AI - Opportunities and Risks in the Context of Customs & Trade Compliance

12th September 2024

MIC - Overview

Leader in Global Customs & Trade Compliance Software

Founded: **1988 in Linz (HQ)** / Austria

550+ employees worldwide – offices in US, Mexico, Thailand, Germany and Switzerland

Family-owned and financial independent global software vendor with 1 single IT platform globally

MIC Customs Solutions

Solutions cover **55+ countries** on 6 continents including **direct custom filing, customs tariff classification, export control management** and **origin calculation**

Software-as-a-Service (SaaS) Deployment and used by **800+ multinational customers**

Customers - excerpt

Automotive

Technology

Machinery

Chemical

Pharma

Textile

Food & Beverages

Retail & Consumer

Logistics

Tobacco

Others

 |

MIC

550+ Employees

1988 founded

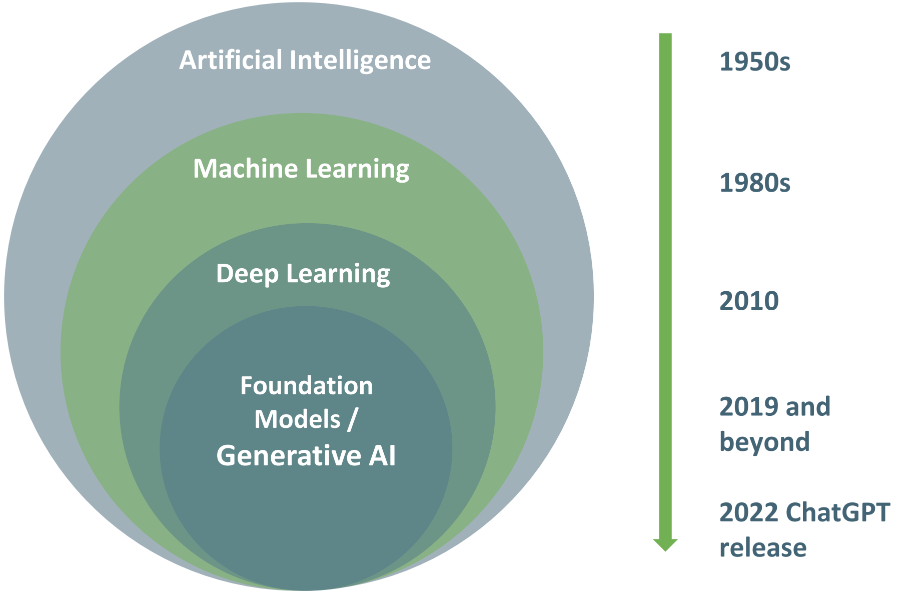
Artificial Intelligence, Machine Learning and Deep Learning

Artificial Intelligence (AI)
Machines mimic human intelligence and behaviors

Machine Learning (ML)
Rooted in statistical learning and methods
Not explicitly programmed, but in general **learns from data**

Deep Learning (DL)
Utilizing and learning from large amount of data
Uses large deep neural networks with billions of parameters
Is computational challenging

Founded dedicated team in 2020 to **drive innovation in AI within MIC**



Generative AI allow for new type of applications

 |

Artificial Intelligence, Machine Learning and Deep Learning

According to the EU Artificial Intelligence Act an ‘AI system’ is defined as

*‘AI system’ means a machine-based system that is designed to operate with **varying levels of autonomy** and that **may exhibit adaptiveness after deployment**, and that, for **explicit or implicit objectives, infers, from the input** it receives, **how to generate outputs such as predictions, content, recommendations, or decisions** that can influence physical or virtual environments;* [1]

[1] https://artificialintelligenceact.eu/article/3/

Deep Learning – Approaches

Discriminative Models

Predict a label or output based on input features **– models have a concrete goal** (e.g. tariff classification)

Trained on historical labelled data (e.g. existing product classifications)

Computer Vision (CV)

Image Classification - Assign one class to the image (e.g. tariff classification)

Natural Language Processing (NLP)

Text Classification - Classify arbitrary strings and assign to classes

Or using a **multimodal** approach (combining CV and NLP)

features → label

62024010

61091000

“T-Shirt print 2-1-D-06
Textil Damen Jersey
Short Sleeves 100%
Baumwolle, gewirkt”

61102099

Discriminative **multimodal models** are already **in production within MIC Cloud**

Deep Learning – Approaches

Generative AI

Generate new data samples from given context (e.g. text, image, audio or multimodal)

Generative AI Models are **Foundation Models**

Allow for AI applications without training on historical data

Self-Supervised Learning (“dark matter of intelligence”[1])

Generative models are **first trained to learn from huge amounts of unlabeled data**. By predicting parts of the input data, models learn useful representations

