

MACHINE LEARNING FOR SIGNAL PROCESSING

24-3-2025

Sriram Ganapathy

*LEAP lab, Electrical Engineering, Indian Institute of Science,
sriramg@iisc.ac.in*

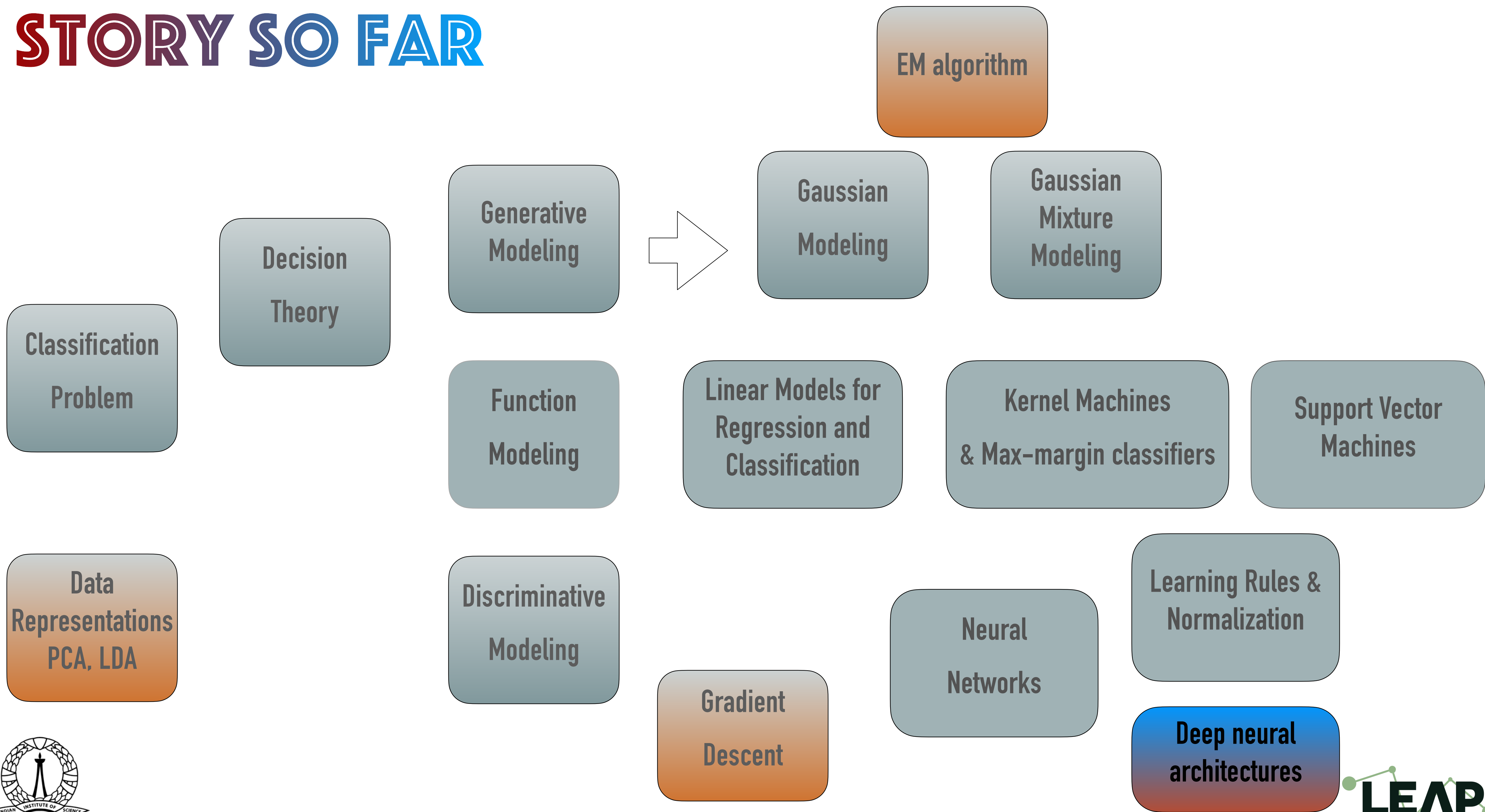
Viveka Salinamakki, Varada R.

