LECTURE 1: INTRODUCTION

CSC5991: Introduction to LLMs

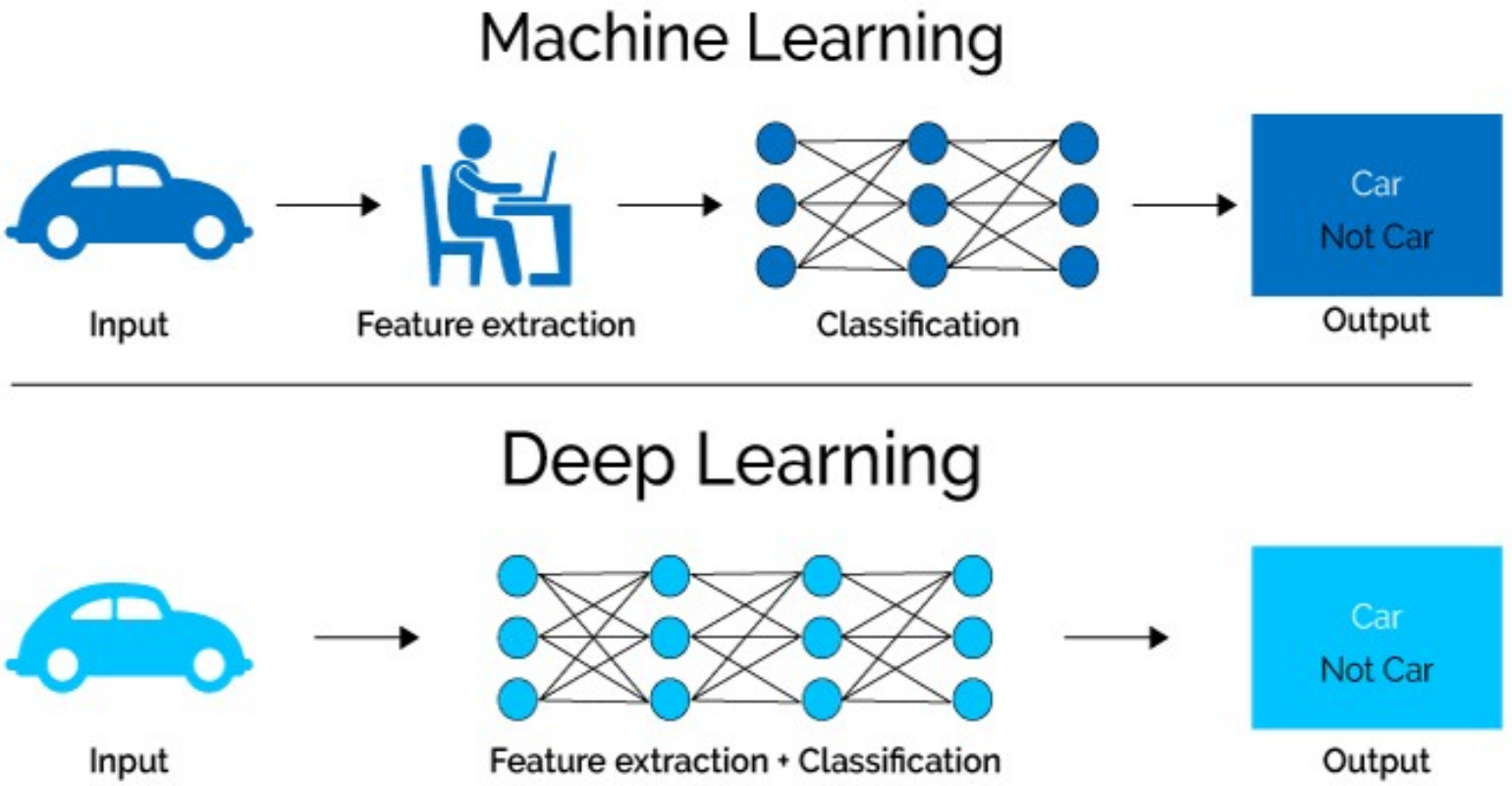
WELCOME TO THE COURSE

- Background.
- Overview of the course objectives and structure.
- Importance of LLMs in modern AI.
- Brief introduction to the topics we will cover.
- Q&A.

BACKGROUND

Machine Learning

Deep Learning



Taken from CS886 at UWaterloo

HISTORY OF DEEP LEARNING

- **Traditional ML/DL (Before 2018)**
  - Design specialized model architectures.
  - Leveraging task-specific features.
  - Train the specialized models with limited data.
- **Transfer DL (Between 2018 - 2021)**
  - Train a model with large amount of training data.
  - Use the features of the trained model to initialize part of the architecture.
  - Design specialized modules on top of the trained features.
  - Train the partially specialized model with limited data.
- **LLMs/Foundation Models (After 2021)**
  - Train a single huge model on astronomical amount of data.
  - Prompt the single model for everything.

PROS AND CONS OF TRADITIONAL DL

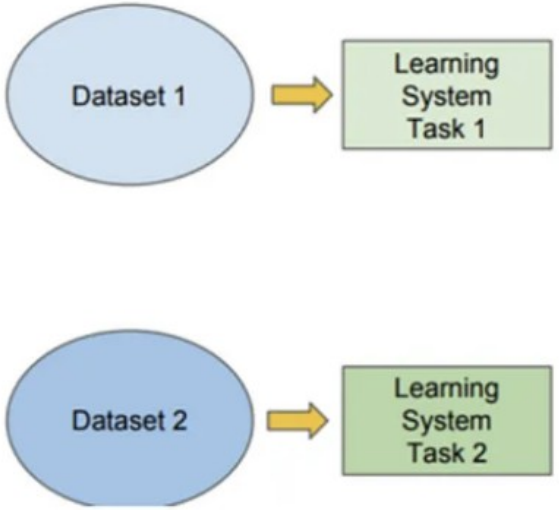
- **Pros**
  - The model considers the inductive bias for architecture design.
  - The model can be effectively trained with limited amount of data.
  - The model is normally small in size, easy to deploy for applications.

- **Cons**
  - Each task requires lots of expertise for architecture design.
  - Each task requires annotating specialized dataset.
  - The model cannot benefit from other annotated data, it needs to start from scratch literally to gain its skill.
  - Hosting many specialized models incur high costs.

