UE21CS334 Natural Language Processing Project

RAM-GPT

PRESENTED BY

SANDEEP RAM

JYOTIRADITYA J

K GANESH VAIDYANATHAN

DEPARTMENT OF CSE PES UNIVERSITY

UNDER THE GUIDANCE OF

DR. MAMATHA.H.R

PROFESSOR

DEPARTMENT OF CSE

PES UNIVERSITY

Agenda

- Introduction
- Problem Statement
- Literature Survey
- References

Introduction

Ram-GPT - your personal guide on the great epic Ramayana

Slides under introduction should be arranged in following fashion

Slide 1: Problem identified

Slide 2: Evidence of problem

Problem Statement

Problem Identified

- Design and develop a Ramayan-focused GPT (Generative Pre-trained Transformer) that serves as a knowledge assistant for users seeking information, teachings, and insights from the Ramayan epic.
- The goal is to create a chat bot capable of answering user queries related to the Ramayan, catering to individuals interested in rediscovering their cultural roots, extracting important teachings, or refreshing their memory on the epic.

Evidence of the problem

1.User Engagement and Exploration:

Problem: Limited engagement and exploration opportunities for users interested in the Ramayan.

Solution: The application encourages users to explore deeper into the epic, providing interactive elements and links to related resources, making the learning experience more engaging and enriching.

2. Lack of educational resource and organized learning

Problem: Many individuals face a challenge in accessing structured and organized learning materials for the Ramayan, resulting in a lack of educational resources that specifically focus on the ethical and moral teachings of the epic.

Solution: To address this issue, the application combines the benefits of organized learning by structuring information in a user-friendly format. It also serves as a comprehensive educational resource, emphasizing the ethical and moral lessons from the Ramayan, thereby offering users a valuable tool for systematic learning, personal growth, and character development. The interactive features and suggested readings further enhance the overall learning experience.

Evidence of the problem

3. Cultural Disconnection:

Problem: Cultural disconnection among individuals who want to reconnect with their roots and heritage.

Solution: The GPT model offers insights into the cultural, moral, and spiritual teachings embedded in the Ramayan, fostering a sense of connection and understanding of one's cultural identity.

Literature Review

THIS IS TO REVIEW HOW THE PROBLEM IS SOLVED IN LITERATURE

RESEARCH AREA

Proposed Solution

1. **Data Preparation**: We plan to collect and process the Ramayana and interpretations of the Ramayana by different authors.

2. **Large Language Models**: Through the usage of existing large models such as BERT, GPT-2, etc, we plan on training these models on our corpus of data and evaluate the performance of these models on validation dataset

3. **Analysis Of The Models**: Compare the performance of each of the different models on our validation dataset.

Datasets & References!

- https://www.kaggle.com/datasets/subodh29/valmiki-ramayan-sanskrit
- https://www.kaggle.com/code/srinivasaraghavanr/ramayana-gpt2

Reference

[1] Thottempudi, S.G., A Visual Narrative of Ramayana using Extractive Summarization, Topic Modeling and Named Entity Recognition.

[2] Srijeyarankesh, S., Kumaran, A., Nithayshri, L. and Shanmuga, M., 2022, December. A Comprehensive Study of Mahabharat using Semantic and Sentiment Analysis. In *Proceedings of the 19th International Conference on Natural Language Processing (ICON)* (pp. 308-317).

[3] de Lhoneux1 Anoop, R.A.M. and Søgaard, K.A., 2021. Itihasa: A large-scale corpus for Sanskrit to English translation. *WAT 2021*, p.191.

Thank You