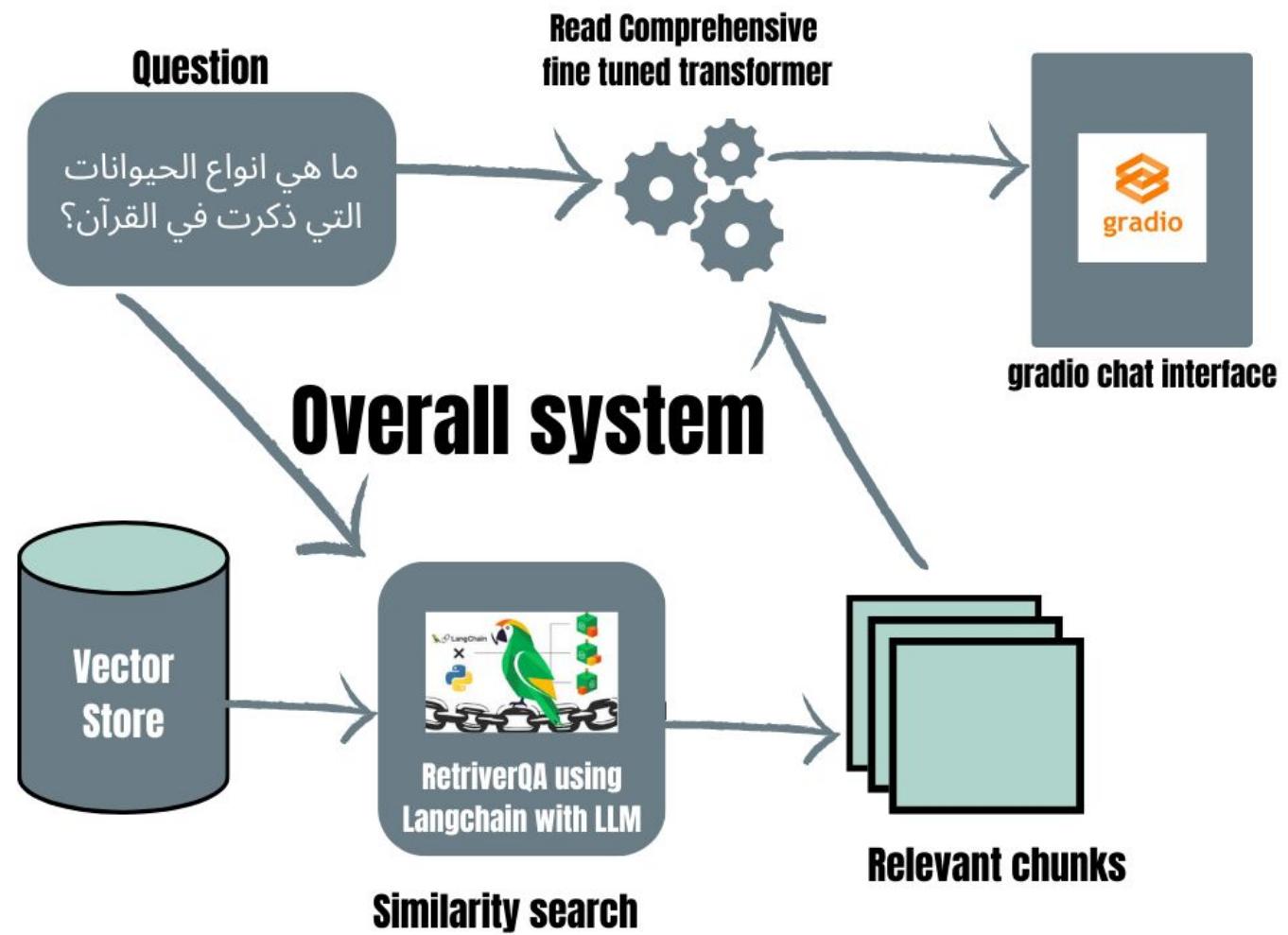


# **Quran Question Answering**



By

Asmaa Ahmed  
Salma Mohamed  
Mariam Gamal



# Overall system

---



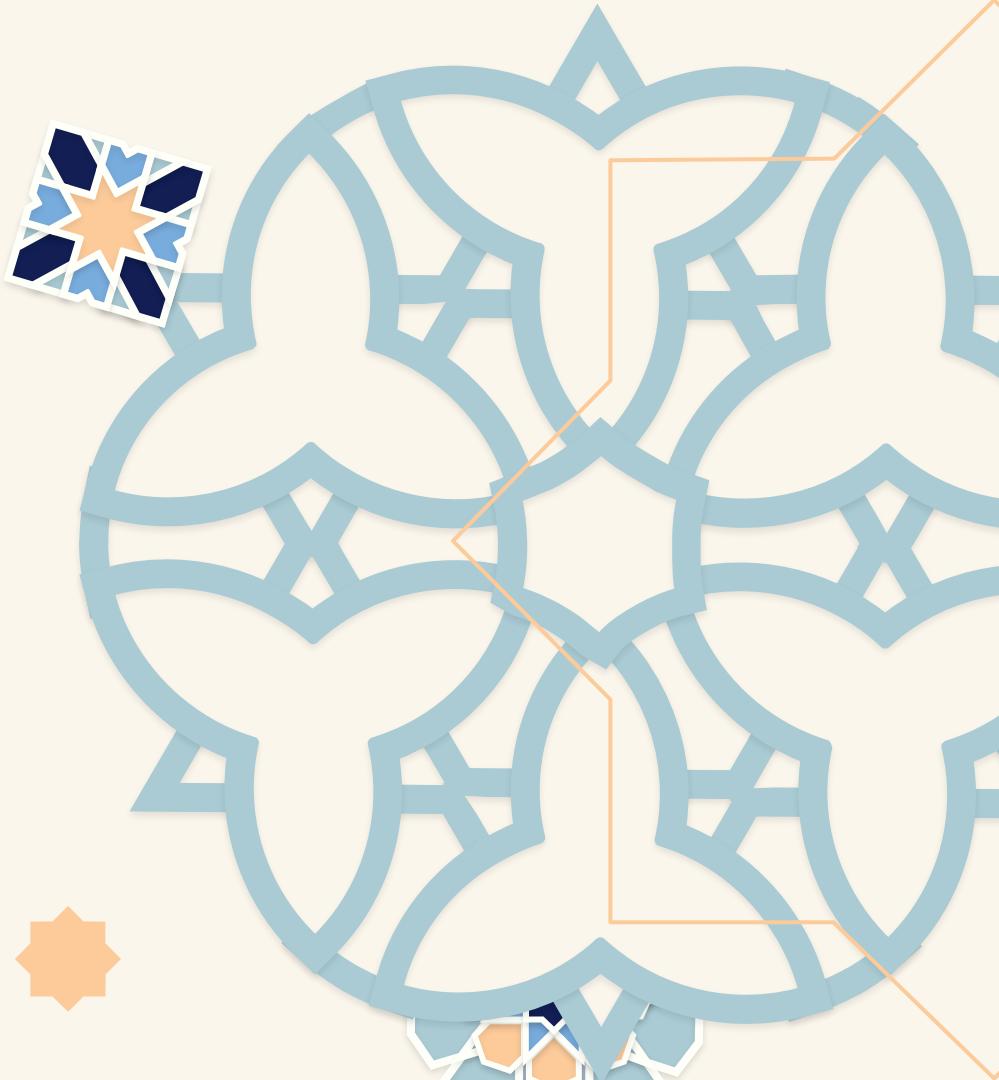
**1- Reading  
Comprehensive  
(main task)**

