



# The Challenge

Large Language Models have taken the world by storm!

This new frontier uncovers a multitude of opportunities and solutions in all industries.

However, it also comes with its own set of challenges...

- You are called to evaluate an LLM on its aviation knowledge;
- This will help the business understand if its use as an AI Assistant is feasible;
- To answer this question, the business provides you with a benchmark in the form of quiz;
- You need to run said questions through the LLM and submit the results;
- In doing so, you need to find ways of improving the LLM's answers.

# Requirements

## DO

Documents:

- Search and utilise publicly available documents that may contain information the LLM can use to answer correctly;

Embeddings:

- Use the provided embeddings;
- Use open-access embedding datasets, or vectorisation models;
- The embeddings used must be of 600 dimensions max – dimensionality reduction allowed;

LLM:

- You can only use Llama2 & Mistral models;
- Only up to 7 billion parameters allowed;
- Quantization up to Q8 permitted;
- Use through LM Studio or Notebook / .py script;

## DON'T

- More than 7 billion parameter models are prohibited;
- No paid services allowed (OpenAI, AWS, etc) for the purposes of embeddings;
- No paid services allowed (OpenAI, AWS, etc) for the purposes of answering the quiz directly;
- No API calls to external models (GPT, HuggingFace) for the purposes of embeddings or answering the quiz directly;

## IMPORTANT

Be prepared to share your code and demonstrate reproducible results.

Ideas:

- Embeddings;
- Prompt engineering;
- Retrieval Augmentation;
- LLM parameters;



# Resources

## Useful Links

Hugging Face

<https://huggingface.co/>

LM Studio

<https://lmstudio.ai/>

Models (gguf variations)

TheBloke/Llama-2-7B

TheBloke/Mistral-7B

Prompt Engineering

<https://www.promptingguide.ai/>

## Kaggle Competition

[https://www.kaggle.com/t/  
08d20a168ddc4471a8dcb08000354e45](https://www.kaggle.com/t/08d20a168ddc4471a8dcb08000354e45)

## Schedule

Saturday	Doors Open	10:00am
	Introduction	11:00am
	Challenge Start	11:30am
	Lunch	1:30pm
Sunday	Lunch	1:00pm
	Final Submissions	3:00pm
	Presentations	3:30pm
	Winners & Conclusion	5:00pm