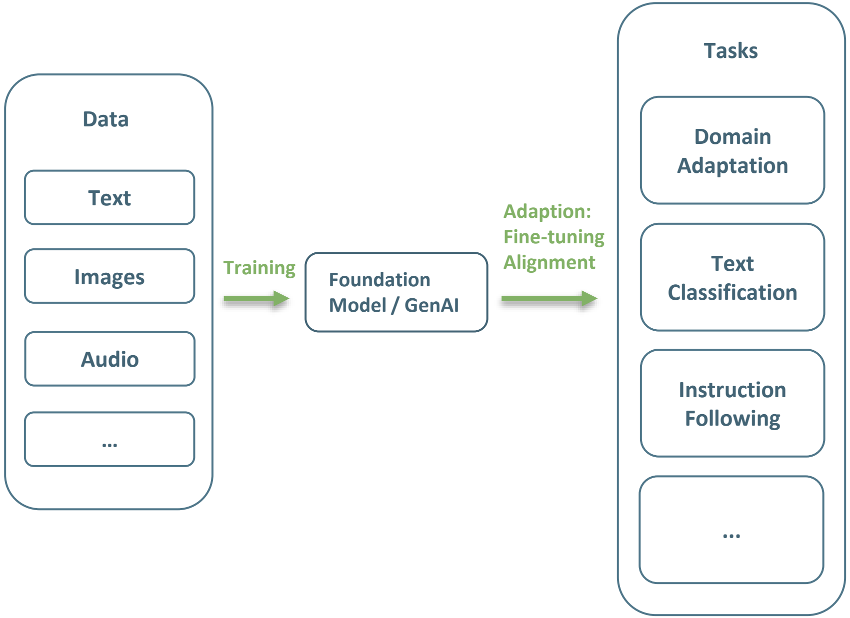
Adaptation - Fine-tuning and Alignment

Foundation models can be **fine-tuned for specific tasks** with smaller labeled datasets or **specialist on more narrow content**

Much **more cost-effective** and effective **than training from scratch**

Multimodal

Most **recent released foundation models are multi modal** – meaning the are trained on various modes of data (e.g. images, text, audio, etc.)



Generative AI models can also be **used without further training**

[1] https://ai.meta.com/blog/self-supervised-learning-the-dark-matter-of-intelligence/

Deep Learning – Foundation Models

Large Language Models (LLM) – text-based foundation models

Text Processing

Text is split into tokens, which are smaller units of meaning

Models assign unique ids to tokens, which are then used as inputs

original text ***- “My name is Tom.”***

tokens ***- [“My”,“ name”,“ is”,“ Tom”,”.”]***

token ids ***- [3666,1438,318,4186,13]***

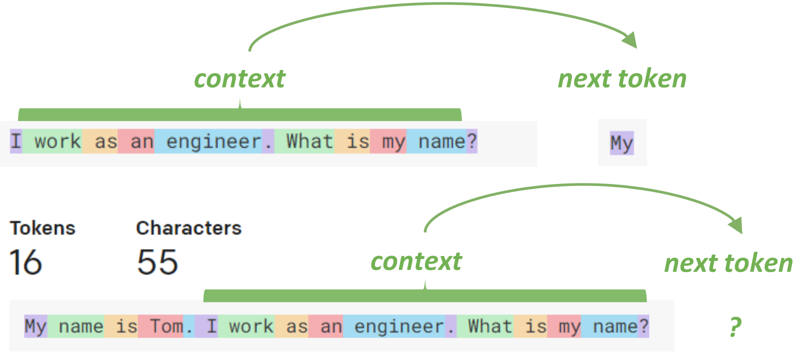
Generative Language Modelling

LLMs process a set of given tokens (context) and sample the next most likely token. This is an iterative process.

Context Length and Transformers

Context length refers to the maximum number of tokens a model can process at once

*Large text language have/**had** a limited context length – for modern foundation model this is becoming less of an issue (e.g. Anthropic Claude Sonnet 3.5 have a context limit of ~200K tokens equal to ~300 pages of text)*



Given a transformer with context length of 12 this question (what is my name) could not be answered

Deep Learning – Foundation Models

Scale

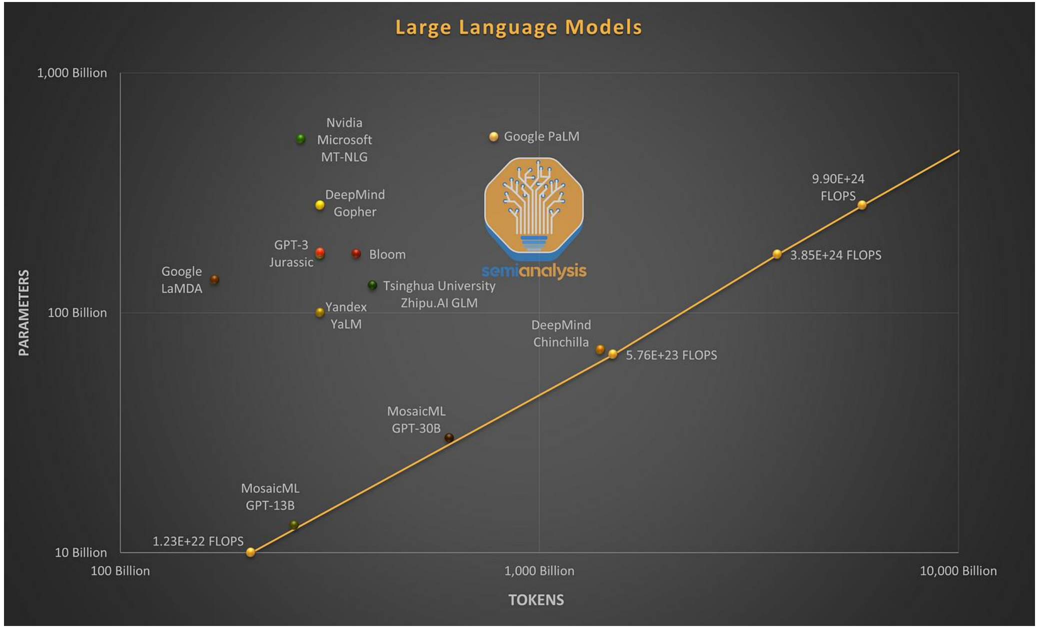
Number of parameters (size) of a model and/or number of tokens a model have seen (training data size) scale an LLM

