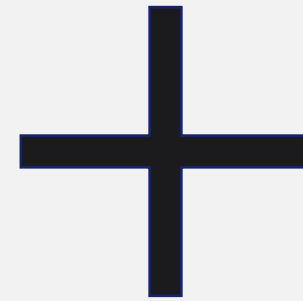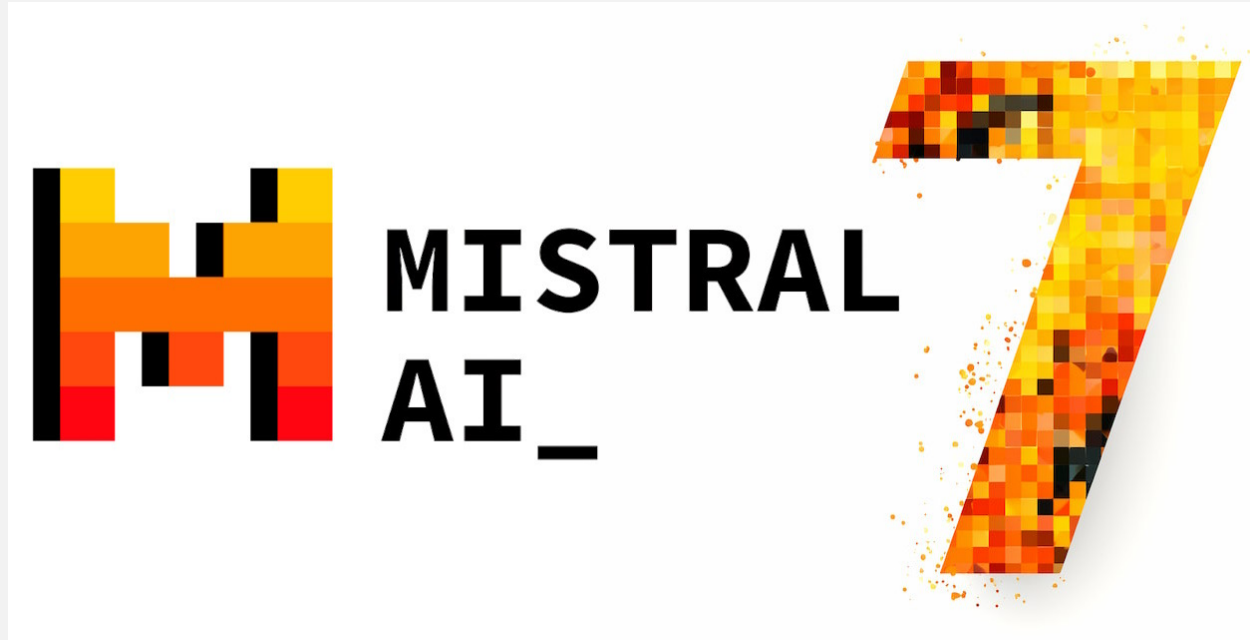# rslt

Hackathon Mistal 7B

RSLT Team| 20/08/2023
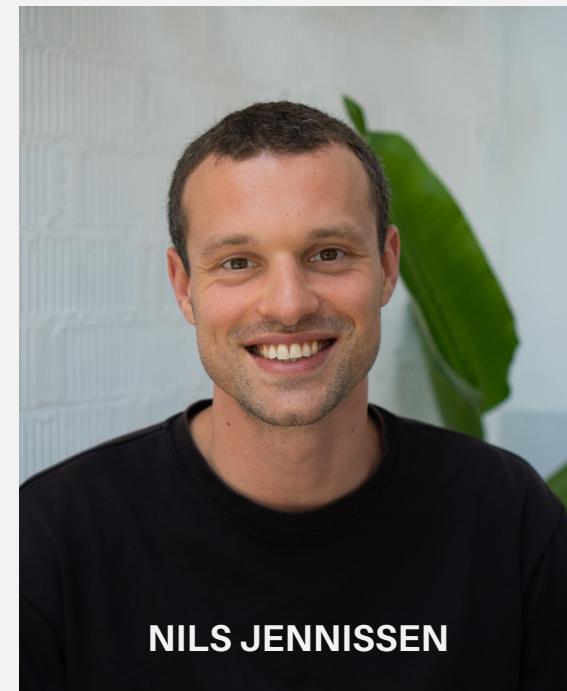
Mistral LLM and Chroma Vector Database RAG Application