**2- Retriver.**

**3- Gradio interface**

# 1-Reading Comprehensive

---

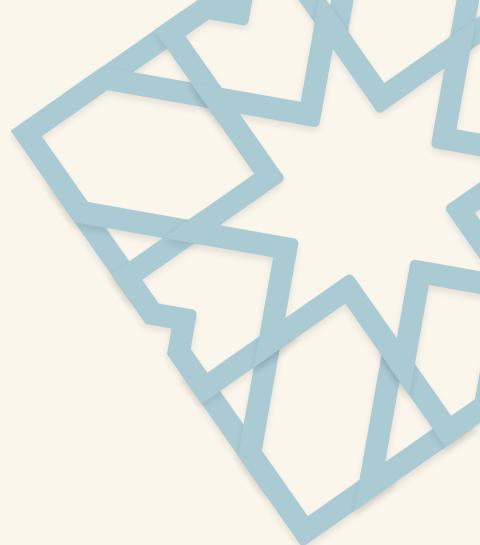




# Dataset



- The dataset for the task is called QRCD (Qur'anic Reading Comprehension Dataset).
- Two versions of the dataset were used
- QRCD is stored as a JSON Lines (JSONL) file, where each line is a JSON object.



# The first version

- It is composed of 1,093 tuples of question-passage pairs.
- The dataset also includes their extracted answers, which constitute 1,337 question-passage-answer triplets.
- QRCD is split into three sets: training (65%), development (10%), and test (25%).

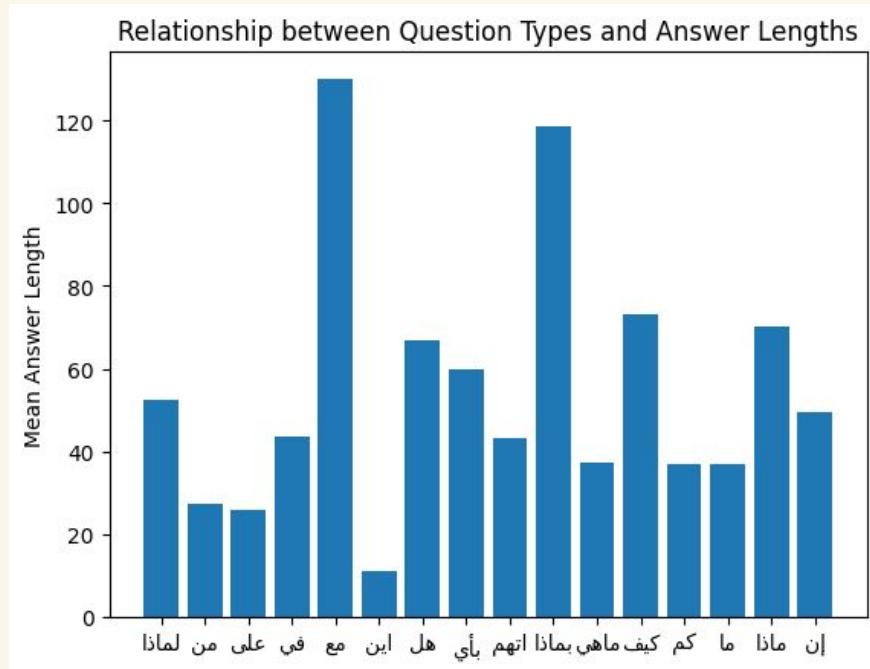
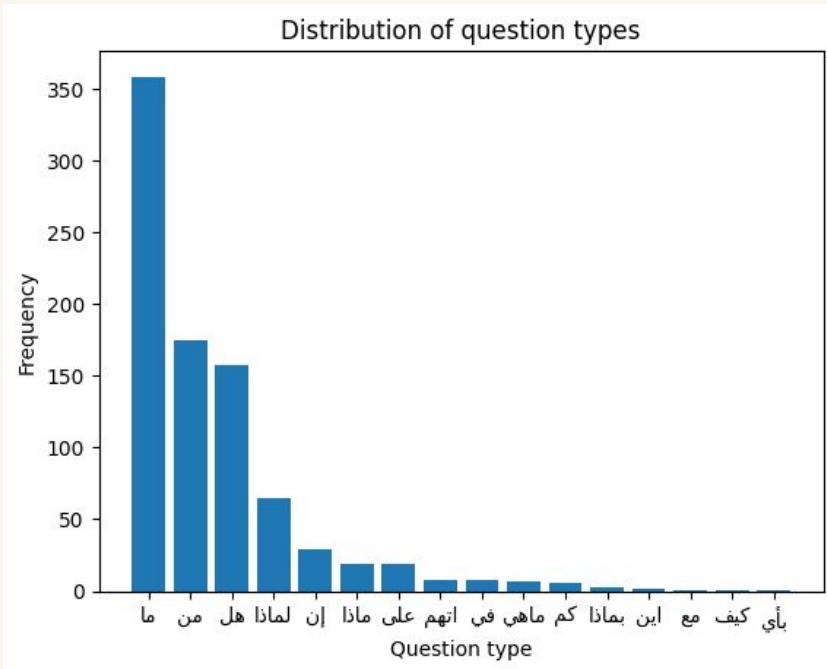
# The second version

- The dataset is composed of 1,155 tuples of question-passage pairs.
- These tuples are coupled with their extracted answers to constitute 1,399 question-passage-answer triplets .
- The test set is currently under development, and the above numbers are expected to increase in the future.