Compute resources limitations - *e.g., recent Facebook LLM called "LLaMA 3.1" used 16000 H100 NVIDIA GPUs for multiple weeks.*

Data

Training from scratch requires a huge amount of data, measured in billion of tokens – example for GPT-3 (original – in 2020)

- *Common Crawl – 410B tokens (~3.3TB disk size)*
- *WebText2 – 19B tokens*
- *Books1 – 12B tokens*
- *Books2 – 55B tokens*
- *Wikipedia – 3B tokens*



[1]

A recent model, LLaMA 3.1 from Meta used 15000B tokens during training

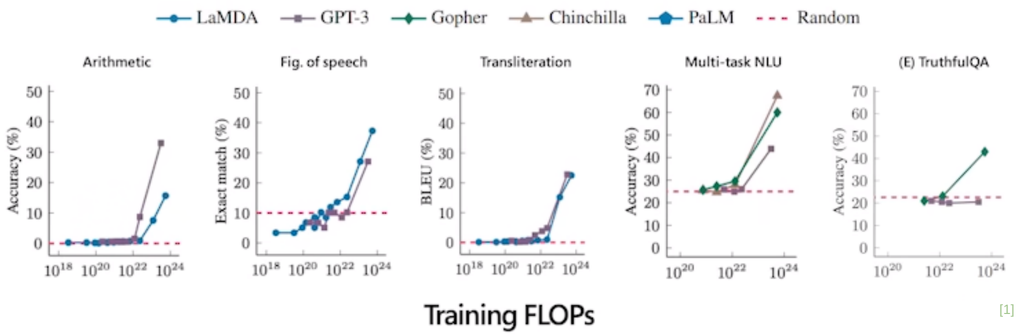
 |

[1] https://www.semianalysis.com/p/the-ai-brick-wall-a-practical-limit

Deep Learning – Foundation Models

Capabilities

Scaling leads to emerging capabilities like arithmetic, reasoning, transliteration, etc.



[1]

In Context Learning

Allows for in context learning/reasoning without the need to fine tune a model

Zero Shot Learning

Ability to classify new examples from previously unseen classes.

```
Given x=3 solve x^2 - 9 = y, only give the final answer
y = 0
```

```
Classify "Tuesday" according to the following classes: year, month, day
Class: day.
```

Few Shot Learning

Given samples as context learn to classify and solve follow up tasks

```
Classify "Tuesday" given Monday=7,Wednesday=5,Thursday=4,Friday=3
Tuesday=6
```

 |

[1] https://arxiv.org/pdf/2206.07682.pdf

Deep Learning – Foundation Models

Challenges

Data availability and quality – finding high-quality data at scale is becoming a problem – solution could be syntactic data

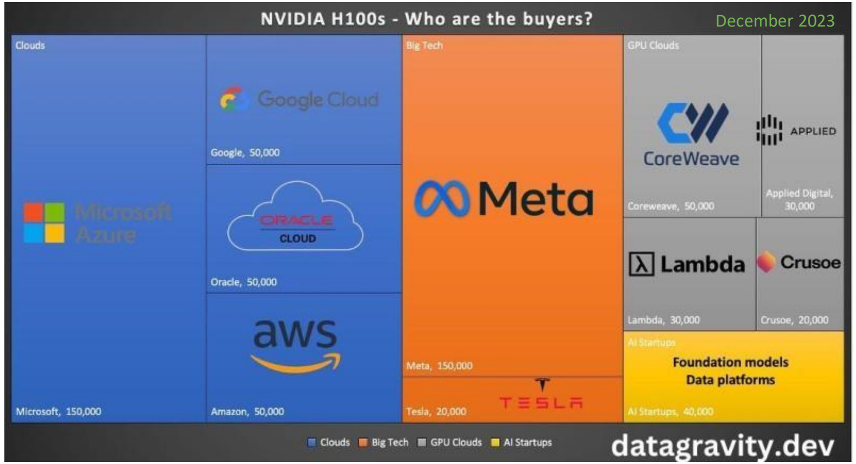
Computational requirements – training requires enormous resources, especially still NVIDIA GPUs (e.g. H100)