RSLT Team

Meet our team

INTERDISCIPLINARY, INTERNATIONAL, INTERCONNECTED

rslt

RUBEN TAK

ONASSIS NOTTAGE

GABRIEL RENNÓ

NILS JENNISSEN

rslt

Introduction

Mistral LLM is a powerful language model that can be used for various applications such as

- text classification,

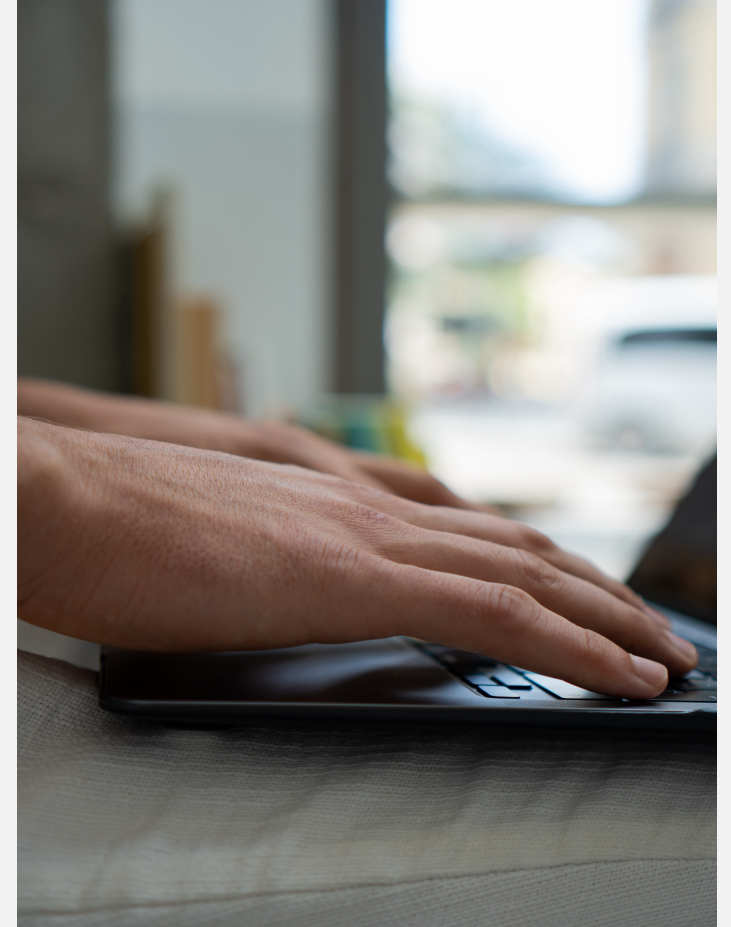- sentiment analysis,