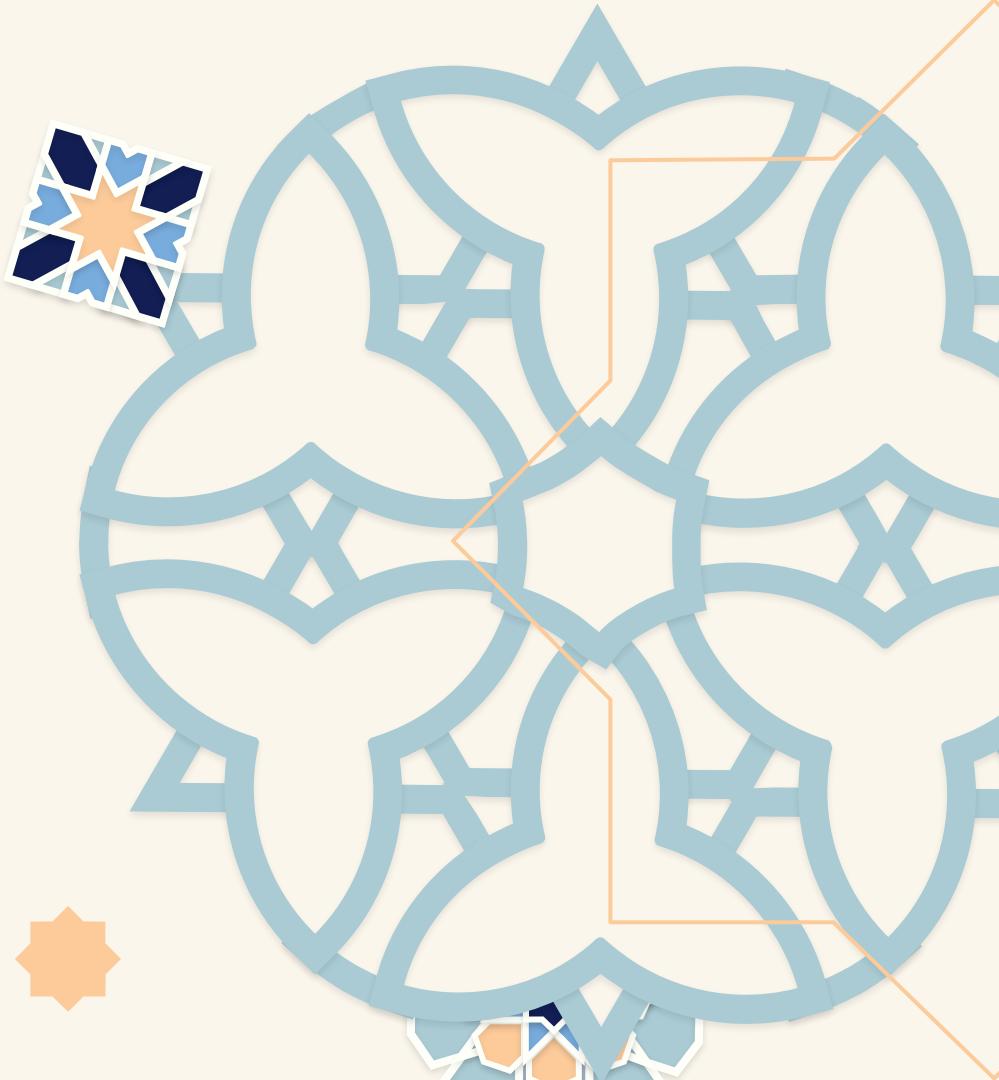
# Samples from the dataset

# Data Explorarion



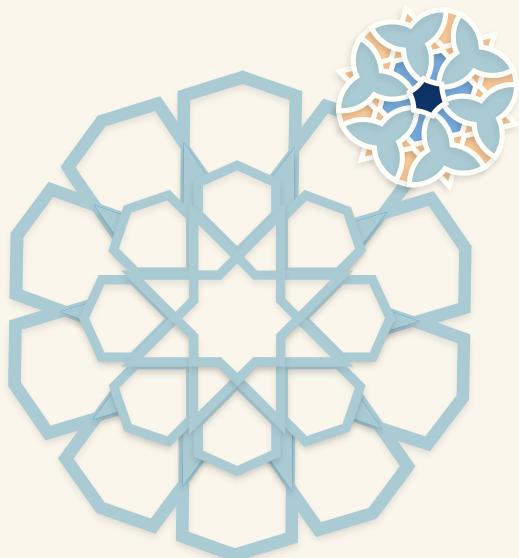
# Methods

---



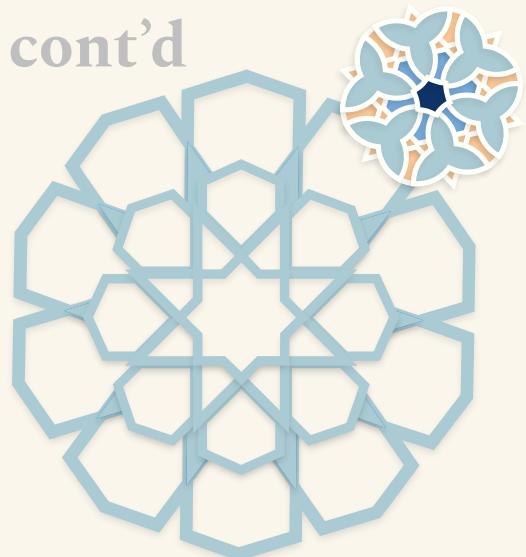
# Latent Semantic Indexing

- Latent semantic indexing (LSI) is a technique used to analyze and represent text documents.
- LSI transforms text data into a two-dimensional space where each document is represented by a point.