Inference speed and latency – typically measure in time to first token (TTFT), total response time (TFT) and tokens per seconds (TPS) (e.g. gpt4o can produce ~100 TPS/~3-4 sentences per second)

Alternatives architectures like xLSTM (by nxAI and Sepp Hochreiter) offer higher inference speed / latency (beside other advantages)

Available Data [1]

- Web Data – 60-160T
- Code – 1T public, 20T private
- Books – 30T
- Academia – 1T
- Private Data – Facebook 140T, Gmail 400T



[1] https://www.educatingsilicon.com/2024/05/09/how-much-llm-training-data-is-there-in-the-limit/

Deep Learning – Foundation Models

Generative AI - Model Landscape

Model	By	Date	Size	Context Length	Open Source	Multimodal
GPT-4o	US, OpenAI	May 2024	NA	~128K	No	Yes
Claude Sonnet 3.5	US, Anthropic	Jun 2024	NA	~200K	No	Yes
Google Gemini 1.5 Pro	US, Google	May-24	NA	~2M	No	Yes
LLaMA 3.1	US, Meta	Jul-24	8B, 70B, 405B	128K	Yes	No
Mixtral 8x7B	FR, Mistral AI	Dec-23	45B	32K	Yes	No
xLSTM?	*AT, nxAI/Sepp Hochreiter*	*May-24*		*INF*		

No major EU based player (yet) – possible solutions is to support Open-Source models and ecosystems / novel approaches

Deep Learning – Foundation Models

Applications

Foundation models/Large Language Models (LLMs)/Generative AI allow for a **new type of application based on their capability to reason and understand the context in natural language**

Agents

Natural language as interface

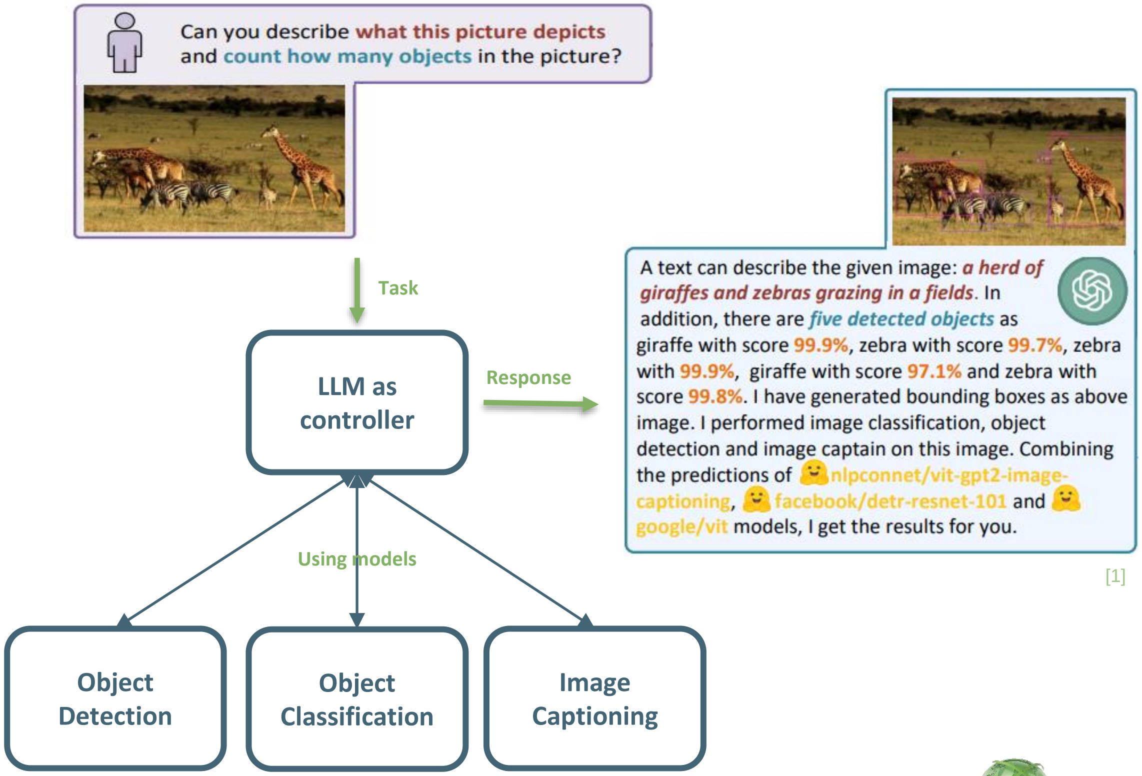
LLM as controller

Calling other models (specialist) using their response in a chain of action to solve a task

Chat Bots

Alignment (e.g., from human feedback) **of LLMs to achieve the desired behavior** (e.g., ChatGPT)

Personal assistant, chatting with the user and interacting with the environment (e.g., searching and summarizing the web, calling APIs, etc.)



