

LLM Engineering

MASTER AI & LARGE LANGUAGE MODELS





PROGRESS REPORT

Less talking, more action!

What you can now do

- Generate text and code with Frontier Models including AI Assistants with Tools, and with open-source models with HuggingFace transformers
 - Create advanced RAG solutions with LangChain
 - Follow a 5 step strategy to solve problems, including dataset curation
-

After today you will be able to:

- Explain the role of a baseline model
- Create a traditional ML solution with features and linear regression
- Apply more advanced NLP techniques including SVR

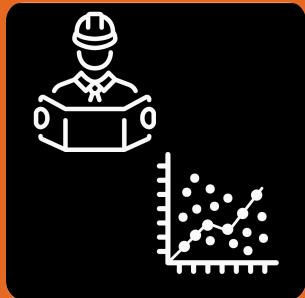
The importance of a Baseline

- Start cheap and simple
 - Gives a benchmark to improve on
 - An LLM might not be the right solution
-

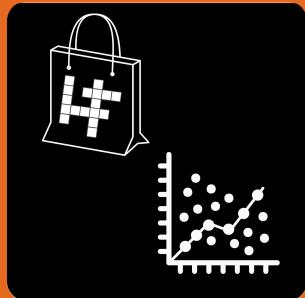


Traditional ML models

We will quickly set up these solutions to give us a starting point



Feature engineering & Linear Regression



Bag of Words & Linear Regression



word2vec & Linear Regression



word2vec & Random Forest



word2vec & SVR

Average absolute prediction error from our baselines



A stylized illustration of a person with a backpack walking along a path through lush green rice terraces under a bright sun.

PROGRESS

Next stop: the Frontier

What you can now do

- Generate text and code with Frontier Models including AI Assistants with Tools, and with open-source models with HuggingFace transformers
- Create advanced RAG solutions with LangChain
- Follow a 5 step strategy to solve problems, including dataset curation and making a baseline model with traditional ML

After the next session you will be able to:

- Build a framework to solve a commercial problem using a Frontier model
- Run our test dataset against GPT-4o-mini
- Run our test dataset against GPT-4o - the frontier version from August