# Latent Semantic Indexing cont'd

- The position of the point is determined by how strongly the document relates to two different topics.
- LSI identifies the most important relationships between the words and the documents.





# Result with similarities

➡ Query (Question): حاسب ونعدب المال يوم القيمة ان كان ""من يضل الله فما له من هاد" كما ورد من قوله تعالى في آية 23 و آية 36 من سورة الزمر؟

Answers: [' أولئك الذين اشتروا الملاة بالهوى': 504]]

Passages sorted by similarity score (TF-IDF):

ومن الناس من يقول آمنا بالله وبالهوى الآخر وما هم بمؤمنين: 0.31:  
وإذا قيل لهم آمنوا كما آمن الناس قالوا أنؤمن كما آمن السفهاء ألا إنهم هم السفهاء ولكن لا يعلمون: 0.11:  
الله يستهزئ بهم ويهددهم في طغيانهم بعهمون: 0.10:  
أولئك الذين اشتروا الملاة بالهوى فيما ربحت تجارتكم وما كانوا مهتمدين: 0.07:  
في قلوبهم مرض فزادهم الله مرضًا ولهم عذاب أليم بما كانوا يذبون: 0.07:  
وإذا قيل لهم لا تفسدوا في الأرض قالوا إنما نحن مصلحون: 0.05:  
يخدعون الله والذين آمنوا وما يخدعون إلا أنفسهم وما يشعرون: 0.04:  
ألا إنهم هم المفسدون ولكن لا يشعرون: 0.00:  
وإذا لقوا الذين آمنوا قالوا آمنا وإذا خلوا إلى شياطينهم قالوا إنما معكم إنما نحن مستهزئون: 0.00:

Passages sorted by similarity score (LSI):

يخدعون الله والذين آمنوا وما يخدعون إلا أنفسهم وما يشعرون: 0.93:  
ومن الناس من يقول آمنا بالله وبالهوى الآخر وما هم بمؤمنين: 0.81:  
أولئك الذين اشتروا الملاة بالهوى فيما ربحت تجارتكم وما كانوا مهتمدين: 0.78:  
ألا إنهم هم المفسدون ولكن لا يشعرون: 0.42:  
في قلوبهم مرض فزادهم الله مرضًا ولهم عذاب أليم بما كانوا يذبون: 0.38:  
وإذا قيل لهم آمنوا كما آمن الناس قالوا أنؤمن كما آمن السفهاء ألا إنهم هم السفهاء ولكن لا يعلمون: 0.34:  
وإذا لقوا الذين آمنوا قالوا آمنا وإذا خلوا إلى شياطينهم قالوا إنما معكم إنما نحن مستهزئون: 0.31:  
الله يستهزئ بهم ويهددهم في طغيانهم بعهمون: 0.18:  
وإذا قيل لهم لا تفسدوا في الأرض قالوا إنما نحن مصلحون: 0.17:

# Transformers

---

- AraElectra-Arabic-SQuADv2-QA
- mdeberta-v3-base-squad2
- araelectra-base-artydiqa
- arap\_qa\_bert
- AraElectra-base-finetuned-ARCD
- Bert\_Arabic-SQuADv2-QA
- arap\_qa\_bert\_large\_v2
- AraELECTRA-discriminator-QuranQA
- AraELECTRA-discriminator-SOQAL
- mdeberta\_Quran\_qa



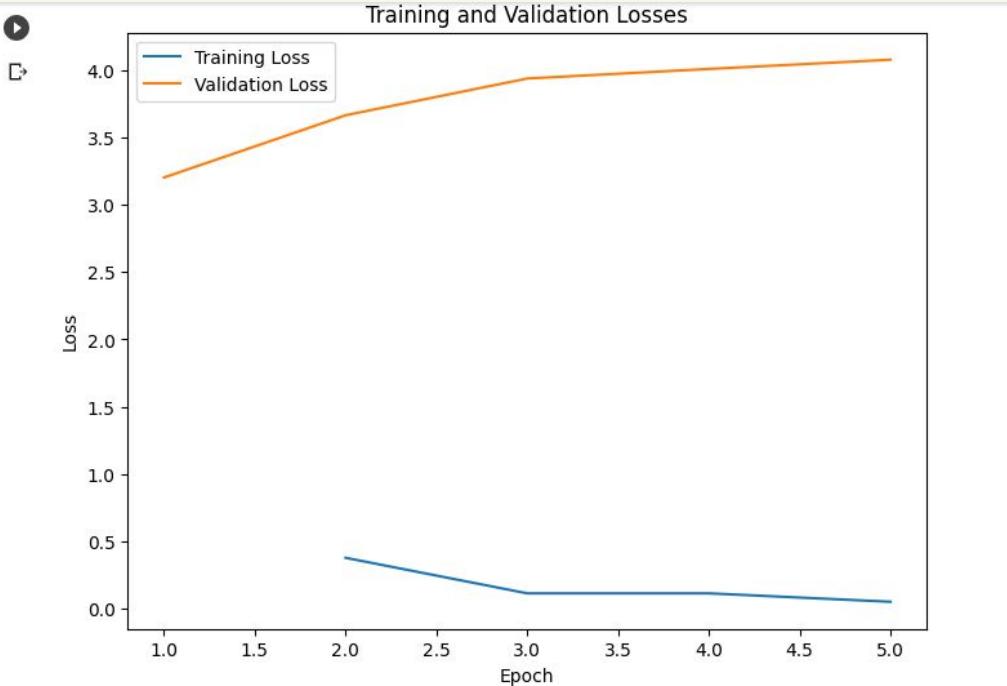
# Models Evaluation



From the figure  
the best model was AraELECTRA-discriminator-QuranQA



# Model fine tuning on version1





# prediction

Question: ما هي مصارف الزكاة؟

top predicted answers:

إنما الصدقات للقراء والمساكين والعاملين عليها والمؤلفة قلوبهم وفي الرقاب والغارمين وفي سبيل الله وابن السبيل: 1

