LA TECHNOLOGY TRANSFER PRESENTS

IVAN REZNIKOV

CHATBOT AND LLM BOOTCAMP

Conversational Agents using Modern Natural Language Processing

ONLINE LIVE STREAMING

MAY 15-16, 2024

DUE TO TIME ZONES, THIS CLASS WILL TAKE PLACE IN 2 AFTERNOONS
FROM 2 PM TO 6 PM ITALIAN TIME

info@technologytransfer.it
www.technologytransfer.it

ABOUT THIS SEMINAR

ChatGPT fundamentally changed the world of Artificial Intelligence (AI). ChatGPT is a Chatbot driven by the GPT3.5/GPT4 Large Language Models (LLMs) from OpenAI. These models are neural networks with billions of parameters, trained on vast amounts of text from the Internet and other sources and provide human-level intelligence on a wide variety of tasks. ChatGPT makes them available through an interactive, conversational prompt. Users can ask it to do anything. ChatGPT power users quickly sought to solve new and unexpected problems with ChatGPT and introduced the world to the field of *Prompt Engineering.*

In Prompt Engineering, the description of the task that the AI is supposed to accomplish is embedded in the input, e.g. as a question, instead of it being explicitly given. Prompt Engineering typically works by converting one or more tasks to a prompt-based dataset and training a language model with what has been called "Prompt-Based Learning" or just "Prompt Learning".

In practice , Prompt Engineering is the science and art of using libraries like Langchain which is an application framework for conversational agents along with document processing libraries to ingest documents, language models and embeddings to encode text on the semantic language of an LLM, vector databases and search engines to provide an agent with memory and prompt formatting modules to provide an LLM with the context to bring its knowledge to bear on a problem.

Why are LLMs and Chatbots important?

Chatbots and LLMs are *highly disruptive AI technologies*. While the importance of Chatbots like ChatGPT is evident to anyone that uses them (try ChatGPT now if you haven't) and considers possible use cases. The most obvious Chatbot use case is customer service, which has seen widespread adoption. Increasingly sophisticated Chatbot customer service representatives are rapidly replacing human jobs, resulting in significant savings. Customer service agents are just the beginning.

This trend will continue, driving mass unemployment across a swath of industries and job functions. If you don't adapt, your company's business model will be disrupted and left behind. Every business with a data science, analyst or engineering team needs to train its staff to take advantage of LLMs and solve task with Chatbots.

The consequences of not embracing these technologies can be severe. Chegg was behind the curve on LLMs and Chatbots and it has cost them dearly - students can now use ChatGPT for homework help rather than Chegg, and as a result, their stock fell 40% overnight. **Sending your technical staff to this course is a way to plant seeds of innovation that will spread across your organization at an accelerated rate**. Asking students to *apply their new skills to prototype a solution to a business problem and then presenting their results to the broader organization* is an excellent way to get maximum benefit from this course!

In this course we will introduce students to building conversational agents using Python, Graphics Processing Units (GPUs), natural language processing, representation learning, language models and text embeddings, document parsing and vector search, prompt templates and large language models combined through the Langchain Chatbot Framework to **conduct *Prompt Engineering* and *Prompt Learning* to solve Machine Learning problems via *conversational agents* driven by Large Language Models (LLMs).**

Students will leave the course with working code for their own Chatbot project using cutting-edge technology that solves real-world problems. This project and its code can serve as the foundation for additional work in students' employers' business domain to create custom Chatbots for their companies.

The course isn't limited to building your own Chatbots - Students will use LLMs to perform several different Machine-Learning tasks. Students will learn how to solve AI problems using LLMs without building Chatbots. Chatbots, LLMs and Prompt Learning are powerful tools, regardless of the form your application takes.

