ثم قتلها، كنوع من الانتقام من النساء distance = 1.0419, بشكل عام، وذلك بسبب الألم الذي شعر به نتيجة خيانة زوجته.
- 2. قرر شهريار أن يقتلها في اليوم التالي بسبب خيانتها له، حيث اكتشف أنها كانت تخونه مع رجل آخر. هذا الأمر distance = 1.0770, أثار غضبه ورغبته في الانتقام، مما دفعه إلى اتخاذ قرار بقتلها.
- 3. قرر شهريار أن يقتلها في اليوم التالي لأنه كان يشعر بالخيانة والقلق من نواياها. بعد أن استمع إلى قصتها، أدرك أنها كانت تخطط لقتل شقيقه، مما جعله يشعر بأن حياتها كانت تهديداً له ولعائلته. هذا الشعور بالخوف من الخيانة distance = 1.1221, والانتقام دفعه إلى اتخاذ قرار بقتلها قبل أن تتمكن من تنفيذ خطتها.

Query :10 ماذا طلبت السيدة التي ترتدي ملابس بالية من الرجل العجوز؟

- Answer Without Context: طلبت السيدة التي ترتدي ملابس بالية من الرجل العجوز أن يساعدها أو يمنحها شيئاً من المال أو الطعام، حيث كانت في حالة من الحاجة والفقر.
- Classical Search Answers:
 - 1. score = 0.1953, طلبت السيدة التي ترتدي ملابس بالية من الرجل العجوز أن يتزوجها ويأخذها معه على متن السفينة.
 - 2. طلبت السيدة التي ترتدي ملابس بالية من الرجل العجوز أن يساعدها في الحصول على بعض الطعام أو المال، score = 0.1938, حيث كانت تعاني من الفقر والحاجة.
 - 3. طلبت السيدة التي ترتدي ملابس بالية من الرجل العجوز أن يساعدها أو يمنحها شيئاً من المال أو الطعام، حيث score = 0.1555, كانت في حالة من الحاجة والفقر.
- Semantic Search Answers:

- 1. السيدة التي ترتدي ملابس بالية لم تطلب شيئاً من الرجل العجوز، بل كانت قد طلبت من علاء الدين أن يعطيها distance = 1.1504, شيئاً من المال، ولكنها لم تتلقَ أي رد.
- 2. طلبت السيدة التي ترتدي ملابس بالية من الرجل العجوز أن يساعدها، حيث كانت في حاجة ماسة إلى المساعدة distance = 1.1641, بسبب ظروفها الصعبة.
- 3. طلبت السيدة التي ترتدي ملابس بالية من الرجل العجوز أن يساعدها في الحصول على بعض الطعام أو المساعدة، distance = 1.2006, حيث كانت في حالة من الحاجة والفقر.

1.3.4 Analysis Of Retrieval-Augmented Generation (RAG) Results

The Retrieval-Augmented Generation (RAG) system results showcase the impact of combining retrieved context with the gpt-4o-mini LLM for answering Arabic questions:

- **Baseline (LLM-only):** Without context, the LLM provides general but sometimes inaccurate answers (e.g., Query :8 من دينارزاد incorrectly identifies her as Shahryar's sister and the storyteller). This reflects the LLM's limited knowledge of specific details in *One Thousand and One Nights*, underscoring the need for context.
- **Classical RAG:** Answers using TF-IDF-retrieved context are more accurate for queries with explicit keywords (e.g., Query :1 من هم شهریار وشاهزمان correctly identifies them as brother kings, score .(0.3299 However, it misinterprets nuanced queries (e.g., Query :3 ما السبب الرئيسي وراء أحزان شاهزمان wrongly attributes his sadness to unrelated events, score ,(0.0795 reflecting TF-IDF's keyword bias.
- **Semantic RAG:** Semantic context improves answers for complex queries (e.g., Query :5 لماذا قرر شهریار أنه لا يمكن الثقة في أي امرأة قط accurately links his decision to his wife's betrayal, distance .(0.8278 Yet, it falters in simpler cases (e.g., Query :8 من دينارزاد misidentifies her as a character in *Aladdin*, distance ,(1.3227 indicating overgeneralization by embeddings.
- **Overall:** RAG significantly enhances answer quality over the baseline, with classical RAG excelling in straightforward queries and semantic RAG better handling nuanced ones. Errors in both methods (e.g., Query 9's misattributions) suggest that improving retrieval relevance and context selection (e.g., including adjacent paragraphs) could further boost performance. The system's ability to display multiple answer types aids evaluation but highlights the need for refined retrieval to minimize inaccuracies.