Answer number 2: للقراء والمساكين والعاملين عليها والمؤلفة قلوبهم: 2

الصدقات للقراء والمساكين والعاملين عليها والمؤلفة قلوبهم وفي الرقاب والغارمين وفي سبيل الله وابن السبيل فريضة من الله: 3

Question: متى يحل الإسلام دم الشخص؟

top predicted answers:

كعب عليكم الفصاص في القتل الحر بالحر والعد بالعد والأنتي بالأنتي: 1

تى فمن عفي له من أخيه شيء فاتباع بالمعروف وأداء إليه بإحسان: 3

والأنتي بالأنتي فمن عفي له من أخيه شيء فاتباع بالمعروف وأداء إليه بإحسان ذلك تخفيف من ربك ورحمة فمن اعذني بعد ذلك فله عذاب أليم: 4

Answer number 5: الحر بالحر والعد بالعد والأنتي بالأنتي فمن: 5

Question: كم فترة رضاعة المولود؟

top predicted answers:

Answer number 1: حولين كاملين

Answer number 2: والوالدات يرضعن أولادهن حولين

Answer number 4: يرضعن أولادهن حولين كاملين لمن أراد أن يتم الرضاعة

Answer number 5: ن حولين كاملين لمن أراد أن يتم الرضاعة وعلى المولود له رزقين وكسوتين بالمعروف

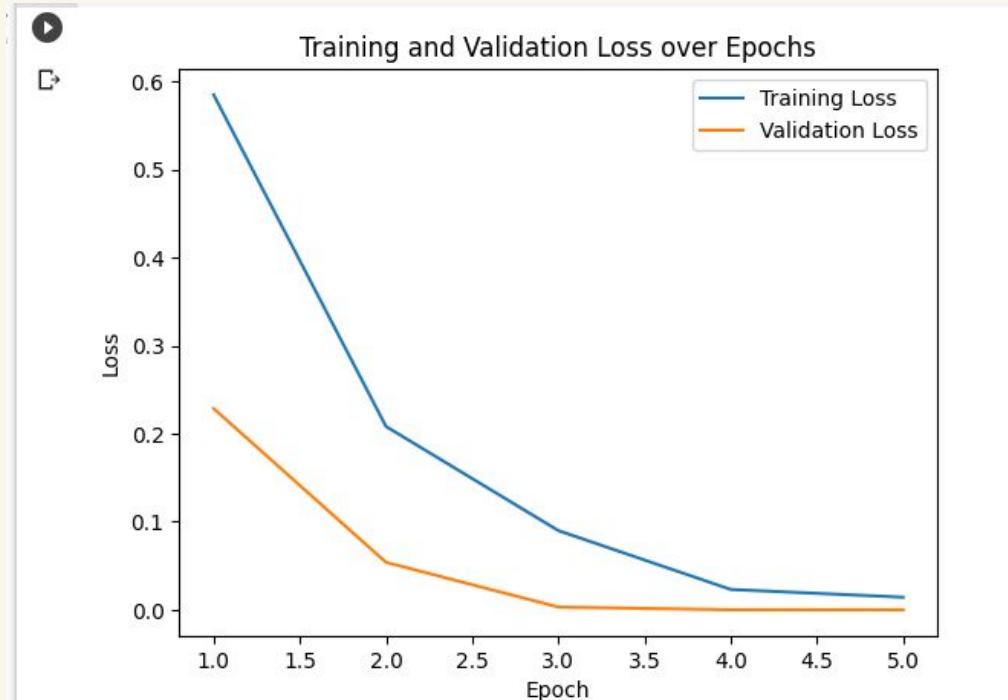
Question: ما حكم التعدد في الزواج؟

top predicted answers:

Answer number 1: فاتكروا ما طاب لكم من النساء متى وثلاث ورباع



# Model fine tuning on all data





# prediction

Question: ما هي كفاره الظهار؟

top predicted answers:

Answer number 1: ير رقة من قبل أن يتماسا ذلكم توعظون به والله بما تعلمون خير. فمن لم يجد فصيام شهرين متتابعين من قبل أن يتماسا فمن لم يستطع إلطعام ستين مسكيتاً فمن لم يجد فصيام شهرين متتابعين من قبل أن يتماسا فمن لم يستطع إلطعام ستين مسكيتاً ذلك: 5

Question: كم عدد حملة العرش؟

top predicted answers:

Answer number 1: ثماني

Answer number 4: واحدة. وحملت الأرض والجبال فدكتا دكة واحدة. فيومئذ وقعت الواقعه. وانتقت السماء فهي يومئذ واهية. والملك على أرجائها ويحمل عرق ربك فرقهم

Answer number 5: #ثماني#

Question: من هم المطقوفون؟

top predicted answers:

Answer number 1: الذين إذا اكلوا على الناس يستوفون. وإذا كالوهم أو وزنوه يخسرون

Question: ملأ يتصل الإحسان؟

top predicted answers:

Answer number 1: وما كان قولهم إلا أن قالوا ربنا أغفر لنا ذنبينا وإسرافنا في أمرنا وتبت أقدامنا وانصرنا على القوم الكافرين

Answer number 3: كان قولهم إلا أن قالوا ربنا أغفر لنا ذنبينا وإسرافنا في أمرنا وتبت أقدامنا وانصرنا على القوم الكافرين

Answer number 4: قولهم إلا أن قالوا ربنا أغفر لنا ذنبينا وإسرافنا في أمرنا وتبت أقدامنا وانصرنا على القوم الكافرين

Answer number 5: أن قالوا ربنا أغفر لنا ذنبينا وإسرافنا في أمرنا وتبت أقدامنا وانصرنا على القوم الكافرين

Question: هل ذكر القرآن أوقات نوم معينة؟

top predicted answers:

Answer number 1: الله الذي جعل لكم الليل لتسكنوا فيه

Answer number 3: الذي جعل لكم الليل لتسكنوا

Answer number 4: الليل

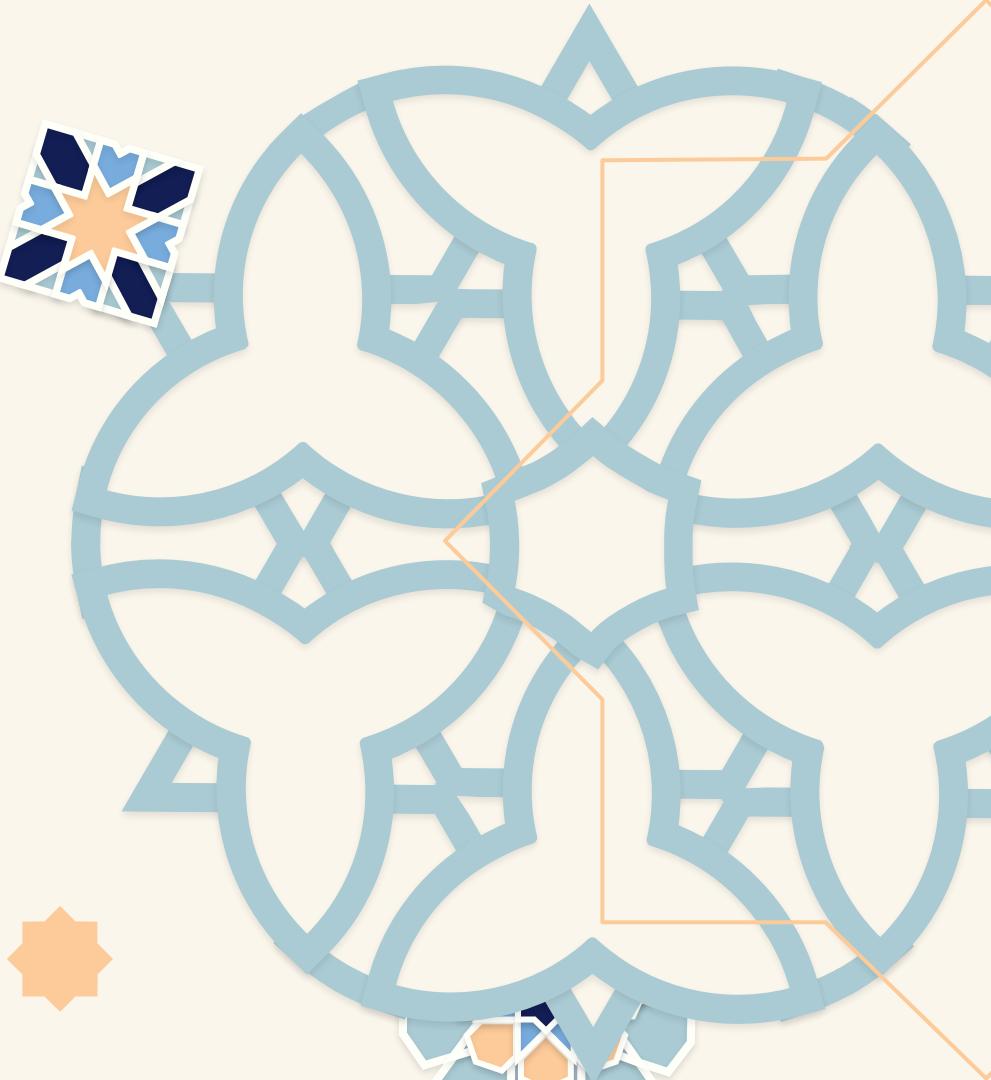
Answer number 5: جعل لكم الليل لتسكنوا فيه والنهار





# Evaluation

---





# Evaluation Matrices

---

**partial Reciprocal Rank (pRR)** : It is a variant of the traditional *Reciprocal Rank* evaluation metric that considers *partial matching*.

**Exact Match (EM)** : the EM metric is a binary measure that rewards a system only if the top predicted answer matches one of the gold answers *exactly*, which are evaluation metrics applied on the *top predicted* answer only.

**F1@1**: the F1@1 metric measures the token overlap between the top predicted answer and the best matching gold answer , to get an overall evaluation score, each of the above measures is averaged over all questions.





# Evaluation Matrices cont'd

Metric	Training	Validation	Test
pRR	0.9889	0.6436	0.5177
Exact Match	0.9761	0.3761	0.2227
F1	0.9880	0.6315	0.4874

On version1 data



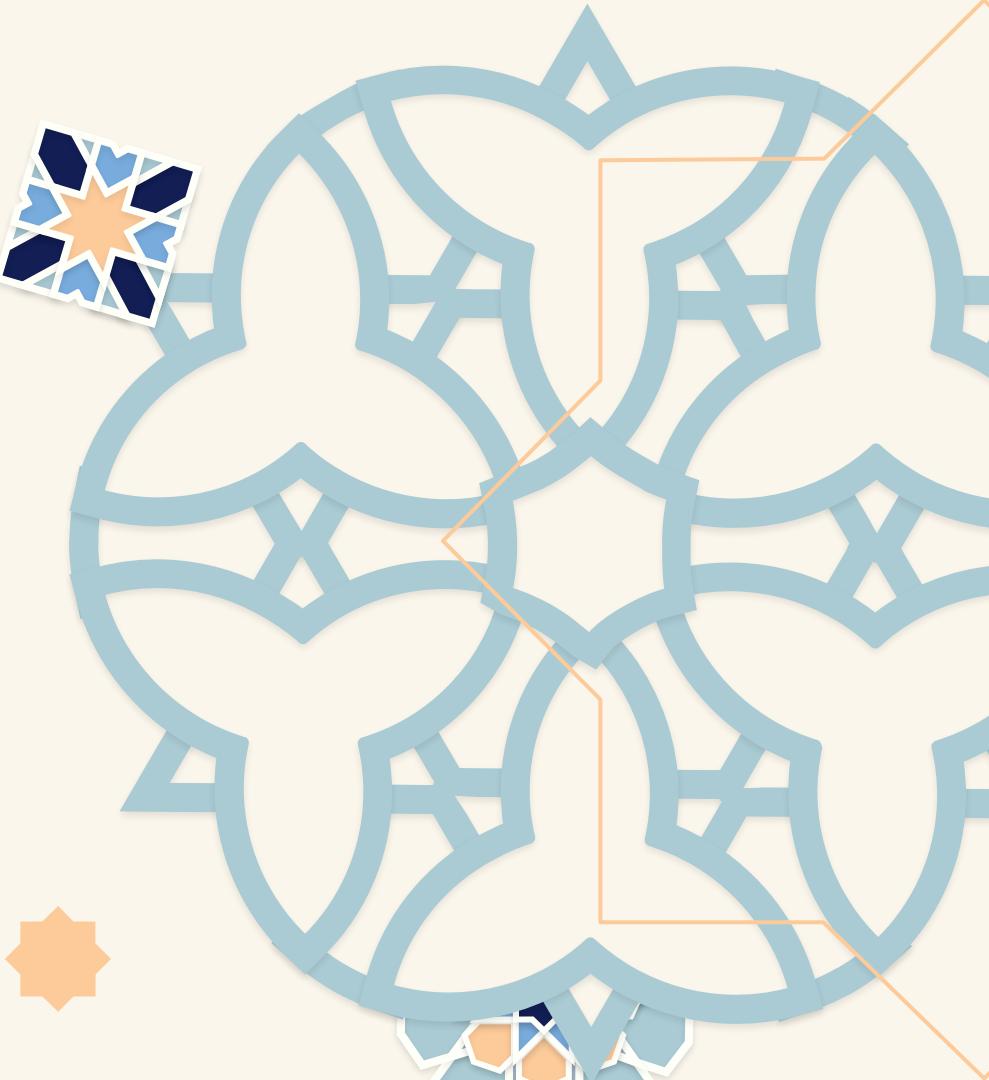
Metric	Training	Validation	Test
pRR	0.9631	0.9903	0.9193
Exact Match	0.9470	0.9358	0.8824
F1	0.9625	0.9903	0.9136