LEAP lab, Electrical Engineering, Indian Institute of Science

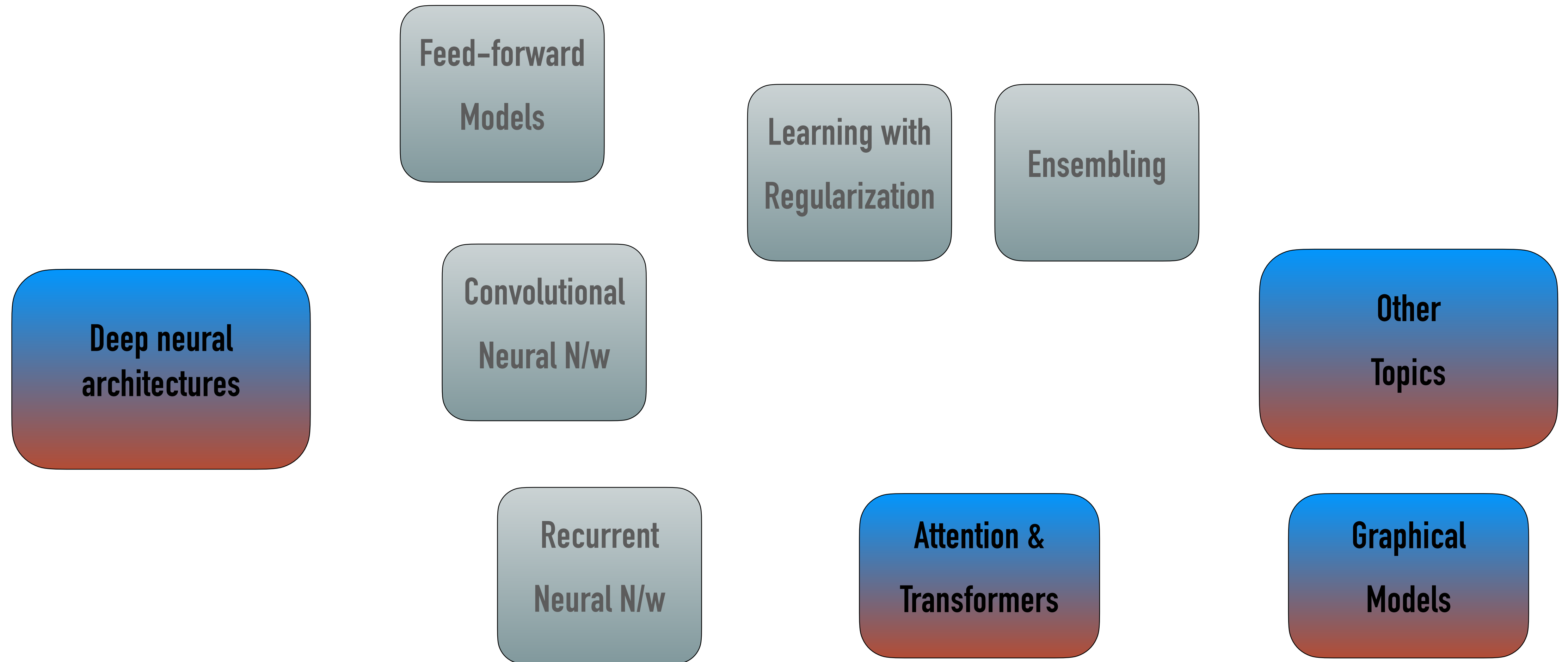
<http://leap.ee.iisc.ac.in/sriram/teaching/MLSP25/>