2 Problem 2: Benchmarking Arabic ASR

2.1 Checklist Table

#	Category	Checklist Item	Status (Yes/No)	Notes
1	Scope Definition	Have you clearly defined the NLT task(s) to be evaluated?	Yes	I have evaluated Automatic Speech Recognition (ASR)
2	Scope Definition	Have you specified the evaluation goals (accuracy, robustness, etc.)?	Yes	Character Error Rate (CER) is the most suitable metric for Arabic ASR
3	Language and Domain Selection	Have you selected relevant languages and dialects?	Yes	I am Using Modern Standard Arabic (MSA), Classical Arabic, and Egyptian dialect, each contributing approximately 0.33 to the dataset
4	Language and Domain Selection	Have you included multiple genres and domains?	Yes	
5	Data Collection	Is the dataset licensed for research/benchmark use?	Yes	All used datasets are open source
6	Data Collection	Have you included real-world or synthetic data?	Yes	All used datasets are from broadcast audios or TV which are real-world data
7	Annotation Quality	Is there a clear annotation schema for the task?	Yes	
8	Annotation Quality	Is inter-annotator agreement calculated and acceptable?	Yes	
9	Data Balance	Does the dataset reflect demographic and topical diversity?	Yes	Modern Standard Arabic (MSA), Classical Arabic, and Egyptian dialect, each contributing approximately 0.33 to the dataset
10	Preprocessing	Is the data cleaned and normalized consistently?	Yes	
11	Preprocessing	Is the preprocessing procedure documented?	Yes	
12	Evaluation Metrics	Are appropriate metrics defined for each NLT task?	Yes	CER is suitable for Arabic ASR
13	Evaluation Metrics	Do you include baseline results or evaluation scripts?	Yes	
14	Documentation	Is the data creation process fully documented?	Yes	
15	Documentation	Are known limitations or biases discussed?	Yes	

Table 1: Checklist for NLT Task Evaluation

2.2 Methodology

2.2.1 Evaluation

We have Choose **Character Error Rate (CER)** to be our Evaluation metric, that is because it is a Suitable Metric for Arabic ASR as it is less strict than **word Error Rate (WER)** and gives more reasonable meaning of the quality of the transcript.

2.2.2 Dataset Preparation

The dataset for benchmarking Arabic Automatic Speech Recognition (ASR) was curated to include three linguistic variants: Modern Standard Arabic (MSA), Classical Arabic, and Egyptian dialect, each contributing approximately 33% to the dataset. The dataset comprises audio clips with corresponding text transcriptions, sourced from publicly available datasets on the Hugging Face platform. The preparation process ensured a balanced representation of dialects, a minimum duration of 30 minutes per variant, and a standardized audio format. The following steps outline the methodology, implemented using Python with the datasets library.

.1 **Loading Raw Datasets:** Three datasets were selected to represent the target linguistic variants:

- **Egyptian Arabic:** The MightyStudent/Egyptian-ASR-MGB-3 dataset (train split) provided conversational and broadcast audio in the Egyptian dialect.
- **Classical Arabic:** The MBZUAI/ClArTTS dataset (train split) included recitations and formal speech in Classical Arabic.
- **MSA:** The halabi2016/arabic_speech_corpus dataset (train split) contained formal speech in Modern Standard Arabic.

Each dataset was loaded using the `load_dataset` function from the datasets library.

.2 **Column Standardization:** To ensure consistency, only the audio and text columns were retained. For the Egyptian dataset, the sentence column was renamed to text. The Classical Arabic dataset required reformatting to align its audio structure with the others, using a mapping function to wrap audio data into the standard format `{"array", "sampling_rate"}`. Unnecessary columns were removed using the `remove_columns` method.

.3 **Subsampling Classical Arabic:** The Classical Arabic dataset)ClArTTS(was significantly larger than needed. To balance the dataset and reduce processing time, it was shuffled with a fixed seed (0) and subsampled to 1000 examples using the `select` method.

.4 **Audio Standardization:** All audio clips were standardized to a 16 kHz sampling rate using the `cast_column` method with the Audio feature from the datasets library. This ensured compatibility across datasets and alignment with standard ASR model requirements.