[1]

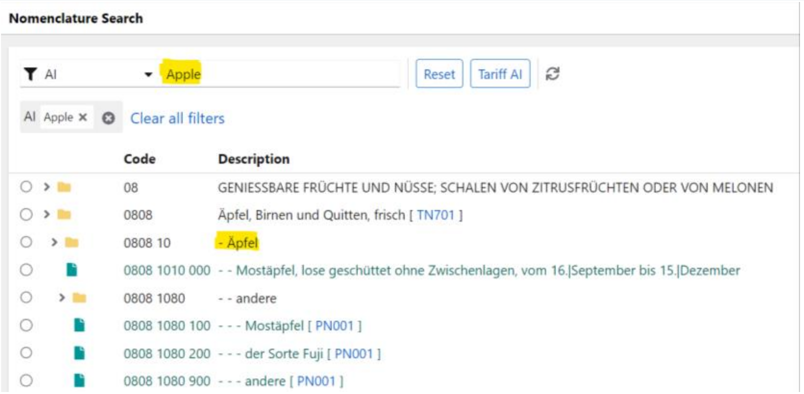
[1] https://arxiv.org/pdf/2303.17580.pdf

Deep Learning - Generative AI Concepts

Embeddings

Allow for a semantic search, finding semantically similar or semantically related items - even multilingual (e.g. searching for “Apple” will find the concepts/items “Früchte”, “Apfel”, etc.)

Used for document retrieval for a given task or question or **finding similar products based a given product image**

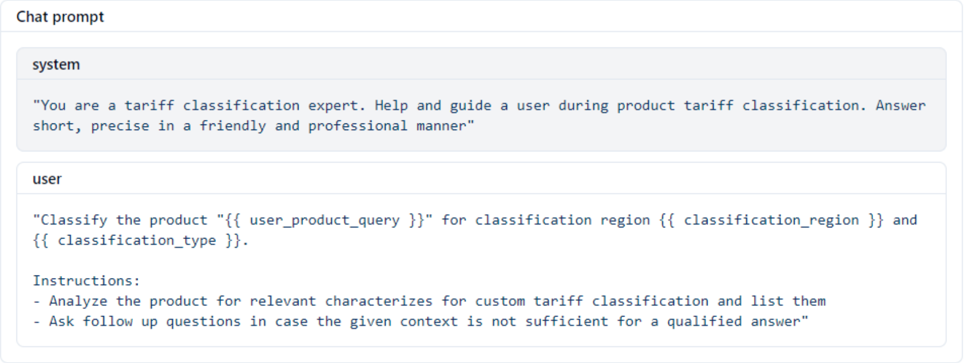


Prompt

A specific instruction to the Generative AI model to generate the desired output.

System message typically define the tone and general context of the conversation

User message defines the instruction – are dynamically generated and only certain parts of the instruction can be defined by an end-user.



Prompt engineering - the practice of designing and refining prompts to achieve accurate and relevant outputs.

Deep Learning - Generative AI Concepts

Tools

Describes a **strictly defined tool or function** which can be **used by a Generative AI** model, deciding when and how to use

Examples - **embedding search**, database lookup, deterministic functions (e.g. calculator), other agents and many more (no limit to ideas and functionality)

Chains

Fixed list of prompts and tools which are **executed sequentially** by a Generative AI model.

Agents

Agents are using Generative AI models, available tools, other chains and agents, defined prompts to perform a tasks
Autonomously decide what tool to use and when

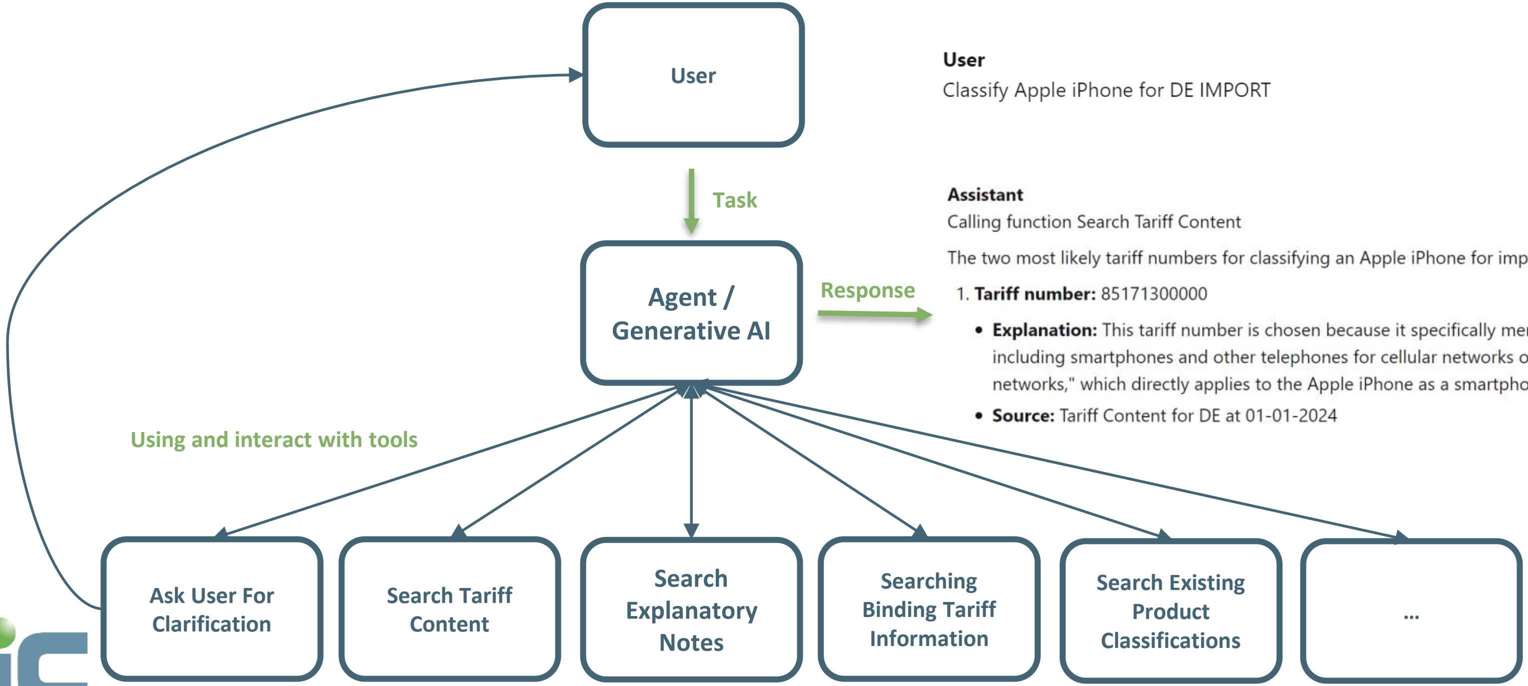
```
tools :
• search_tariff_content
      • Searches the tariff content for a given query
        using embeddings
      • parameters:
            • query
            • region – (e.g. US, EU, CN)
• search_bti
      • Searches the binding tariff information which
        are valid for the EU using embeddings
      • parameters
            • query

agents:
• tariff_classification_agent
      • Classify the product given the {{
        user_product_query }} for region {{ region }}.
        Autonomously decide which tool to use and how.
      • available_tools:
            • search_tariff_content
            • search_bti
            • …
```