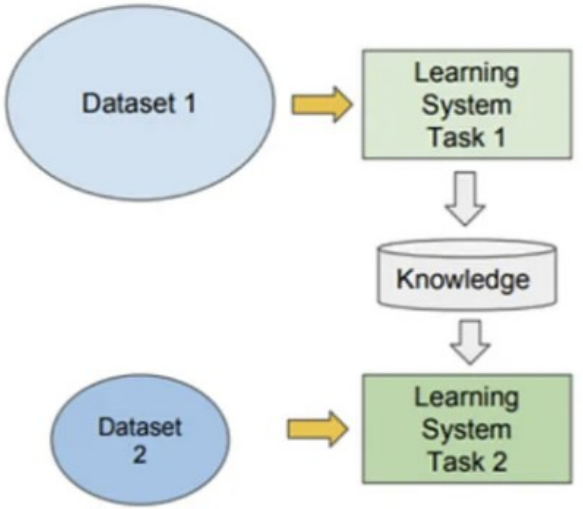
TRANSFER LEARNING

Traditional ML vs Transfer Learning

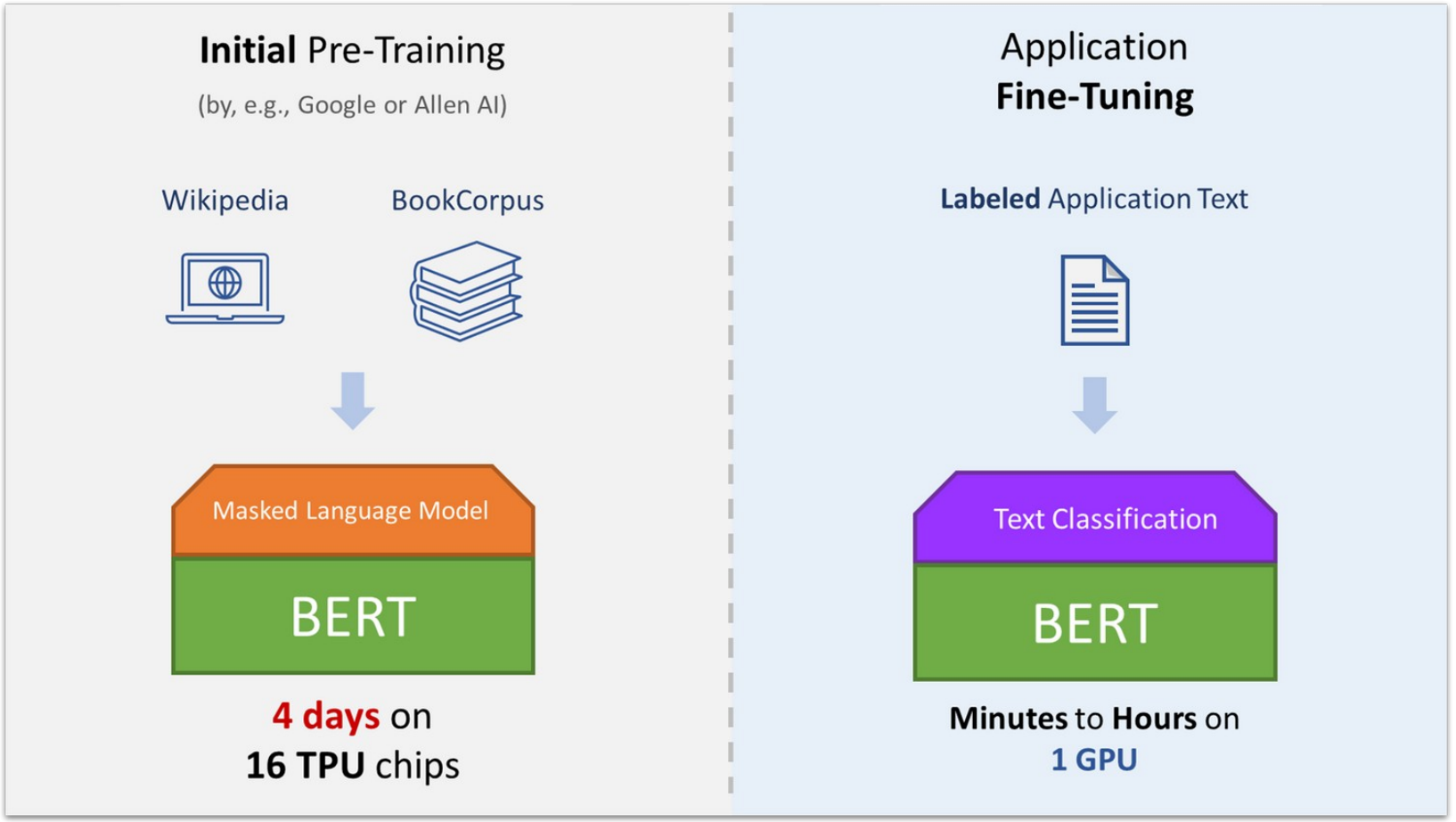
- Isolated, Single task learning.
- Knowledge is not retained or accumulated. Learning is performed w.o. consideration for knowledge learned from other tasks.



- Learning new tasks relies on previously learned tasks.
- Learning process can be faster, more accurate and/or need less training data.