On version1&2 data



# 2-Retriver

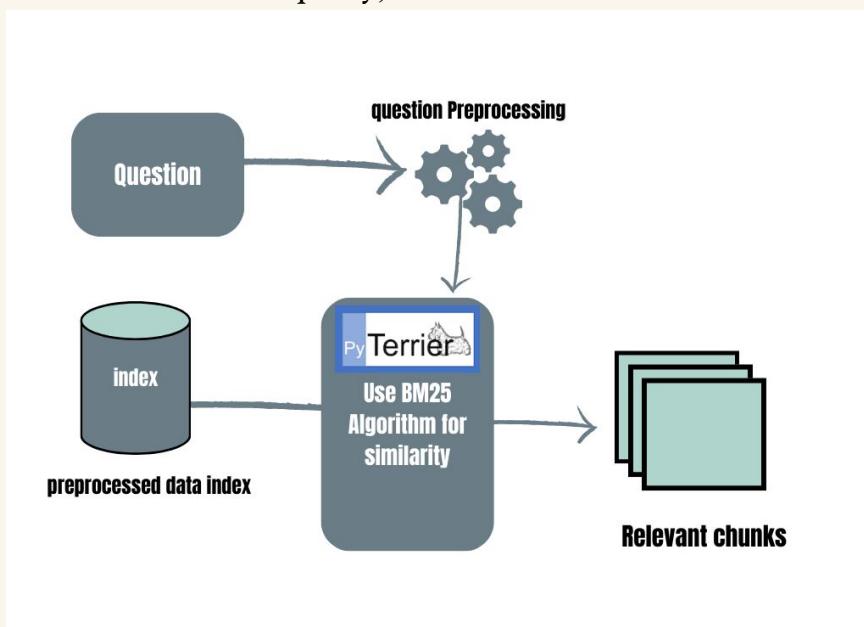
---





# First Method

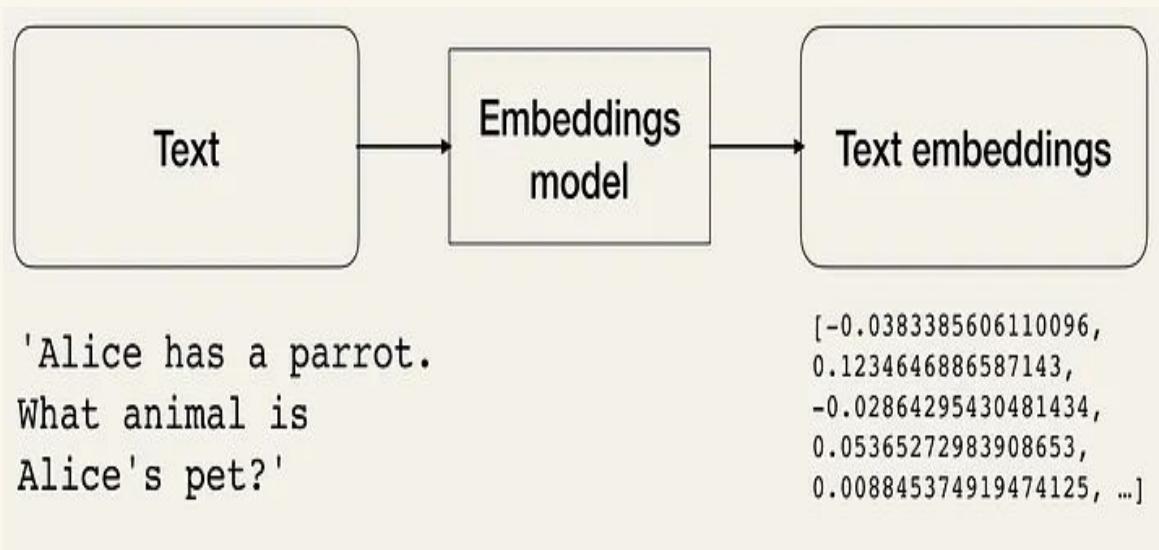
BM25 is a widely used probabilistic retrieval model that is commonly used in information retrieval. It is based on the bag-of-words model and considers the frequency of terms in the document and the query, as well as the inverse document frequency of the terms.





