

CLOUDILIC-HHEM MODEL FOR HALLUCINATION DETECTION

Unhallucinate Challenge



Table of CONTENTS

01

Challenge

Overview

02

Demo

Overview

03

Demo

Description

04

Demo

Test



CHALLENGE OVERVIEW

- The Unhallucinate Challenge: 24 Hours Fight aims to build a flawless RAG-based app without any hallucinations using the newest Vectara's HHEM model.
- Participants have 24 hours to complete the challenge.
- The task involves implementing an RAG-based app by utilizing the HHEM model.



DEMO OVERVIEW

Our team has developed Cloudilic-HHEM, an application that uses Large Language Models (LLMs) to summarize PDF documents. It detects 'hallucinations' in summaries, ensuring accuracy and reliability. The application uses dynamic temperatures when invoking LLMs to understand their behavior better. Presented through a user-friendly Streamlit GUI, it allows users to interact with PDFs effectively.



Demo Description:

Chat with PDFs

Users can input any PDF document, and our tool extracts the text content for interactive chat sessions.

Model Selection:

- ChatGPT 3.5: Users can choose the ChatGPT 3.5 model for natural language interactions.
- LLama 2: Alternatively, they can opt for the LLama 2 model.

Temperature Control:

- A slider allows users to adjust the model's temperature dynamically.
- 0: No hallucinations; responses are grounded and factual.
- 1: Increased hallucinations; responses may become more creative but less reliable.