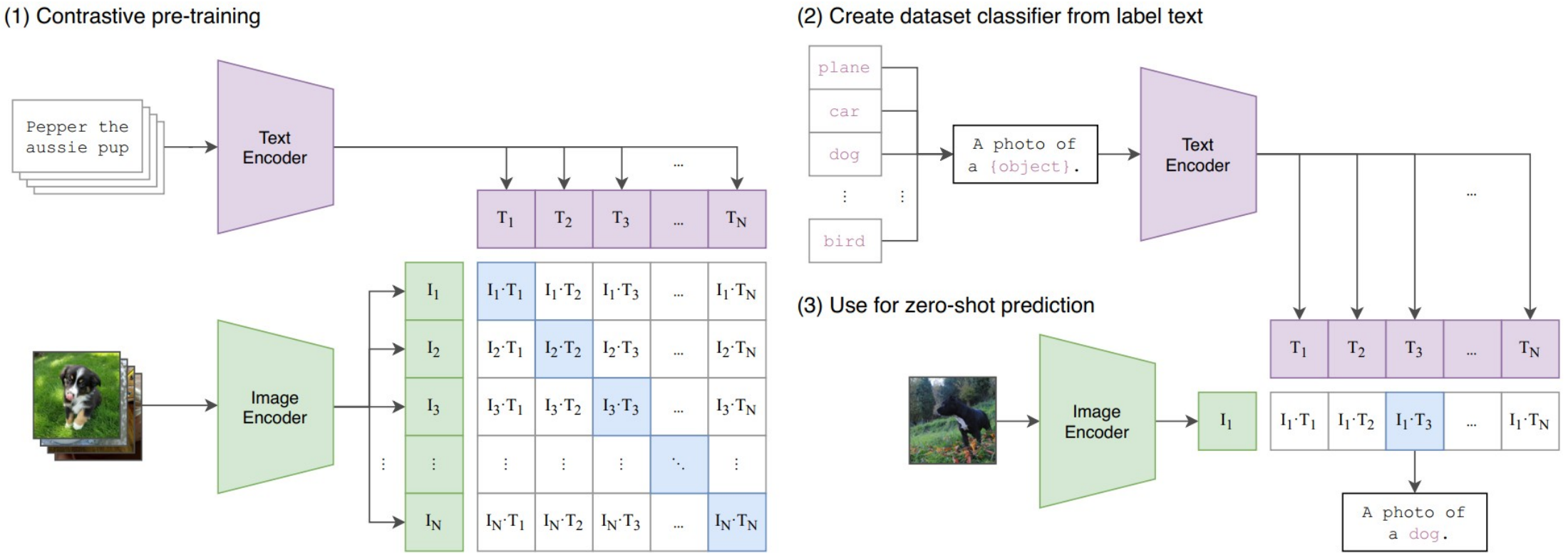


TRANSFER LEARNING IN BERT



Taken from CS886 at UWaterloo

TRANSFER LEARNING IN VISION



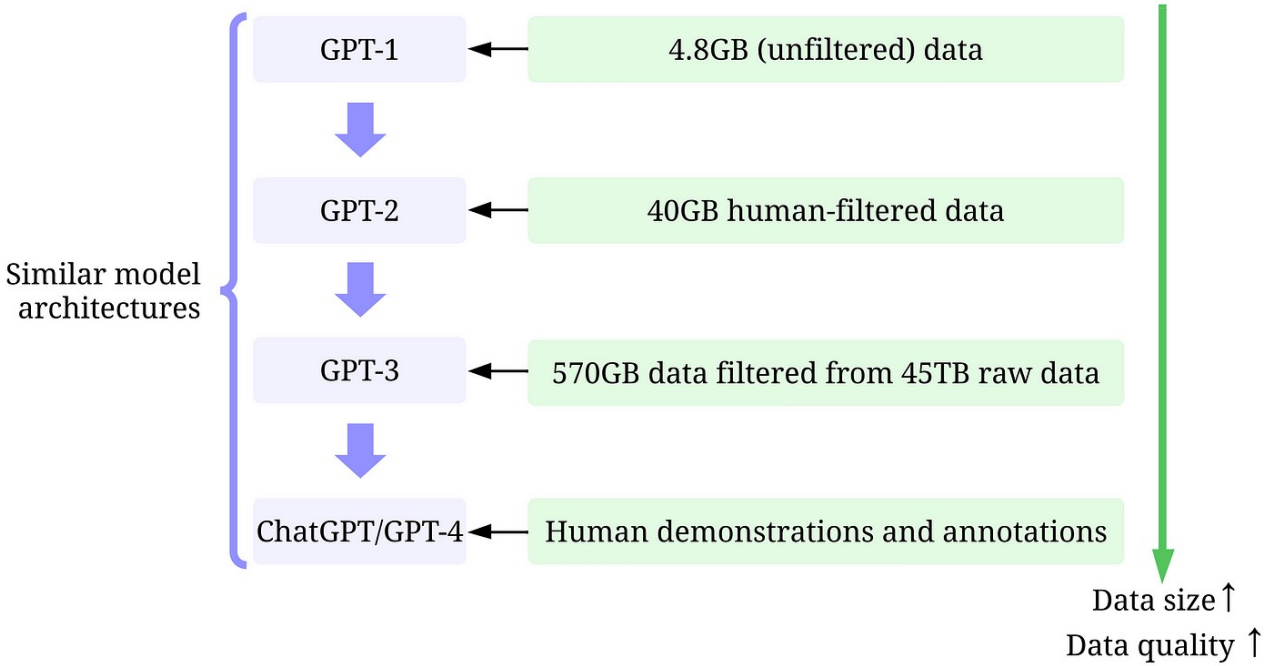
PROS AND CONS OF TRANSFER DL

- **Pros:**
  - The model shows much stronger capability than Traditional DL.
  - The model can generalize to unseen cases.
  - The model requires very few fine-tuning.

- **Cons:**
  - The model's performance is still not perfect.
  - There is still fine-tuning needed for the downstream tasks.

LARGE LANGUAGE MODELS

- **Task versatility**: LLMs handle multiple tasks (e.g., summarization, translation) without retraining.
- **Cost efficiency**: No need for task-specific labeled data or retraining, saving time and resources.
- **Generalization**: Broad knowledge across domains.
- **Faster deployment**: Prompt-based interaction.

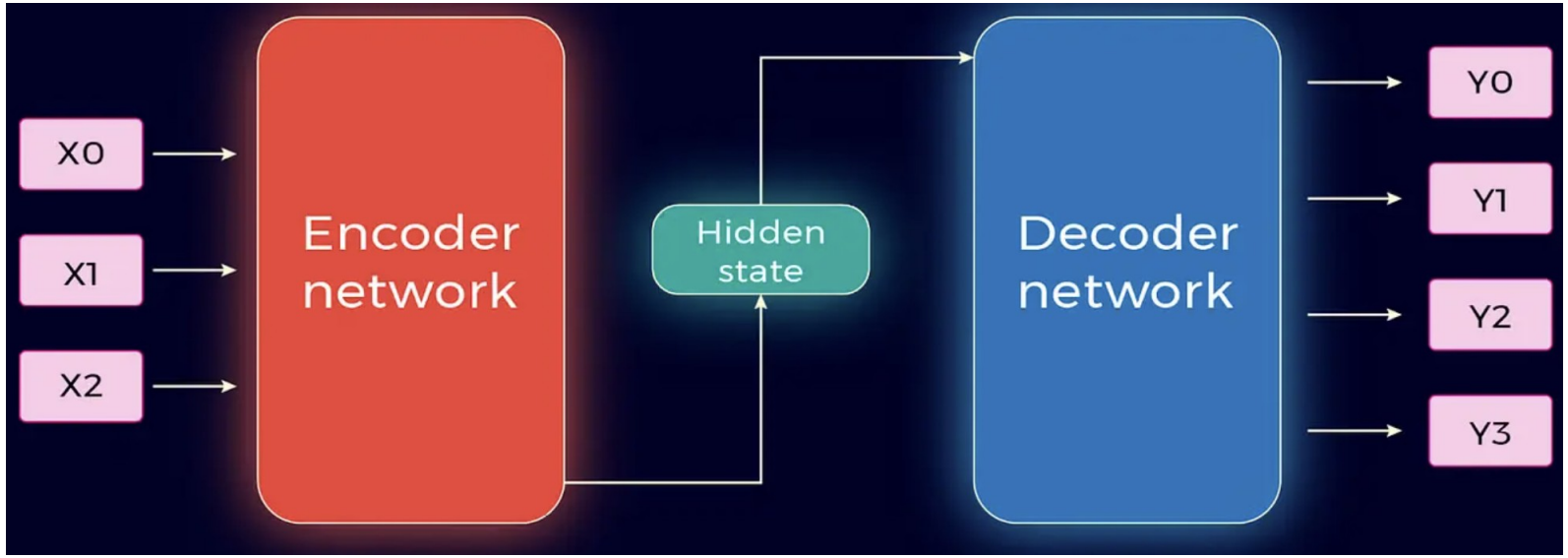


PARAMETERS OF LLMS

Parameters	Specialized DL	Transfer DL	LLMs
Model Size	< 100M parameters	100M -> 1B parameters	7B -> 1T parameters
Data Size	10K -> 1M tokens	100M -> 10B tokens	100B -> 30T tokens
Architecture	Specialized	General	General
Generalization	None	Reasonable	Strong