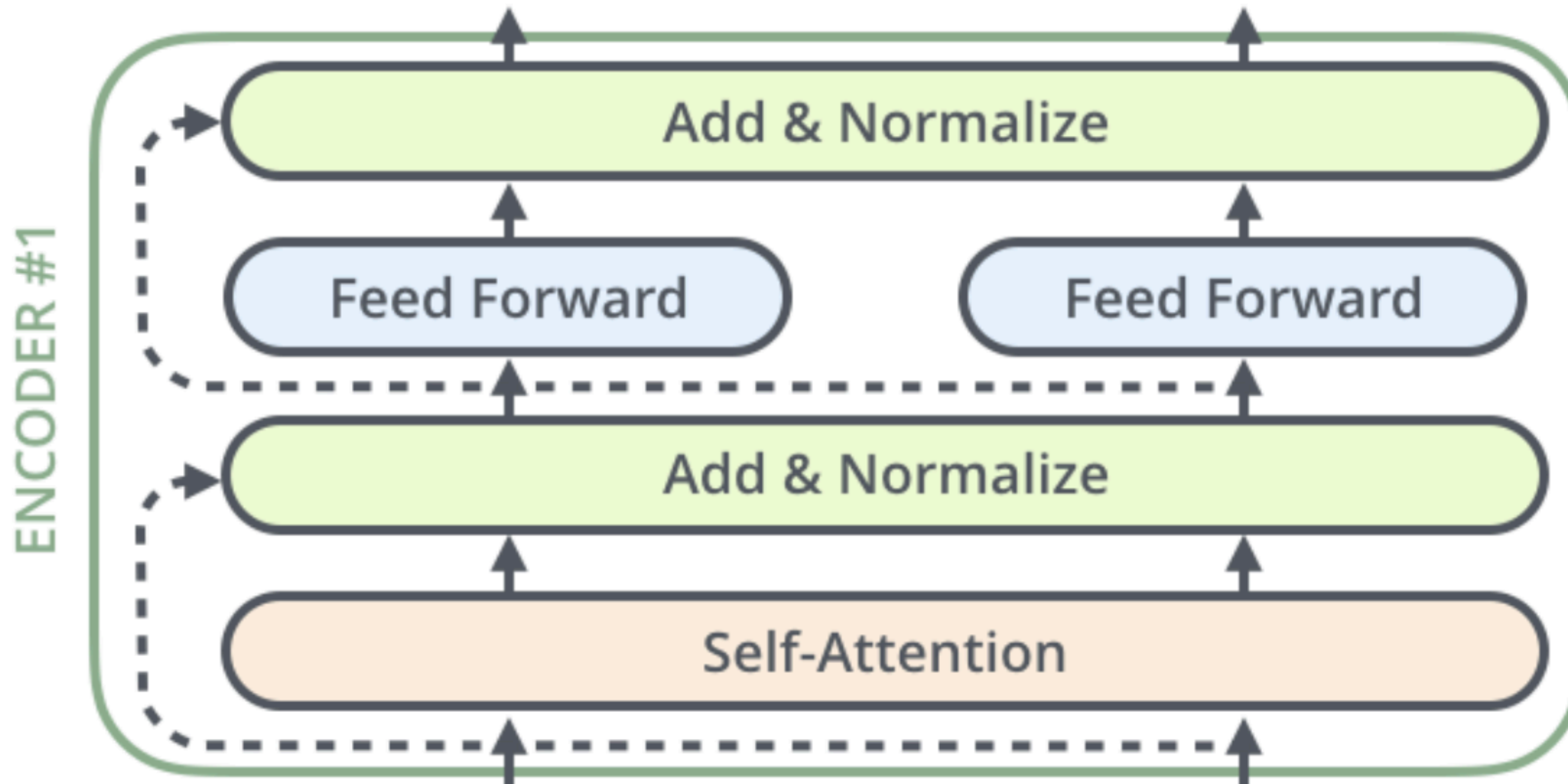
STORY SO FAR



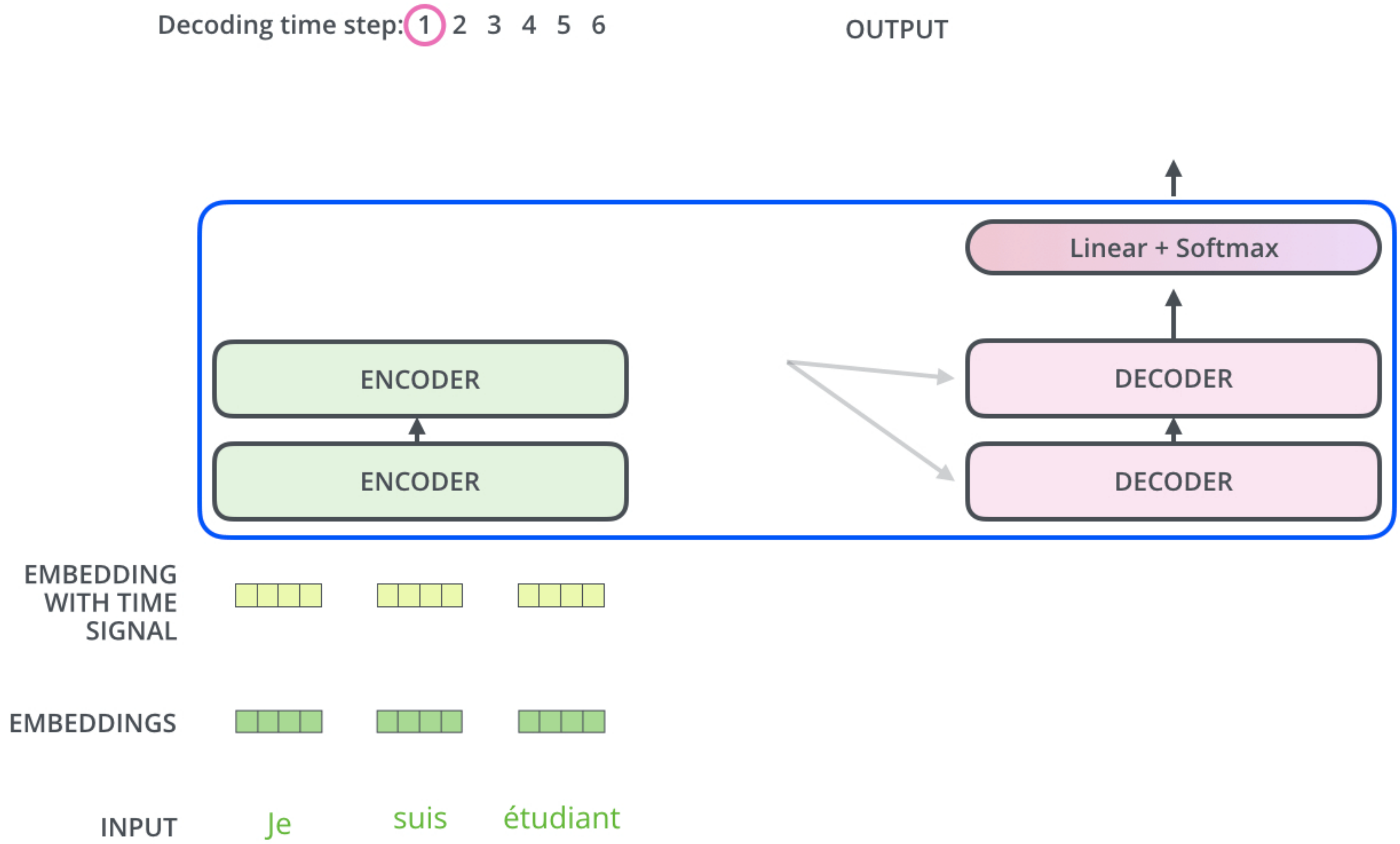