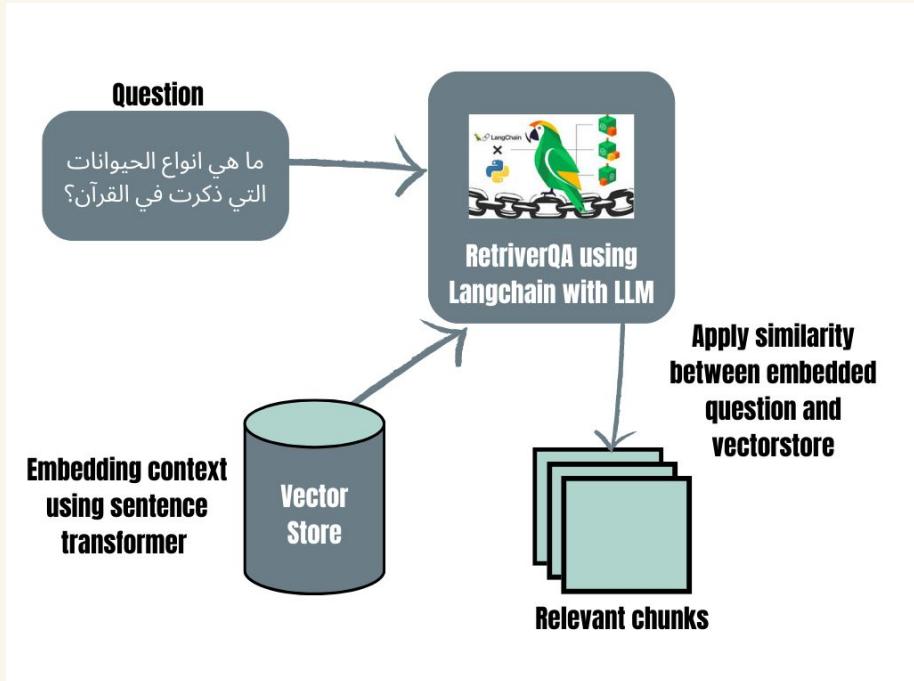
# Second Method

Lang chain with LLM



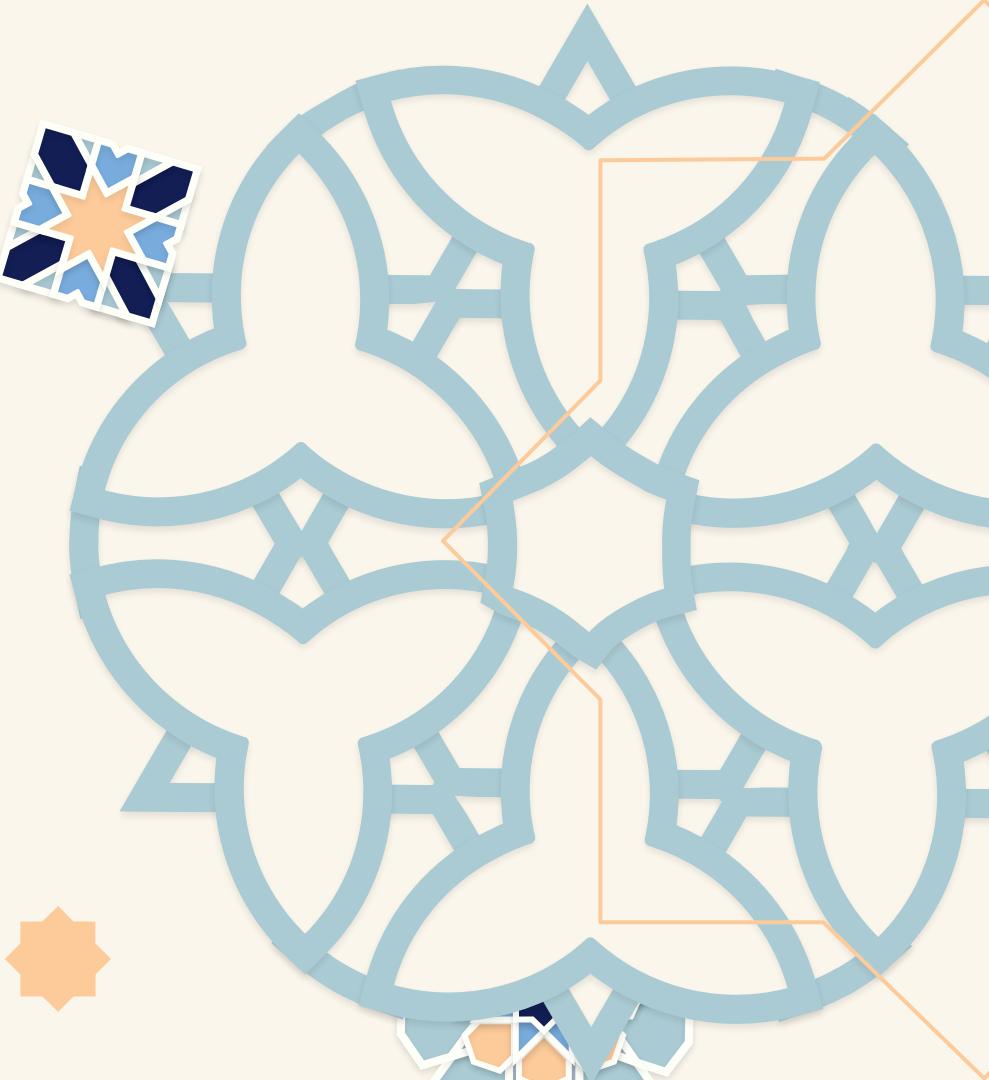
# Second Method

Lang chain with LLM



# 3- Interface

---



# Interface

---

- Gradio is a Python library used for creating web interfaces for machine learning models.
- It allows users to interact with the model and see its predictions in real-time.
- Gradio can be customized to match your branding or preferences.
- Gradio is easy to use and can be integrated with most machine learning frameworks.



# Demo

The screenshot shows a Microsoft Teams browser-based interface titled "استفسارات قرآنية" (Quranic Questions). The interface includes a yellow input field labeled "السؤال" (Question) and a green output field labeled "الإجابة" (Answer). The output field displays the following text:

شات بوت القرآن  
شك برب ربب على أسلنل الله الدليل من القرآن الكريم.  
كيفية الاستخدام:  
1. اكتب سوالك في مربع النص.  
2. انقر على زر "إرسال" للحصول على إجابة.

السؤال

الإجابة

من هم قوم موسى؟  
بني إسرائيل.

هل ورد في القرآن إشارة لمستوى ذيتأثير إيجابي على جسم الإنسان؟ ولذلـل من القرآن ما هو شفاء ورحمة للمسلمين

input

output 0

# Thank You