- question-answering

It is a strong and impressive model. However, the full potential of Mistral LLM can only be realized when it is integrated with a vector database like Chroma Vector Database to create a RAG (Recallable AI Guide) application.
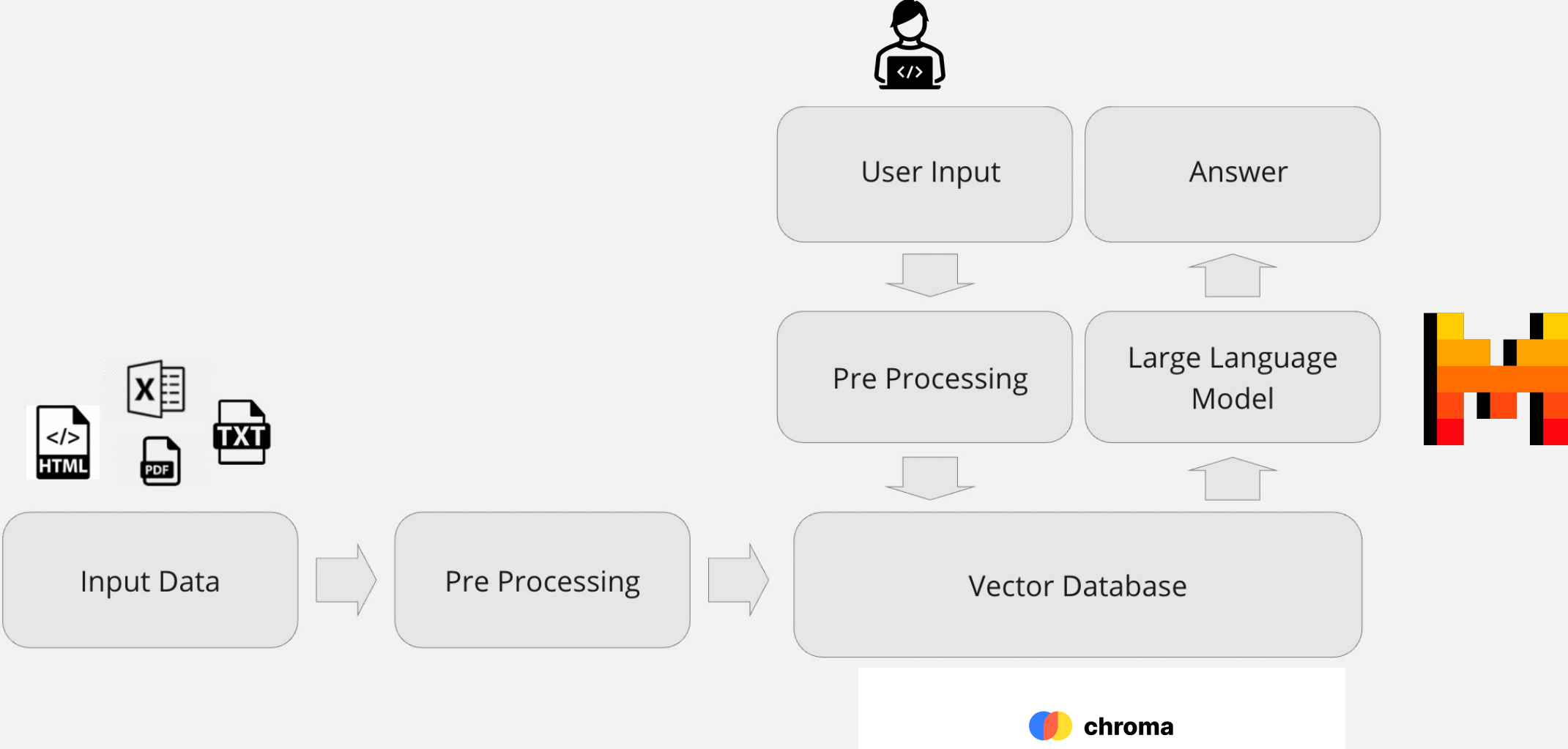