Deep Learning - Generative AI Concepts

Applications and Agents

Generative AI models allow for a **new type of application** based on their **capability to reason and understand the context in natural language – do not require training on historical data**



User
Classify Apple iPhone for DE IMPORT

Assistant
Calling function Search Tariff Content

The two most likely tariff numbers for classifying an Apple iPhone for import into Germany are:

1. **Tariff number:** 85171300000
   - **Explanation:** This tariff number is chosen because it specifically mentions "Telephone sets, including smartphones and other telephones for cellular networks or for other wireless networks," which directly applies to the Apple iPhone as a smartphone.
   - **Source:** Tariff Content for DE at 01-01-2024

 |

Deep Learning - Discriminative AI Concepts

Dataset

Usually describes a set of labeled features (e.g historical data)

Labels

this is the **objective** or target what we want **to predict/learn**.

Features

can be multimodal - textual, numerical, categorical, but also signal data like audio, video or images are possible (or a combination of those)

Features		Labels
Image	Textual Description	HTS
	car mobile charger	85044090
	gear box	87084000

Deep Learning - Discriminative Model/AI Concepts

Model

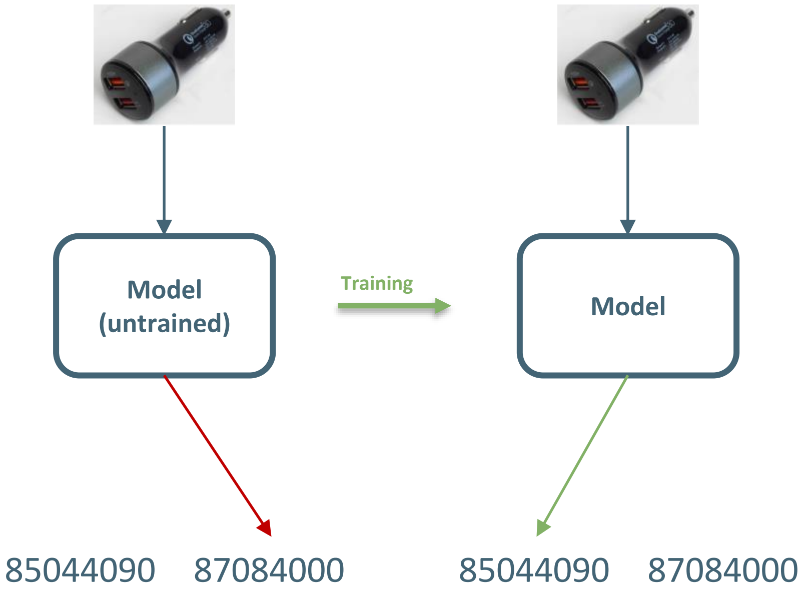
Accepts **features as inputs** and **returns a label**

We can return a label and a score (e.g., probability) but reasoning about the result (e.g., why this label was returned) is difficult

Training

A Model is **trained** using a dataset **to fulfill defined objective** (e.g., predict a given set of labels) – referred to as **learning**

(an untrained model will return most likely a random label - a trained model will return more likely the correct label)



 |

Deep Learning - Discriminative Model/AI Concepts

Model Evaluation

After training provide a concreate **statement about the performance of the model** (e.g., simple accuracy or other statistical measurements/metrics)