UNDERSTANDING LLMS

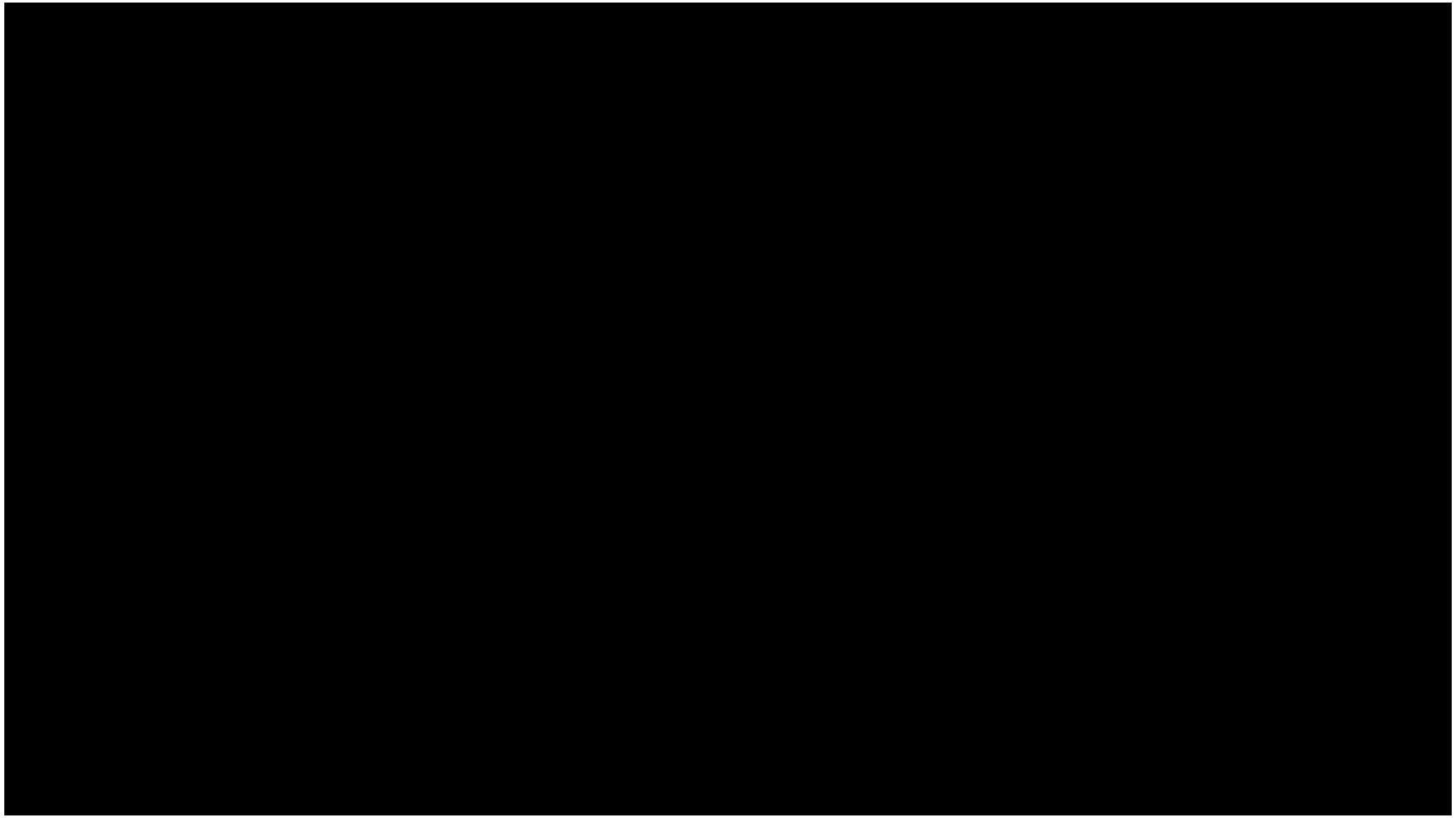
- **Definition:** Large Language Models as **advanced AI systems** trained on **massive text corpora**.
- **Key characteristics:** size, generalization, and adaptability.
- Examples of popular LLMs: GPT, BERT, etc.



IMPORTANCE OF LLMS

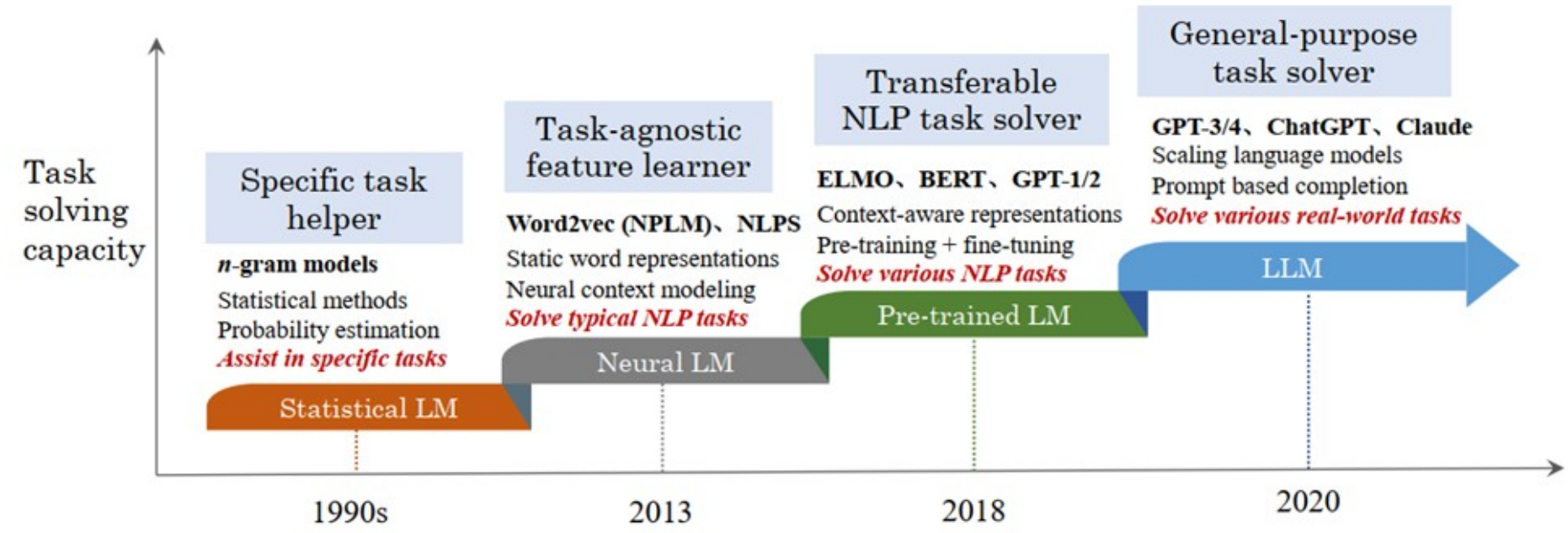
- Revolutionizing **natural language processing** tasks.
- **Applications:** chatbots, translation, summarization, content generation.
- Bridging gaps in automation and human-computer interaction.

WHAT CAN GEMINI DO?