STORY SO FAR



Transformer encoder

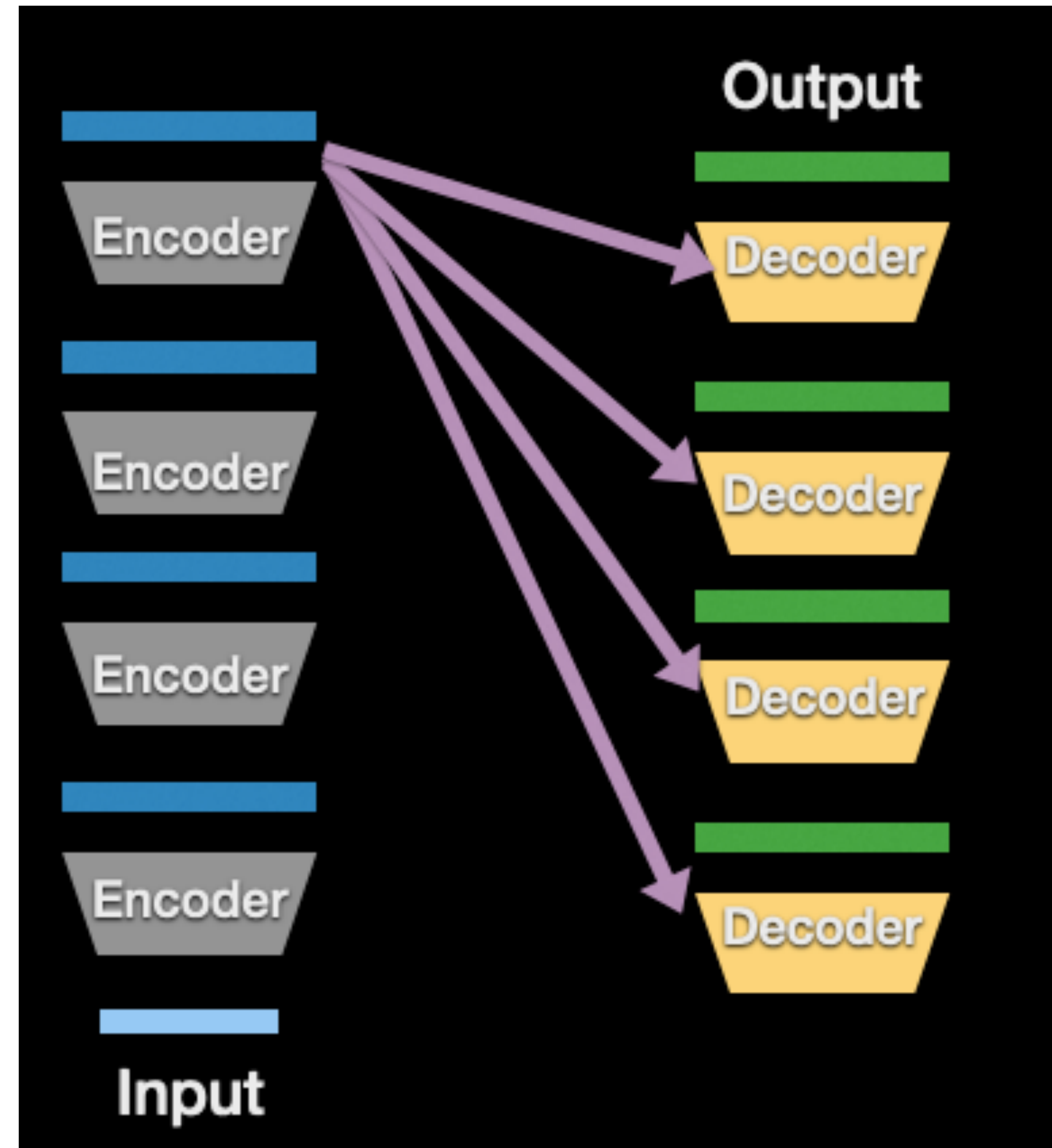


