HABA

GenAI Capability for Progressive SMEs

Created - Stephen Adebola & Andrita

HABA

HABA

Welcome to the AI Capability Workshop tailored for SMEs.

This workshop is designed to provide both technical and non-technical people with a comprehensive understanding of GenAI, its applications, and best practices especially around containment and alignment. Guardrails are going to be crucial for GenAI adoption in the real world

Over the course of two days,
participants will delve into the intricacies of models, frameworks, and the ethical implications of AI, create and experiment with their own GenAI applications, agents and chatbot among other topics.

HABA

Day One

HABA

Understanding AI and Its Foundations

Day One focuses on introducing participants to AI and its foundational concepts. The sessions are designed to be interactive, with hands-on exercises complementing theoretical knowledge.

The day's highlights include:

- Demystifying Models, Weights, and Biases
- The differences between tunning vs Prompting
- Introduction to AI frameworks with a spotlight on Langchain
- Exploring no-code AI solutions with LangFlow and Flows

HABA

Day Two

HABA

Implementing, Testing, and Safeguarding AI

Day two shifts the focus towards implementing, testing, and safeguarding AI solutions. Participants will gain insights into testing AI models, addressing safety concerns, and understanding the ethical considerations in AI.

The day's highlights include:

- Testing and evaluating models using tools like LangSmith and LangForge
- Introduction to GuardRails for AI to ensure business safety
- Practical hands-on session to build a simple AI project

HABA

Key Takeaways at the end of this workshop, participants will

- Have a clear understanding of AI, models, weights, and biases.
- Be familiar with various AI frameworks and tools.
- Gain hands-on experience in building AI applications.
- Understand the importance of safety and ethics in AI.
- Be equipped with the knowledge to scale AI solutions for their businesses.

HABA

Tunning Vs Prompting

HABA

Tuning

Modular Fine-Tuning of Pre-existing Models
For domain-specific tasks, Fine Tuning of existing models is vital. It customizes pre-trained models for specific tasks using domain data. However, this can be costly, as it modifies millions of parameters, needing extensive data and infrastructure. Full fine-tuning risks catastrophic forgetting and results in large, non-modular models. Solutions like Parameter Efficient Fine Tuning (PEFT) offer modular fine-tuning that conserves resources and is cost-efficient.

Your Objectives
&
Key Results

GenAI
Frameworks &
Applications

Your Models LLM

Tunning:(PEFT-LoRA, QLoRA)

Llama 2
GPT-3-Turbo
PALM 2

Prompting

Prompt Engineering Pre-existing Models
Fundamentally, obtaining anticipated results from Large Language Models requires meticulous engineering of prompts. This method entails designing appropriate prompts and inputs to draw out the specific reactions from the model. Engineering prompts is a crucial approach for numerous applications, particularly when general feedback is adequate.

Prompting: (MRKL, REAct, COT, RAG, Zero Shot)

Llama 2
GPT-3-Turbo
GPT-4
PALM 2
Claude 2

HABA

Glossary Terms

Attention
Model architecture
Completion
Low rank adaptation (LoRA)
Fine tuning
MRKL
Artificial Intelligence (AI)
Reinforcement Learning from Human Feedback (RLHF)
One-shot / Few-shot
Parameters
Transformer
Prompt
REAct
Retrieval Augmented Generation (RAG)
System prompt
Chain-of-thought
Generative pretrained transformers (GPT)
Plugins / tools
Generative AI
Embeddings
Multi-modal
Hallucination
Prompt injection
Large language model (LLM)
Training
Agents
Prompt engineering
Neural network
Token
ChatGPT
Foundational model
Alignment

HABA

Term	Description
Neural network	Network modelled on the brain
Parameters	Weights that control neural network calculations
Model architecture	Components of a complex AI model
Training	Improving model performance on data
Generative AI	Models that generate text/images from prompts
Generative pretrained transformers (GPT)	Popular large language model
ChatGPT	Conversational version of GPT
Large language model (LLM)	AI model that handles language
Transformer	Popular neural network architecture
Token	Encodes text numerically for models
Embeddings	Represent words/text semantically
Attention	Allows models to understand context

HABA

Term	Description
Alignment	Steers models towards ethical output
Foundational model	Broadly trained model
Fine tuning	Tailoring model to specific tasks
RLHF	Reinforcement Learning from Human Feedback to improve models
Low rank adaptation (LoRA)	Efficient fine-tuning method
Multi-modal	Handle mixed text/image input
Prompt	Text input to models
Completion	Text output from models
Hallucination	Fictional/incorrect output
One-shot / Few-shot	Types of prompting
System prompt	Defines model characteristics
Prompt engineering	Developing effective prompts

HABA

Term	Description
Prompt injection	Subverting models via input
Chain-of-thought	Improves reasoning via breakdown
REAct	Read Evaluate and React
MRKL	Modular Reasoning, Knowledge & Language
Agents	Versatile AI tools
Plugins / tools	Expand capabilities via APIs
Retrieval Augmented Generation (RAG)	Supplementing with searches

HABA

Rate Card – GenAI Services

HABA

GenAI Services	Rate
GenAI Pro Introduction 180 mins	£999
GenAI Basic Introduction 90 mins	£499
GenAI Pro Hands-on Workshop	£2499
GenAI Enterprise 2-day Hands-on Workshop	£24,499
General Cloud Architecture	£999 p/d
GenAI Application Architecture	£2499 p/d
GenAI 15min Inro/Demo 15min	Free