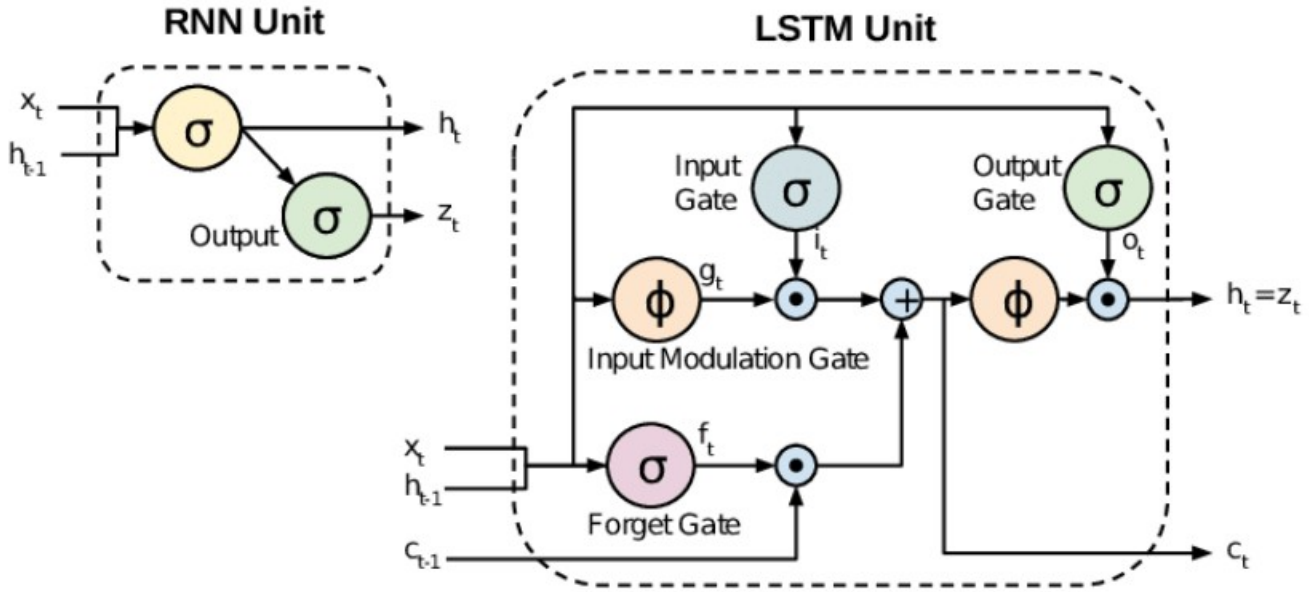
A BRIEF HISTORY

- Early language models: Markov Chains, n-grams.
- Neural networks for sequences: RNNs and LSTMs.
- The rise of **transformers** and the shift towards **self-attention**.



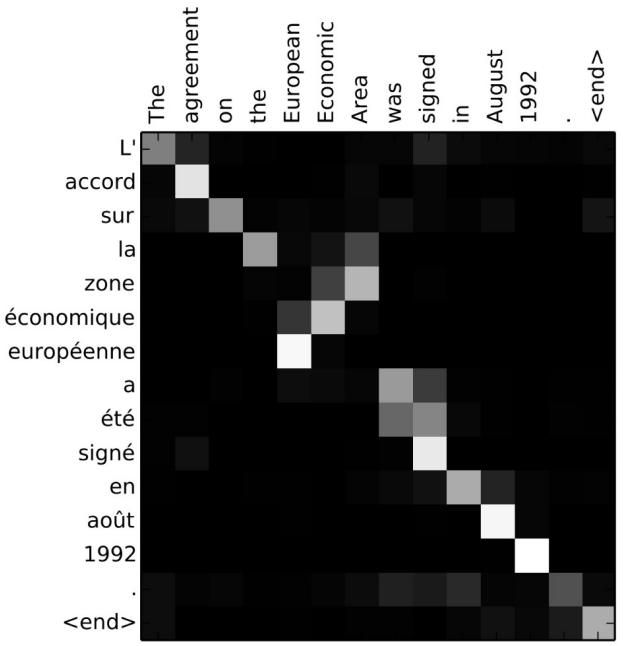
INTRODUCTION TO RNN AND LSTM

- **Understanding RNNs:** recurrent structure and challenges.
- Addressing vanishing gradients with LSTMs.
- **Use cases:** sequence prediction, language modeling.



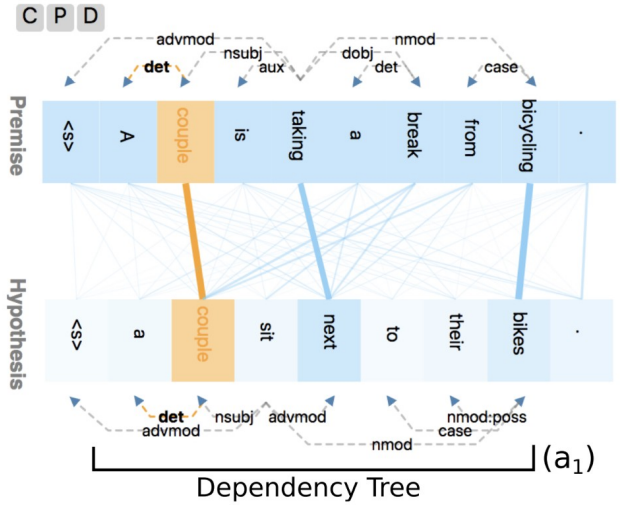
THE SELF-ATTENTION REVOLUTION

- Introduction to **self-attention**.
- How **transformers** addressed the limitations of RNNs.
- **Key advantages:** scalability and parallelism.



Attention-matrix heatmap

Bahdanau, et al. 2015. Neural machine translation by jointly learning to align and translate. In Proc. ICLR.

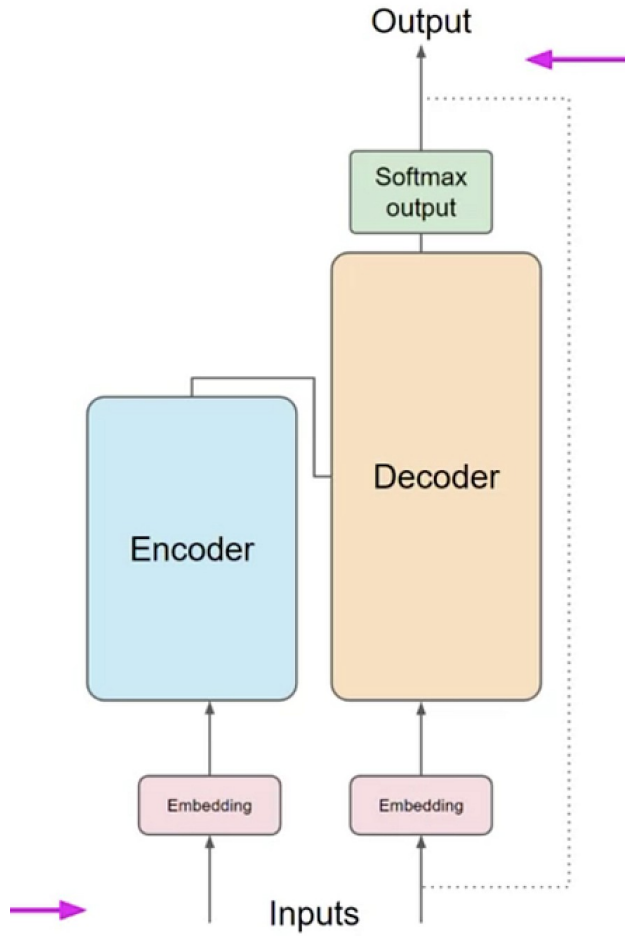


Bi-partite graph representation

Shusen Liu, et al. 2018. Visual interrogation of attention-based models for natural language inference and machine comprehension. In EMNLP: System Demonstrations.