Process

Complete dataset is **split** into a **training and test** part

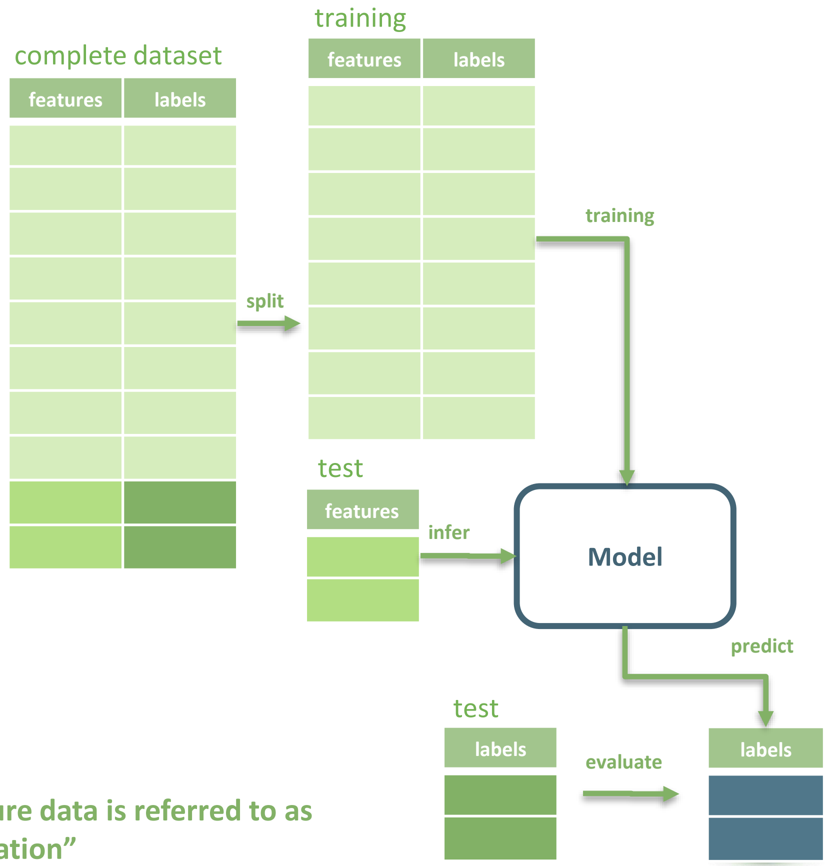
Model is **trained on the training** (features + labels) and **evaluated on the test** (only features) dataset.

Finally, model predicts labels on the test dataset and is evaluated with the true labels

Metrics

Based on the **true labels and predicted labels using the test dataset** (e.g., simple accuracy or other statistical measurements)

Gives an outlook of the performance on unseen/future data



Performance on unseen/future data is referred to as "Generalization"

Deep Learning - Discriminative Model/AI Concepts

Metrics and Model Evaluation

Single metrics (e.g., overall accuracy) can be **misleading**.

Multiple metrics should be considered

Beside macro or average metrics (e.g., overall accuracy) **individual per class metrics** should be analyzed

Imbalanced Support

Imbalanced number of samples can lead to **unreliable metrics** for classes with fewer sample

Costs of Error

Which type of error are more important (FP or FN)

Important to consider per class (some classes are more relevant and important than others)

true labels

predicted labels

	T-Shirt	Shoe	Pants
T-Shirt	71	1	4
Shoe	8	1	1
Pants	1	3	10
support	80	5	15
class-accuracy	89%	20%	33%
accuracy	82/100 = **82%**		
balanced-accuracy	(89 + 20 + 33)/3 = **47%**		

Typically, a single metrics is not sufficient to evaluate a model

Deep Learning - Discriminative Model/AI Concepts

Top K

Consider the **top k prediction** of a model during inference and evaluation

The model assigns all classes a score. The certainty of the suggestion is directly connected to this score.

For example, the **top 5 accuracy** describes that the correct class is in the first 5 suggestions of the model.

	Score %	Tariff Number	Tariff Description
1	90.90	61091000	T-shirts, singlets and other vests of cotton, knitted or crocheted
2	2.74	61099020	T-shirts, singlets and other vests of wool or fine animal hair or man-made fibres, knitted or crocheted
3	0.23	61051000	Men's or boys' shirts of cotton, knitted or crocheted (excl. nightshirts, T-shirts, singlets and other vests)
4	0.09	61102099	Women's or girls' jerseys, pullovers, cardigans, waistcoats and similar articles, of cotton, knitted or crocheted (e
5	0.05	61102091	Men's or boys' jerseys, pullovers, cardigans, waistcoats and similar articles, of cotton, knitted or crocheted (exc

Deep Learning – Use Cases GTM

Tariff Classification Approaches

Discriminate Models

Using NLP – from part description, engineering codes and other meta information to classification

Using CV – from images, drawings and similar to classification

Or a **combination** of multimodal data types

Generative AI

Using Embeddings – find semantic similar tariff numbers based on the tariff content itself (tariff description, BTI) – this **allows for AI tariff suggestions without explicit model training** on data (e.g. customer data)