.5 **Selecting Balanced Subsets:** To achieve a minimum duration of 30 minutes (1800) sec) per variant, a custom function was implemented to:

- (a) Shuffle each dataset with a fixed seed (1) for reproducibility.
- (b) Iterate through examples, calculating the duration of each audio clip `len(array) / sampling_rate`.
- (c) Selected examples until the cumulative duration reached or exceeded 1800 seconds.
- (d) Stored the duration as a new column `duration` for later analysis.

The selected examples were converted back into Dataset objects using `Dataset.from_list`.

- .6 **Text Normalization for MSA:** The MSA dataset's text was encoded in Buckwalter transliteration. A `buckwalter_to_arabic` function was applied to convert these transcriptions to standard Arabic script, ensuring consistency with the other datasets.
- .7 **Combining Datasets:** The balanced subsets for Egyptian Arabic, Classical Arabic, and MSA were concatenated into a single dataset using `concatenate_datasets`. This combined dataset served as the final benchmarking corpus.

The resulting dataset is balanced across the three linguistic variants, with each contributing at least 30 minutes of audio. The use of real-world datasets ensures diversity in speakers, background noise, and recording conditions, enhancing the robustness of the benchmarking process. However, limitations include:

- Potential biases in the source datasets (e.g., over-representation of formal speech in MSA).
- Limited dialectal diversity (only Egyptian dialect included among dialects).
- Dependence on the quality of original transcriptions, which may contain errors.

2.2.3 Used Models

- .1 **wav2vec2-large-xlsr-53-arabic:** The facebook/wav2vec2-large-xlsr-53-arabic model, is a fine-tuned variant of the wav2vec2-large-xlsr-53 model, specifically optimized for Arabic automatic speech recognition (ASR). The base model was pre-trained on 56,000 hours of multilingual audio data across 53 languages using a self-supervised learning approach with the wav2vec 2.0 framework, which combines contrastive learning and masked language modeling to extract robust speech representations. The Arabic fine-tuning was performed on datasets such as Common Voice Arabic and MGB-2, enhancing its ability to transcribe Arabic speech accurately. The model features 24 transformer layers, a model dimension of 1,024 and 317 million parameters, supporting 16kHz sampled audio. It is licensed under Apache 2.0 and intended for research purposes.
- .2 **wav2vec2-large-xlsr-53-arabic-egyptian:** The facebook/wav2vec2-large-xlsr-53-arabic-egyptian model, is another fine-tuned variant of the wav2vec2-large-xlsr-53 model, tailored specifically for Egyptian Arabic, a dialect of Arabic. It inherits the pre-training from the base model. The fine-tuning process focused on Egyptian Arabic datasets, such as those from the Arabic Speech Corpus, to improve recognition of dialectal features. The architecture remains consistent with the base model, featuring 24 transformer layers, a 1024-dimensional model, and 317 million parameters, supporting 16kHz audio. It is also licensed under Apache 2.0 and aimed at research applications.
- .3 **whisper-large-v3:** The openai/whisper-large-v3 model, is an advanced version of OpenAI's Whisper family, designed for robust multilingual speech recognition. Pre-trained on 680,000 hours of audio data across 98 languages, including Arabic, Whisper-large-v3 uses an encoder-decoder transformer architecture with 1550 million parameters. It was trained on a diverse dataset covering various accents and noise conditions, enabling it to handle a wide range of speech inputs. The model supports tasks beyond ASR, such as translation and language identification, and is optimized for 16kHz audio. It is licensed under the MIT License and intended for research and commercial use, with improvements over previous versions in transcription accuracy and robustness.

- .4 **Whisper Small-Ar:** The Whisper Small-Ar model, This model is a fine-tuned version of openai/whisper-small for Arabic ASR . It achieves more efficient while retaining good performance for Arabic transcription. It supports 16kHz audio and is typically licensed under MIT, aimed at research and lightweight ASR applications.
- .5 **HuBERT Egyptian CTC:** The HuBERT Egyptian CTC model, Is a fine-tuned version of the HuBERT (Hidden-Unit BERT) model for Egyptian Arabic on the MGB-3 and Egyptian Arabic Conversational Speech Corpus datasets, achieving a state of the art for Egyptian Arabic with WER.
- .6 **facebook/seamless-m4t-v2-large:** The facebook/seamless-m4t-v2-large model, is part of the SeamlessM4T family, a multimodal and multilingual model for speech-to-text, text-to-speech, and translation across nearly 100 languages, including Arabic. It features 2.3 billion parameters and was pre-trained on 443,000 hours of speech and 29 billion text sentences, utilizing web-scale data like SeamlessAlign for aligned speech-text pairs. The model employs a dual-encoder architecture (one for speech, one for text) with a shared transformer decoder, enabling seamless modality transitions. It supports 16kHz audio and is licensed under CC BY-NC ,4.0 intended for research purposes.