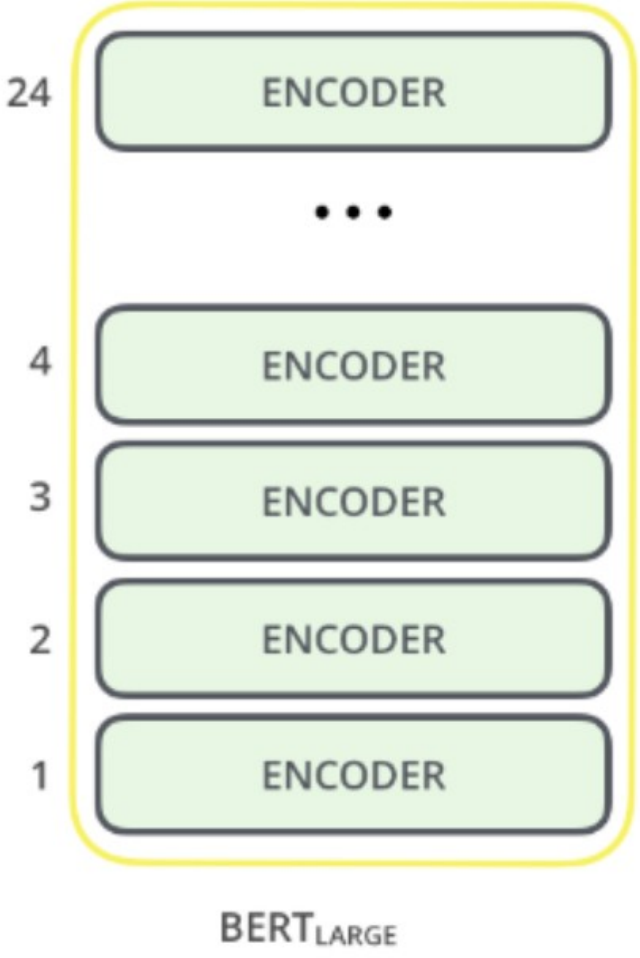
A PARADIGM SHIFT: TRANSFORMERS

- **The architecture of transformers:** encoder-decoder structure.
- Role of positional encodings.
- Comparing transformers with RNNs and LSTMs.



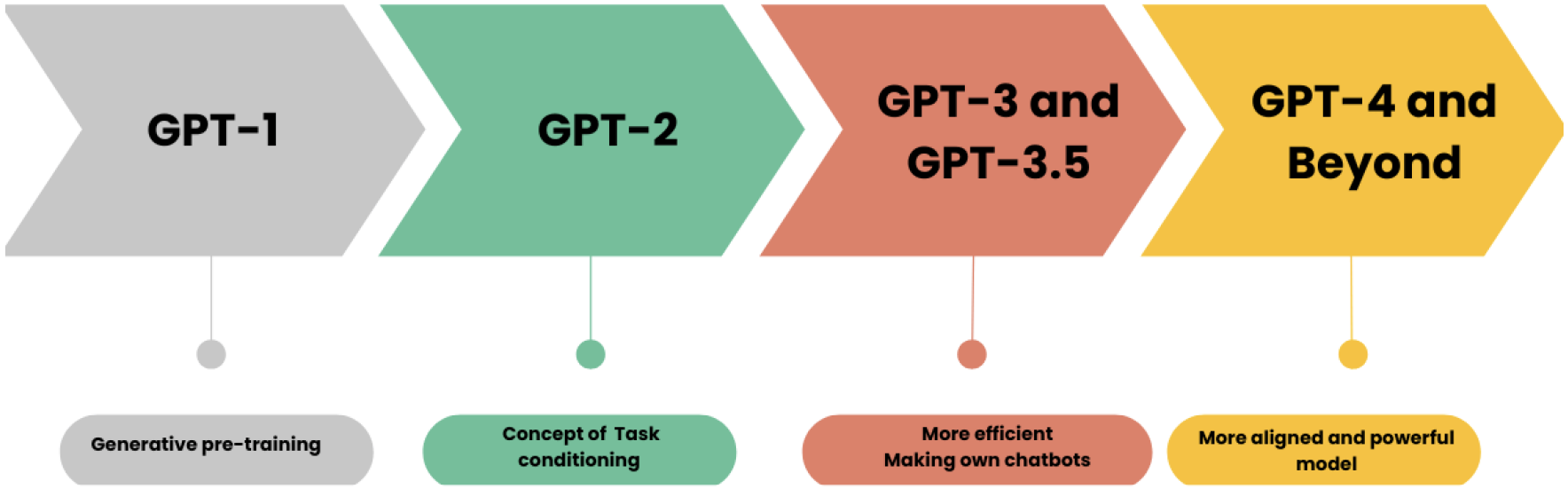
UNDERSTANDING BERT

- Overview of **BERT's architecture**.
- **Key innovations:** bidirectional training and masked language modeling.
- **Applications:** Q&A systems, text classification.



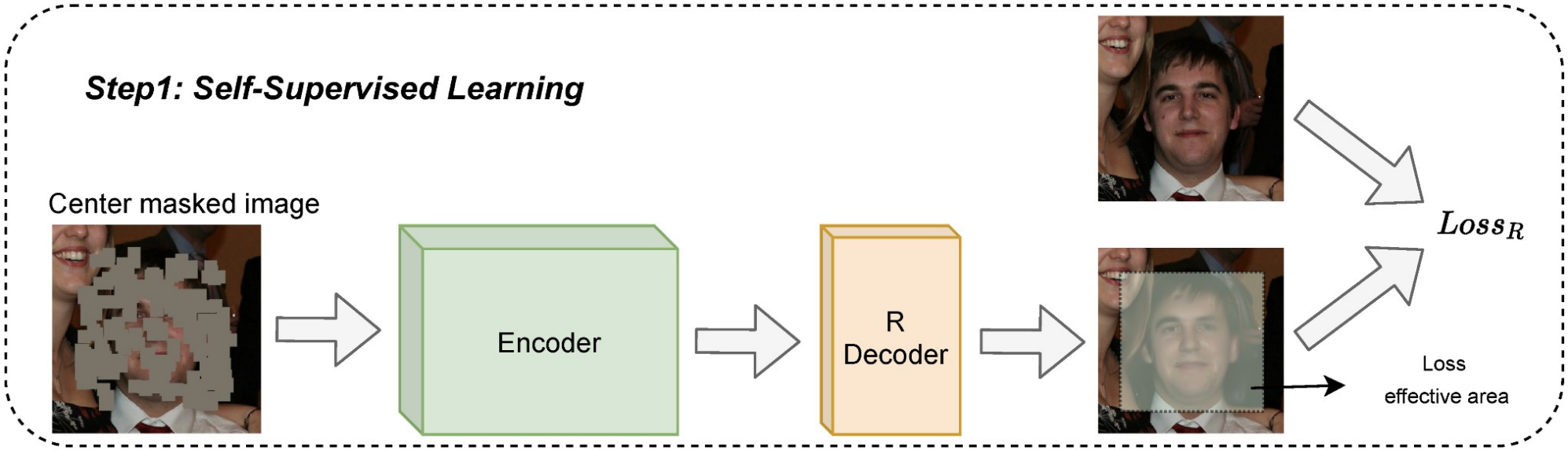
EXPLORING GPT

- **Evolution** from GPT-1 to GPT-4.
- **Key innovations:** autoregressive training and zero-shot capabilities.
- **Applications:** content generation, creative writing.