Using Chat Models - Use foundation models and tariff content to **ask question about tariff content using natural language.** (not limited to tariff content of course)

Or a **combination approach** using discriminate models, embeddings and chat models to guide through the classification process

Deep Learning – Use Cases GTM

Sample Use Cases in GTM

Customs declaration co-pilot

Provide **suggestions** and **guide** for various fields while filing a customs declaration

Using NLP, and other techniques such as auto regression to **learn from previous declarations**

Document Information Retrieval

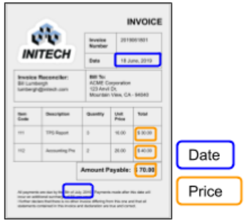
From in paper documents to customs declaration

Using **multimodal foundation models**/Generative AI which understands/interpret images/documents

Compliance and auditing

Continuously **learn** from approved data such as successful declarations

Apply learned knowledge to **verify and audit data** – are values plausible, are there outliers, would a trained deep learning model come to the same conclusion



customs declaration

Deep Learning – Use Cases GTM

Sample Use Cases in GTM

Foundation Model Custom

Using existing models and **fine-tune them on various custom data** (tariff content, FTAs, rulings, custom authority legal documents, etc.). Can be multimodal (text and images)

Use as **base model for various downstream task like** (tariff classification, LLM controller for agents, custom related chat bots, etc.)

Custom Co-Pilots

Use foundation models to **ask question about content, business data using natural language.**

AI - Demos

Live Demos by MIC

AI Classifier - Discriminative Models

Trained using **descriptions and images** as well as a combined approach to provide tariff classifications suggestions

Demo Machine Learning Model facts

Trained for a retail customer on a large amount of data over **600000** articles with over **3.1 million product images**.

Methods use state of the art techniques for NLP and CV as well as a **unique combined approach**.

Generative AI - Demos

Live Demos by MIC

Embedding Search

Semantic retrieval of relevant tariff content given product description

Tariff Classification Agent

Basic **tariff classification agent**, using various tools and sources (e.g. combined nomenclature, explanatory notes, binding tariff information, etc.)

Prompt and Prompt Engineering

Show a tariff classification agent – dynamically **adapt the instruction and prompt for an agent**

Multimodal Image based Agents

Show **multimodal tariff classification** agents understanding product images or plain invoices

Chatting with Data

Simple showcase showing **chatting with data**, asking questions using natural language

Questions?

European Offices

Headquarters
MIC Datenverarbeitung GmbH
4020 Linz, Hafenstraße 24
Austria
Tel.: +43 (0)732-778496
Fax: +43 (0)732-778496-600
sales@mic-cust.com
www.mic-cust.com

MIC Datenverarbeitung GmbH
1100 Vienna, Gertrude-Fröhlich-Sandner Straße 2, Austria
Icon Tower

MIC Datenverarbeitung GmbH
4232 Hagenberg, Softwarepark 35, Austria
IT-Center

MIC Logistik AG
8200 Schaffhausen, Neustadt 23
Switzerland
Tel.: +41 (0)52-630 22 00
Fax: +41 (0)52-630 22 01
info@mic-logistik.ch
www.mic-logistik.ch

DHF Logistik Software- und Datenverarbeitungs GmbH
5020 Salzburg, Jakob-Haringer-Str. 1
Austria
Tel.: +43 (0)662-887326-0
Fax: +43 (0)662-887326-820
office@dhf.at
www.dhf.at

Contact Office for Benelux & France
2030 Antwerp, Noorderlaan 147, bus 19
Belgium
Tel.: +32 (0)3-3049531
sales-be@mic-cust.com

MIC Logistik GmbH
91052 Erlangen
Allee am Röthelheimpark 45
Germany
Tel.: +49 (0)9131-91619-155
office-de@mic-cust.com

LDV-Systema Software GmbH
5023 Salzburg, Linzer Bundesstr. 61a
Austria
Tel.: +43 662-664640-0
Fax: +43 662-664640-16
office@ldv.at
www.ldv.at

Americas Offices

East Coast
MIC Logistics Corporation
29777 Telegraph Road, Suite 1130
Southfield, MI 48034, USA
Tel.: +1 (248)304-4460
info@mic-cust.com
www.mic-cust.com

West Coast
MIC Logistics Corporation
915 Broadway St., Suite 150
Vancouver, WA 98660, USA
Tel.: +1 (248)304-4460
info@mic-cust.com
www.mic-cust.com

MIC Customs Solutions México, S.A. de C.V.
Paseo de la Reforma 296, Piso 38
Col. Juárez, Delegación Cuauhtémoc
C.P. 06600, Ciudad de México
Mexico
Tel.: +52 1 55 7100 1266
office-mx@mic-cust.com
www.mic-cust.com

Asian Offices

MIC Customs Solutions TH Ltd.
Level 32 Interchange 21 Building,
399 Sukhumvit Road (Soi 21),
North Klongtoey, Wattana,
Bangkok 10110
Thailand
Phone: +66 2 660 6252
office-th@mic-cust.com

mic

www.mic-cust.com