2.3 Results

2.3.1 Results Table

The Table 2 presents the Character Error Rate (CER) results for the previously mentioned Automatic Speech Recognition (ASR) models evaluated on The defined dataset.

Table 2: ASR Model Performance (CER)

	Model	CER
1	wav2vec2-large-xlsr-53-arabic	0.232978
2	wav2vec2-large-xlsr-53-arabic-egyptian	0.430803
3	whisper-large-v3	0.396615
4	Whisper Small-Ar	0.384210
5	HuBERT Egyptian CTC	0.395884
6	facebook/seamless-m4t-v2-large	0.41177

2.3.2 Results Analysis

Individual Model Performance:

- .1 **wav2vec2-large-xlsr-53-arabic CER: 0.232978:** It demonstrates strong performance for Arabic ASR, likely due to its fine-tuning on Arabic datasets like Common Voice Arabic and MGB-2, following pre-training on 56,000 hours of multilingual audio. Its architecture (24) transformer layers, 317 million parameters) and the wav2vec 2.0 framework’s focus on contextualized speech representations enable it to effectively capture Arabic phonetic nuances, aligning with its design for robust multilingual ASR.
- .2 **wav2vec2-large-xlsr-53-arabic-egyptian CER: 0.430803:** This model targets Egyptian Arabic but shows a higher error rate. Despite pre-training on 56,000 hours of audio, its fine-tuning on Egyptian Arabic datasets like the Arabic Speech Corpus may be limited by data scarcity or the dialect’s divergence from Modern Standard Arabic (MSA), This suggests challenges in dialectal ASR without extensive dialect-specific data.

- .3 **whisper-large-v3 CER: 0.396615**: The moderate performance may reflect its broad multilingual focus, potentially diluting Arabic-specific accuracy, as it also supports translation and language identification, introducing trade-offs in pure ASR tasks.
- .4 **Whisper Small-Ar CER: 0.384210**: It suggests that Arabic-specific fine-tuning enhances efficiency, as noted in Whisper’s documentation, highlighting the value of targeted optimization over model size for specific languages.
- .5 **HuBERT Egyptian CTC CER: 0.395884**: Indicates challenges with Egyptian Arabic, possibly due to limited fine-tuning data or dialectal complexity, despite HuBERT’s strength in acoustic representation learning via self-supervised clustering, as per its documentation.
- .6 **facebook/seamless-m4t-v2-large CER: 0.41177**: This 2.3 billion-parameter model was pre-trained on 443,000 hours of speech across 100 languages. Its higher CER may stem from its multimodal design (supporting speech-to-text, text-to-speech, and translation), which, as noted in the documentation, lacks Arabic-specific fine-tuning, impacting its ability to handle Arabic speech nuances as effectively as specialized models.

Collective Analysis:

The models exhibit a CER range from 0.232978 to 0.430803 reflecting varying success in Arabic ASR. The

wav2vec2-large-xlsr-53-arabic model achieves the lowest CER, benefiting from targeted Arabic fine-tuning, as highlighted in its documentation. Dialect-focused models (wav2vec2-large-xlsr-53-ar-HuBERT Egyptian CTC) show higher CERs, suggesting challenges with Egyptian Arabic due to limited dialect-specific data or phonetic divergence from MSA. General-purpose models like whisper-large-v3 and facebook/seamless-m4t-v2-large perform moderately, likely due to their broader multilingual focus, which may reduce Arabic-specific accuracy. Notably, Whisper Small-Ar outperforms its larger counterpart, indicating that fine-tuning can be more effective than model size, a trend supported by Whisper’s documentation. Overall, the results underscore the importance of language-specific fine-tuning for Arabic ASR, particularly for dialects, and suggest that multimodal or generalized models may face trade-offs in performance for targeted tasks like Arabic speech transcription.