rslt
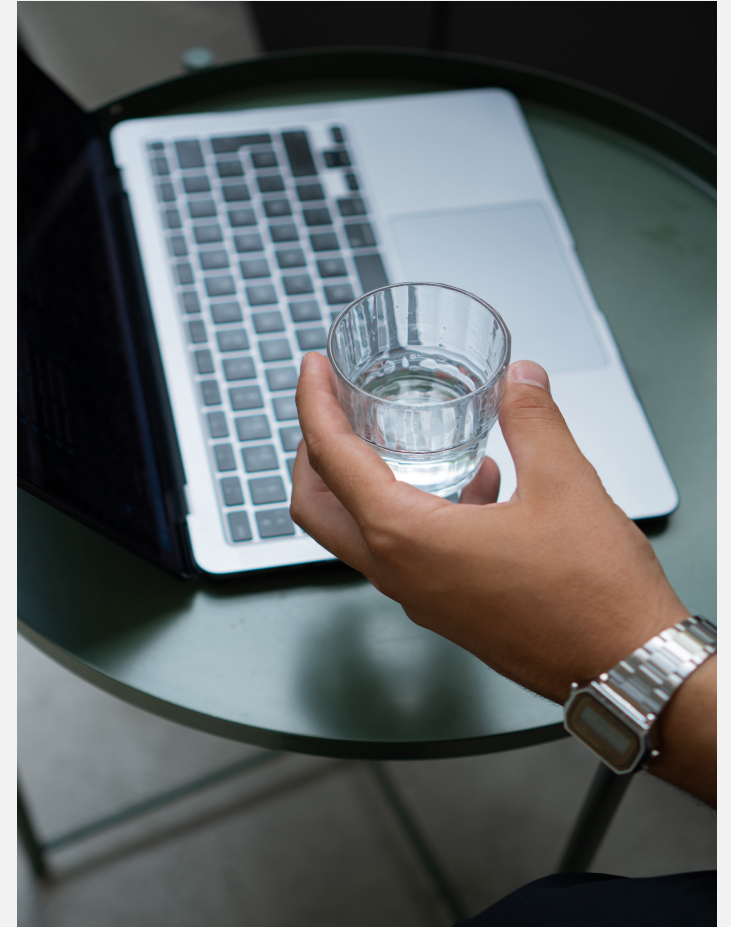
Problem Statement

Mistral LLM is a good and strong LLM for its size. However, it is not perfect and makes a lot of mistakes about specific topics.

There is a integration necessary with external know ledge. Also known as a RAG application.

How do LLM based applications work? (RAG)

Input Data → Pre Processing → Vector Database

User Input → Pre Processing → Vector Database → Large Language Model → Answer

chroma

Solution Overview

Our solution involves creating a script that connects Mistral LLM with Chroma Vector Database to create a RAG application.
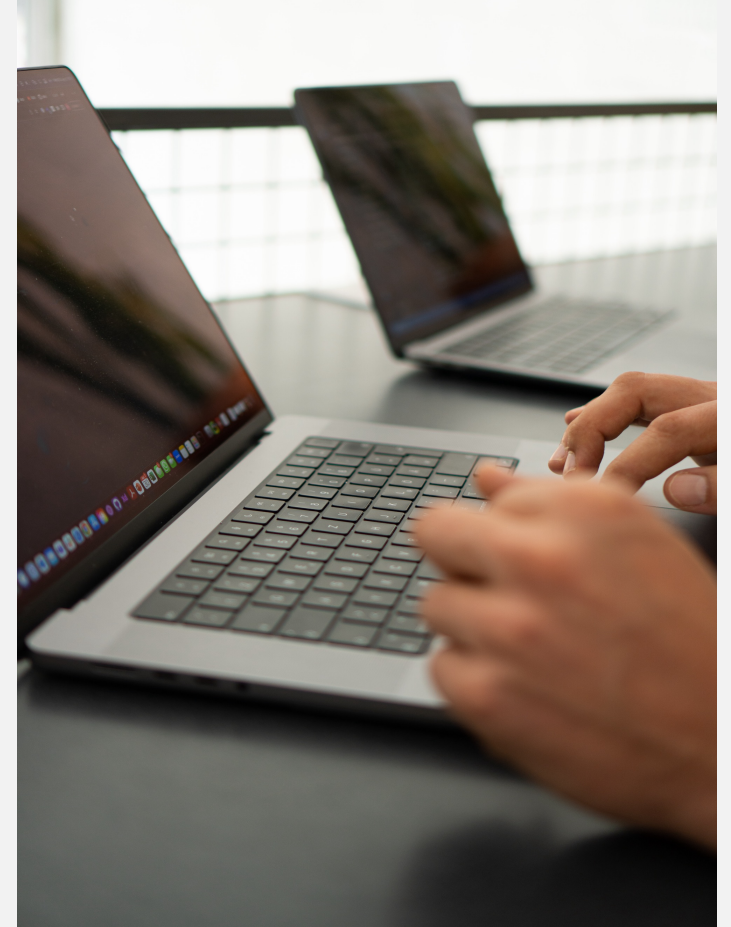
The Streamlit application that we created does 3 things