What you will learn

Students will use their basic Python data skills to learn to:

- Build their own Chatbots driven by LLMs
- Grasp the theory and fundamentals behind modern LLM technologies
- Build Chatbots using the popular Langchain framework
- Interact with remote LLM APIs like ChatGPT
- Use open-source LLMs locally for privacy
- Fine-tuning open-source LLMs like LLaMa for your business domain
- Entity Resolution (ER): de-duplicate records using Prompt Engineering
- Utilize vector database and search engines for semantic search
- Add memory to Chatbots to add context to chat sessions

Who should attend this course

This course is for anyone who knows Python!

- Data Scientists
- Data Analysts
- Software Engineers
- Data Engineers
- Machine Learning Engineers
- Data Science and Engineering Managers who want to extend their work using Large Language Models (LLMs) and conversational agents or Chatbots

Skills required to be successful in this course

- Students should know basic Python programming.
- **Basic experience** working with data for at least one of:

- Spreadsheets
- Reports
- Analysis
- SQL queries
- Problem-solving
- Machine Learning
- Data Processing

3. Basic mathematical concepts are used to explain how LLMs and Chatbots work, but **it is possible to build Chatbots without understanding the math behind Machine Learning.**

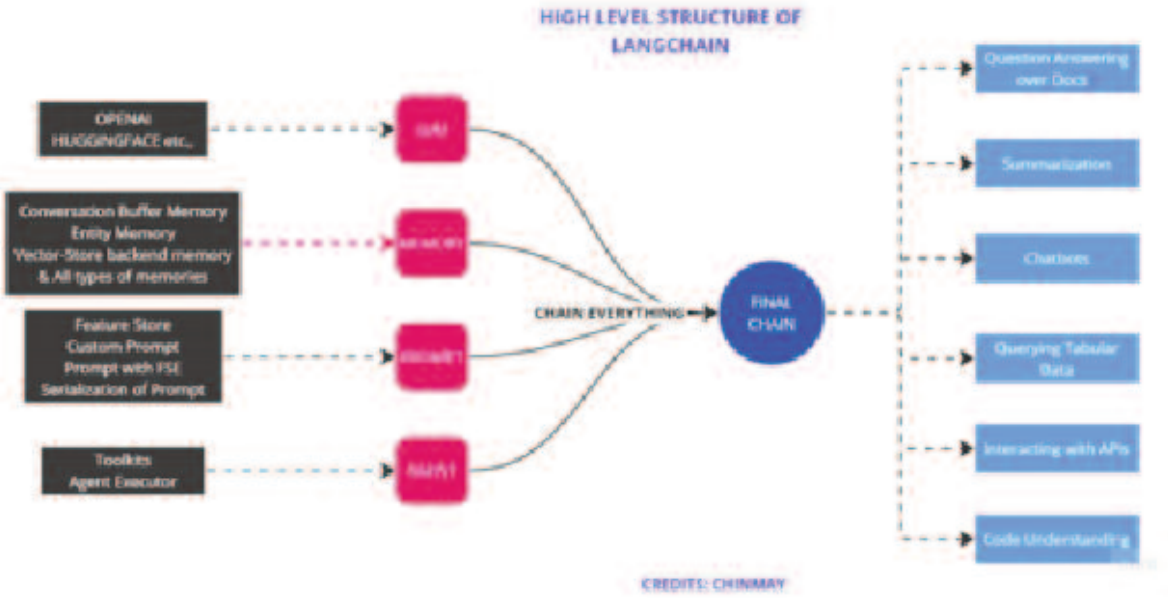
OUTLINE

The course outline is subject to change as technology rapidly develops but will roughly cover:

- Day 0 - **Preparation**: during the week before class begins, students will prepare their computers to run the software for the course. I will be available by email to help them debug any problems. This lets us spend classroom time learning. Setup is painless using Docker Desktop (download here) with a cu stom docker image for the course. This image contains web-based, interactive Python code using Jupyter notebooks that contain the interactive course content.

  At the start of Day 1 of the bootcamp, we will verify all software and course content works for all students and debug any issues that remain. Docker images should make this process seamless, as there is nothing to configure or change - the images contain a complete operating system and the course software.