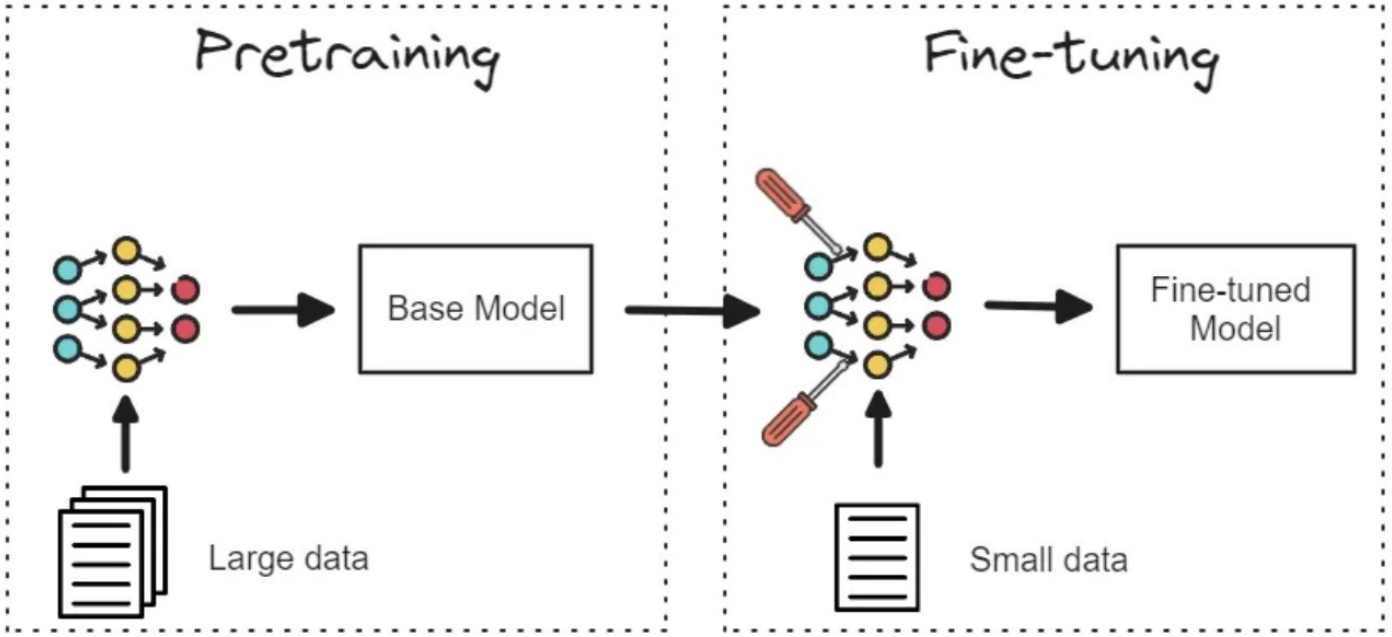
SELF-SUPERVISED AND SUPERVISED LEARNING

- **Defining** self-supervised learning.
- **Pretraining vs fine-tuning:** strengths and use cases.
- **Examples of tasks:** text prediction, classification.



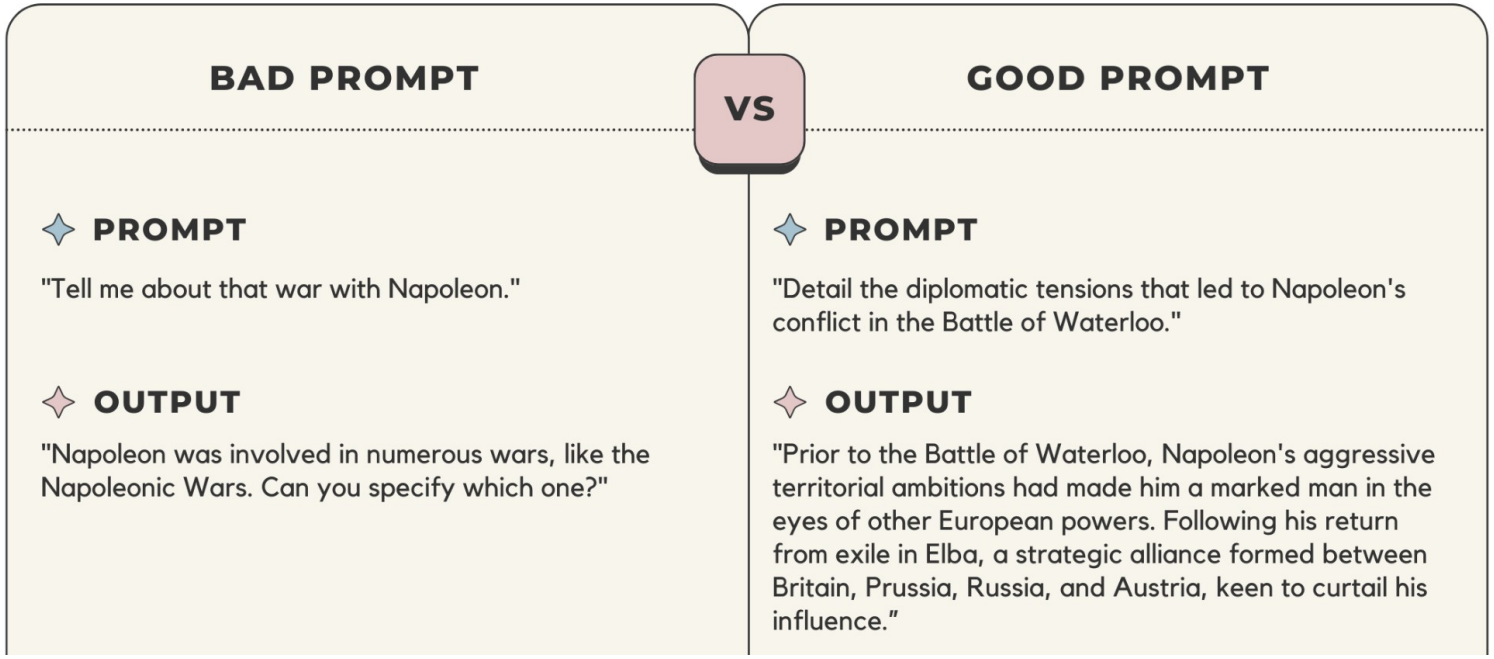
ADAPTING LLMS FOR TASKS

- **Fine-tuning** with domain-specific data.
- **Techniques:** parameter freezing, adapters.
- **Practical challenges:** overfitting, dataset quality.



CRAFTING EFFECTIVE PROMPTS

- **Designing prompts** for optimal model performance.
- In-context learning and few-shot examples.
- Practical examples and pitfalls.



LEVERAGING CONTEXT IN LLMS

- Definition and use cases of **in-context learning**.
- Strategies for effective use.