- Create a **normal chat** interface to converse with Mistral 7B

- Create a **RAG application** to let you chat with your documents and save them to a Chroma DB

- Make a connection with an **existing Chroma DB** and chat with this

rslt
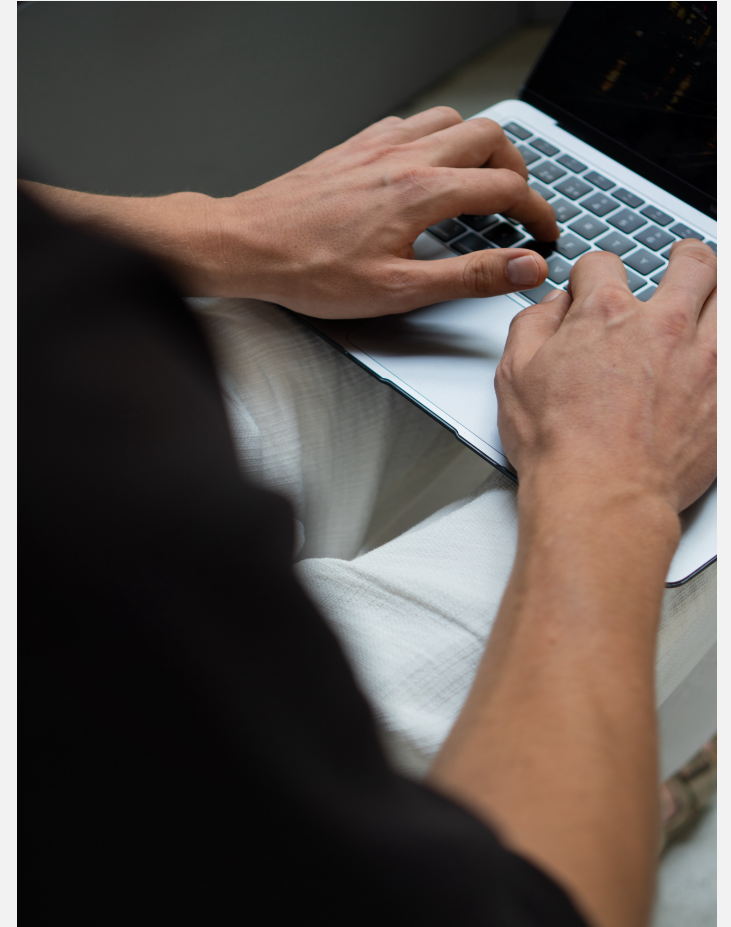
Benefits of the Solution

- **Improved accuracy and relevance**: RAG models can access and incorporate up-to-date information from a vast knowledge base, which helps them to generate more accurate and relevant responses to user queries. This is especially beneficial in domains where precision is critical, such as healthcare, finance, and legal services.

- **Context-aware responses**: RAG models can better understand the context and nuances of user queries by retrieving and incorporating relevant information. This enables them to provide more context-aware and helpful responses.

- **Adaptability**: RAG models can adapt to a wide range of user inputs, including open-ended, challenging, and strange questions. This makes them ideal for applications such as chatbots, question-answering systems, and content generation.

rslt

Demo!

- https://rslt.agency/

- https://www.linkedin.com/company/96695171/admin/feed/posts/

- https://github.com/nilsjenn

- https://medium.com/@gabriel_renno

- https://github.com/GabrielRenno

- https://github.com/rubentak

- https://www.linkedin.com/in/ruben-tak-665b66194/

- https://medium.com/@rubentak

rslt