SARAS AI
INSTITUTE

SARCathon 2024

Presented By : **Team METAVERSE**

Dev Kumar

Divya Gandhi

Devansh Upadhyay

IIT (BHU) Varanasi | 2024

INTRODUCTION

PROBLEM STATEMENT:
The task was to build a model that takes a user query and returns the most relevant questions from the list of FAQs.

- **To achieve this, we employed GloVe embeddings to convert words into vector representations. We developed a prediction function which uses similarity scores between the questions and queries to identify the closest matching questions.**
- **The model was deployed on Streamlit and hosted on Streamlit Cloud platform, ensuring fast, accurate results while being both scalable and resource-efficient.**

OVERVIEW

1. Utilized GloVe embeddings (50-dimensional) to convert words into vector representations, effectively overcoming the challenge of insufficient data that limited the extraction of nuanced contextual embeddings. This approach enhanced the model's semantic understanding.

2. Developed a prediction mechanism using Cosine similarity with normalized vector embeddings to efficiently identify the closest matching questions to user queries, achieving accurate and relevant results compared to traditional sentence vector similarity methods.

3. Optimized the program's import structure to enhance resource consumption, reducing its RAM footprint and minimizing data retrieval overhead during runtime. This refinement, alongside the deployment on Streamlit, significantly improves overall performance, scalability, and user experience.

METHODOLOGY

Traditional Method

Compute cosine similarity between the sentence vectors of the query and questions by averaging their respective word embeddings.

Proposed Method

Calculate cosine similarities between individual words of questions and queries and use the mean of the highest scores as the overall score for the question.

```python
def word_similarities_glove(user_query):
    blob = TextBlob(user_query)
    corrected_query = str(blob.correct())

    query_vector = get_glove_words_vector(corrected_query, glove_embeddings)
    query_normalised = query_vector/np.linalg.norm(query_vector, axis=1, keepdims=True)
    max_length = max(len(q) for q in question_vectors)
    question_array = np.array([np.pad(question, ((0, max_length - question.shape[0]), (0, 0)),
                    mode='constant', constant_values=0) for question in question_vectors])
    norms = np.linalg.norm(question_array, axis=2, keepdims=True)
    questions_normalised = np.where(norms != 0, question_array / (norms + 1e-8),question_array)

    similarity = np.tensordot(questions_normalised, query_normalised, axes=([-1], [-1]))
    similarity = np.nan_to_num(similarity, nan=0)
    similarity = similarity.reshape(22, -1)
    similarity = np.mean(np.sort(similarity, axis=1)[:, -(max(2, len(query_vector))):], axis=1)

    best_match_idx = np.argsort(similarity)[-5:][::-1]
    results = []
    for idx in best_match_idx:
        results.append((questions[idx], answers[idx]))

    return results
```

Intuition: The mean sentence vector can obscure context and word similarities. Our method focuses on the top matching scores to preserve relevance and improve the accuracy of question matches.

RESULTS

Saras AI - FAQ Search System

Ask a question to find relevant answers:

Enter your query:

fees structure of the course

Top 5 FAQs based on your query are:

Q: Are there any payment plans or options available for tuition fees?

A: Yes, we offer flexible payment plans to help students manage their tuition fees. At Saras AI Institute, you can pay your annual tuition fees in 5 installments, before the commencement of every semester.

Q: What are the tuition fees for your courses?

A: You can find detailed information and breakdown of the fee on 'Programs' page on the website

---

Q: What does the program structure look like, and how is the curriculum delivered?

A: Each year is divided into 5 semesters which last for 8 weeks each. Our programs feature a mix of recorded and live sessions, allowing for flexibility in learning.

Q: Can you provide more details about the role-based curriculum feature and how it benefits students?

A: Our role-based curriculum is designed to provide targeted training and develop specialized skills in students that are highly relevant to their desired job-roles from day one.

Q: Do employers require an accredited degree?

A: An accredited degree is not an absolute requirement from employers. We ensure our students are among the most-skilled and ready individuals to crack the best of jobs

Saras AI - FAQ Search System

Ask a question to find relevant answers:

Enter your query:

duration of the course

Top 5 FAQs based on your query are:

Q: What does the program structure look like, and how is the curriculum delivered?

A: Each year is divided into 5 semesters which last for 8 weeks each. Our programs feature a mix of recorded and live sessions, allowing for flexibility in learning.

Q: Can I connect with mentors outside of class?

A: Yes, we encourage mentorship and provide opportunities for students to connect with mentors outside of class through live sessions as well as 24x7 mentor support to help resolve your doubts or queries.

Q: What are the tuition fees for your courses?

A: You can find detailed information and breakdown of the fee on 'Programs' page on the website

---

Q: What is pre-enrollment assessment test? How do I prepare for it?

A: It is a fully online assessment which takes less than 15 minutes. It is designed to evaluate your personal traits and basic English communication skills. You can take it at the time of filling out the application. It idoes not require any specific preparation

---

Q: Does Saras AI Institute offer any scholarships for students? How can I apply for them?

A: Yes, we offer various scholarships to eligible students based on academic merit, financial need, and other criteria. You can apply for scholarships after you're offered admission. Go ahead with filling out the application to check your eligibility.

Saras AI - FAQ Search System

Ask a question to find relevant answers:

Enter your query:

can you please provide me the details of the program

Top 5 FAQs based on your query are:

Q: Can you provide more details about the role-based curriculum feature and how it benefits students?

A: Our role-based curriculum is designed to provide targeted training and develop specialized skills in students that are highly relevant to their desired job-roles from day one.

Q: Are there any specific requirements or prerequisites for admission into the programs?

A: To be a successful professional in AI, you need to possess basic mathematical proficieny - which can be demonstrated by your math scores in high school or beyond. At Saras, you learn with global peers and faculties and should possess basic communication skills in English. These make up for the basic eligibility critera

---

Q: Are there any payment plans or options available for tuition fees?

A: Yes, we offer flexible payment plans to help students manage their tuition fees. At Saras AI Institute, you can pay your annual tuition fees in 5 installments, before the commencement of every semester.

Q: What does the program structure look like, and how is the curriculum delivered?

A: Each year is divided into 5 semesters which last for 8 weeks each. Our programs feature a mix of recorded and live sessions, allowing for flexibility in learning.

Q: Does the university offer internship placement assistance?

A: Yes, we assist students in finding internships by connecting them with potential employers and offering guidance on applications and interviews.

CONCLUSION

- In conclusion, our model demonstrates exceptional scalability and flexibility, effectively returning relevant FAQs based on user queries.
- Its innovative approach, leveraging custom built function based on cosine similarity, enhances its effectiveness in delivering accurate results.
- The user-friendly web application further simplifies access to this robust solution, making it a powerful tool for efficiently addressing user inquiries and improving overall user experience.

Links: GitHub Streamlit