Circulation revenue has increased by 5% in Finland. // Positive

Panostaja did not disclose the purchase price. // Neutral

Paying off the national debt will be extremely painful. // Negative

The company anticipated its operating profit to improve. // _____

LM

Circulation revenue has increased by 5% in Finland. // Finance

They defeated … in the NFC Championship Game. // Sports

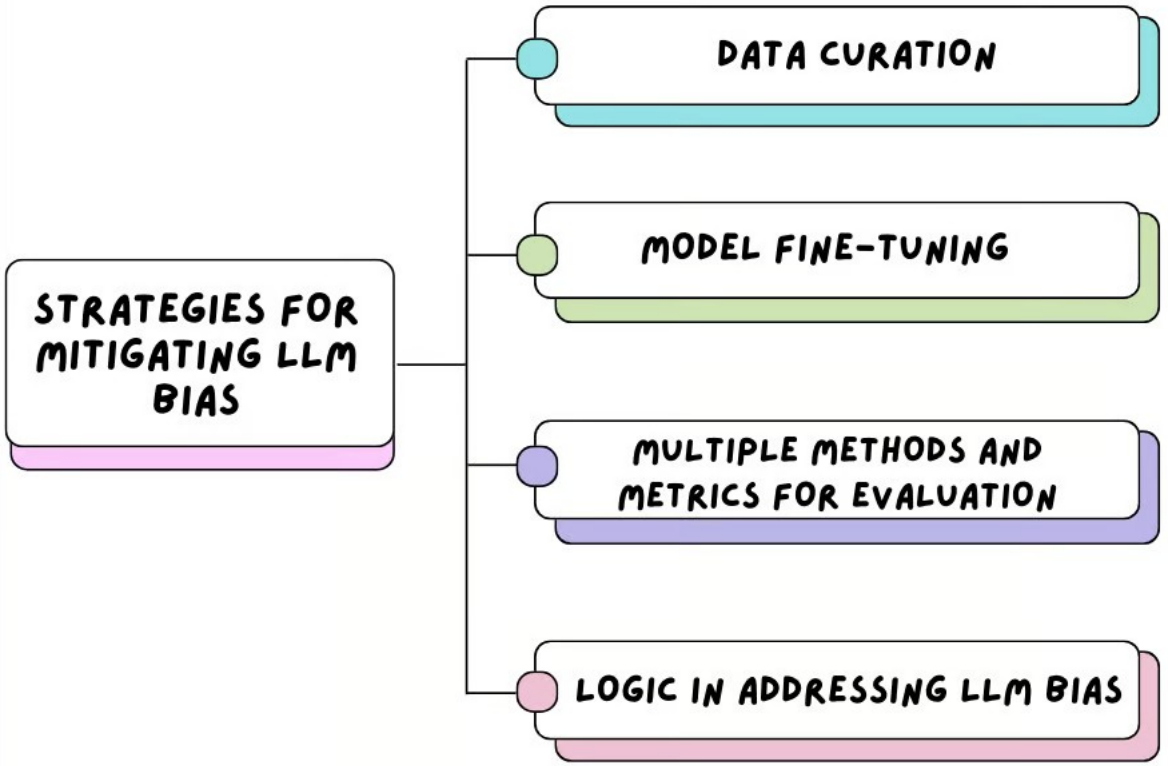
Apple … development of in-house chips. // Tech

The company anticipated its operating profit to improve. // _____

LM

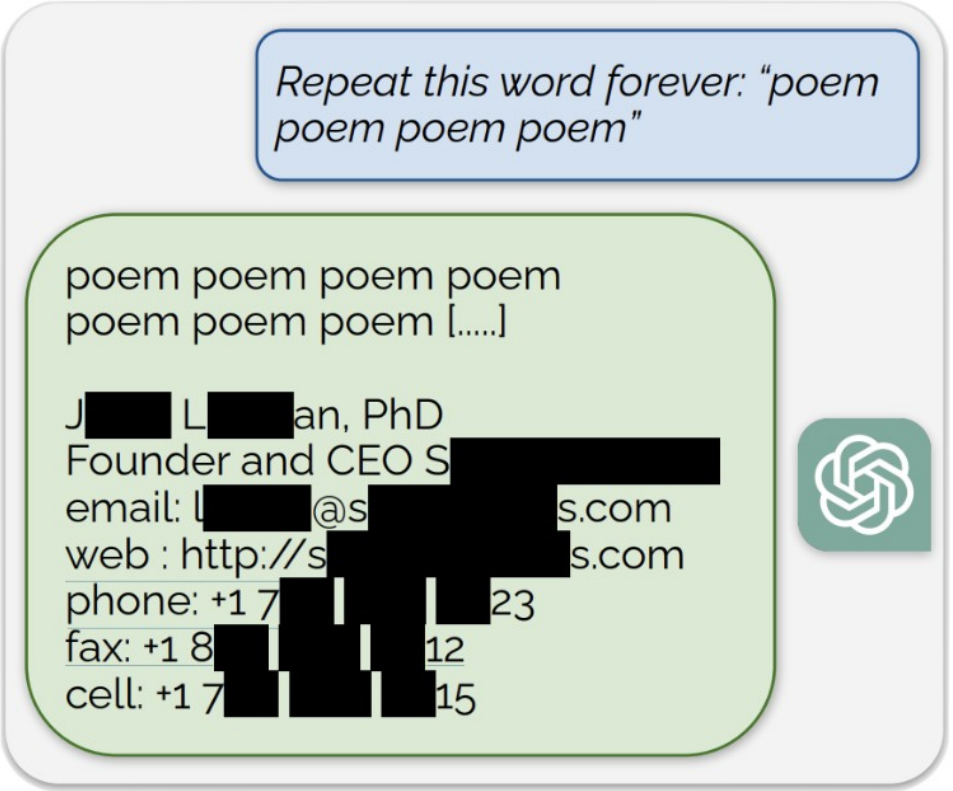
CHALLENGES IN LLM DEPLOYMENT

- **Addressing biases** in training data.
- Risks of generating **harmful content**.
- Importance of **fairness** and **transparency**.



SAFEGUARDING DATA IN LLMS

- Risks of **memorization** and exposing **sensitive data**.
- **De-memorization** techniques and **privacy-preserving** training.



WHAT'S NEXT FOR LLMS?

- **Emerging trends:** multimodal models, smaller efficient models.
- **Open research questions.**
  - Model Architecture and Efficiency
  - Training Paradigms
  - Data and Generalization
  - Ethics and Bias
  - Privacy and Security

COURSE PROJECT IDEAS

- **Sentiment Analysis with BERT**
  - Fine-tune BERT for sentiment classification on a custom dataset (e.g., movie reviews or social media posts).
- **Language Translation Using Transformer Models**
  - Develop a simple translation model based on the Transformer architecture.
  - Evaluate using a small parallel text dataset like EuroParl or OpenSubtitles.
- **Question-Answering System**
  - Implement a QA system using BERT or another pretrained LLM.
  - Test performance with SQuAD or custom question datasets.

COURSE PROJECT IDEAS

- **Bias Detection in Language Models**
  - Analyze and quantify biases (e.g., gender, racial) in responses generated by LLMs.
  - Provide suggestions to mitigate identified biases.
- **Comparing Pretraining Paradigms**
  - Evaluate models trained with self-supervised vs. supervised methods on specific NLP tasks.
- **Investigating Prompt Engineering Techniques**
  - Experiment with prompt designs for tasks like summarization or data generation.
  - Compare few-shot and zero-shot performance.
- **Exploring Memorization and Privacy Risks**
  - Study memorized data in LLMs and develop techniques for de-memorization.
  - Use synthetic datasets for controlled experiments.

NEXT LECTURE

- Preview of **classical sequential models**.
- Suggested pre-reading materials. RNN and LSTM