- Afternoon 1 - **Theory and Practice**: The first day is heavy on theory to provide the background students need to understand Prompt Engineering concepts.



We will familiarize students with the following topics:

- Generative AI Strategy - how your company can adapt!
- Natural Language Processing (NLP)
- Natural Language Understanding (NLU)
- Seq2Seq Machine Learning Tasks
- Basic Linear Algebra for Understanding LLMs
- Text Embeddings
- The Inner Workings of Word2Vec and Doc2Vec
- Introduction to Transformer Models
- Sentence Transformers
- HuggingFace Hub
- Large Language Models (LLMs) - Training, Bias, Hallucinations!
- Vector Distance Metrics
- Vector Databases: FAISS, OpenSearch
- Introduction to Langchain: Concepts and APIs
- Streamlit Web Apps in 10 Minutes

- Afternoon 2 - **Building Conversational Agents with LLMs**: On the second afternoon we will put theory into practice to solve Machine Learning problems using LLMs and build our own Chatbot!

We will work through the following Python projects:

- Chatbot Power Users: ChatGPT as your Coding Partner
- Data Labeling using the Facebook LLama LLM
- LLMs: Entity Resolution with ChatGPT4
- Building our First Chatbot with Langchain
    - Basic Streamlit/Langchain Chatbot web apps
    - Vector search in OpenSearch
    - Ingesting Business Documents
    - Adding Memory to our Chatbot
    - Creating complex Chains

INFORMATION

PARTICIPATION FEE

€ 850

The fee includes all seminar documentation.

SEMINAR TIMETABLE

2.00 pm - 6.00 pm (Italian time)

HOW TO REGISTER

You must send the registration form with the receipt of the payment to:
info@technologytransfer.it

TECHNOLOGY TRANSFER S.r.l.
Piazza Cavour, 3 - 00193 Rome (Italy)

PAYMENT

Wire transfer to:
Technology Transfer S.r.l.
Banca: Cariparma
Agenzia 1 di Roma
IBAN Code:
IT 03 W 06230 03202 000057031348
BIC/SWIFT: CRPPIT2P546

GENERAL CONDITIONS

DISCOUNT

The participants who will register 30 days before the seminar are entitled to a 5% discount.

If a company registers 5 participants to the same seminar, it will pay only for 4.

Those who benefit of this discount are not entitled to other discounts for the same seminar.

CANCELLATION POLICY

A full refund is given for any cancellation received more than 15 days before the seminar starts. Cancellations less than 15 days prior the event are liable for 50% of the fee. Cancellations less than one week prior to the event date will be liable for the full fee.

CANCELLATION LIABILITY

In the case of cancellation of an event for any reason, Technology Transfer's liability is limited to the return of the registration fee only.

IVAN REZNIKOV
CHATBOT AND LLM BOOTCAMP

May 15-16, 2024

Registration fee:
€ 850

If registered participants are unable to attend, or in case of cancellation of the seminar, the general conditions mentioned before are applicable.

first name .....

surname .....

job title .....

organisation .....

address .....

postcode .....

city .....

country .....

telephone .....

fax .....

e-mail .....

Stamp and Signature

Send your registration form with the receipt of the payment to:
Technology Transfer S.r.l.
Piazza Cavour, 3 - 00193 Rome (Italy)
Tel. +39-06-6832227 - Fax +39-06-6871102
info@technologytransfer.it
www.technologytransfer.it

Speaker

Ivan Reznikov is principal Data Scientist at QBurst, brings over a decade (12 years) of proficiency in Python and Data Science. He is currently dedicating his expertise to authoring LangChain for Life Sciences, a book set to be published by O'Reilly. His accolades include being recognized as a Kaggle Competition Expert and teaching university in Master Data Science courses. Ivan has spoken on TEDx, GITEX and PyCON MEA.