Transformer

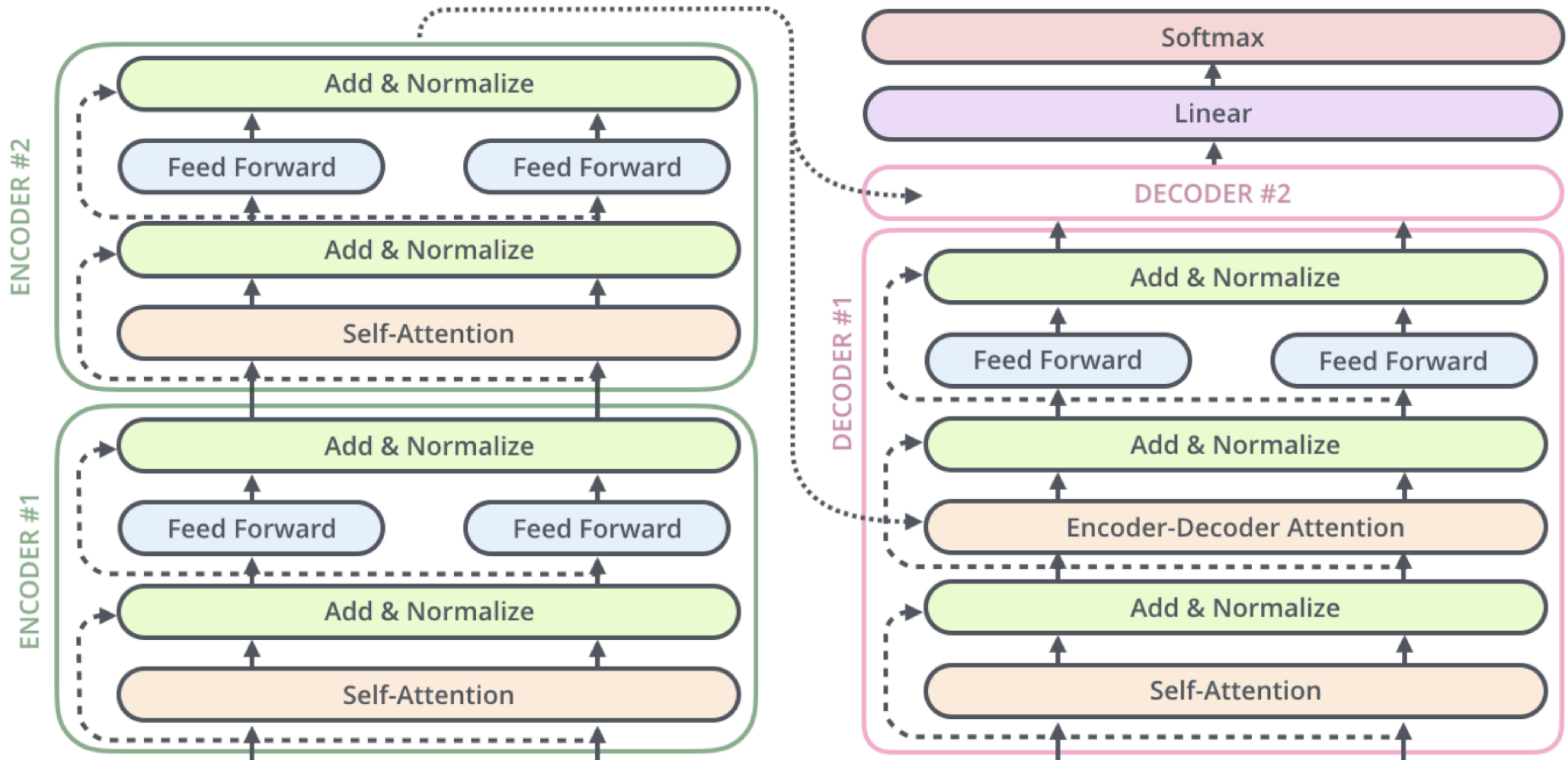


Pics taken from : <https://jalammar.github.io/illustrated-transformer/>

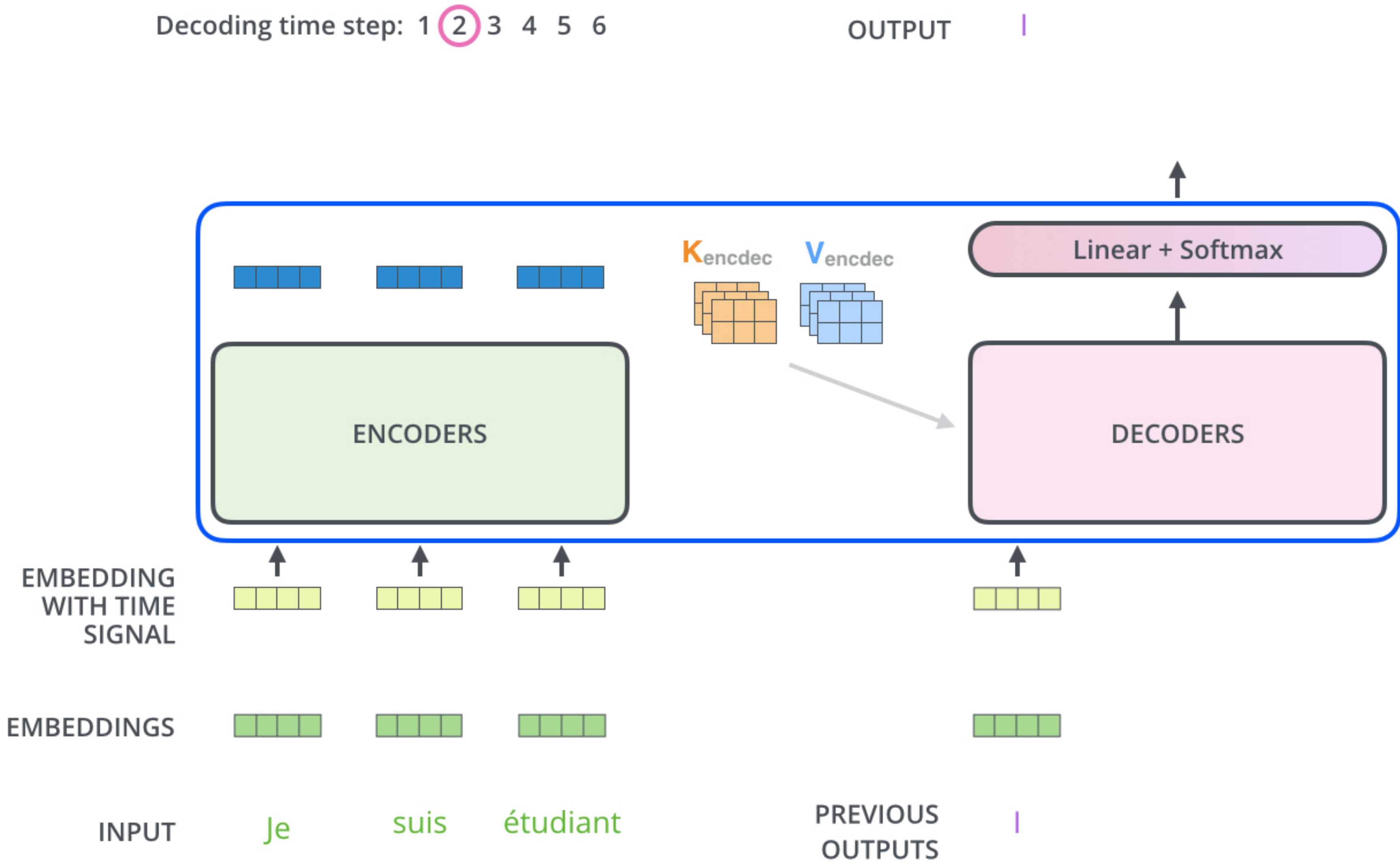
Encoder- Decoder Attention



Transformer decoder



Transformer Example

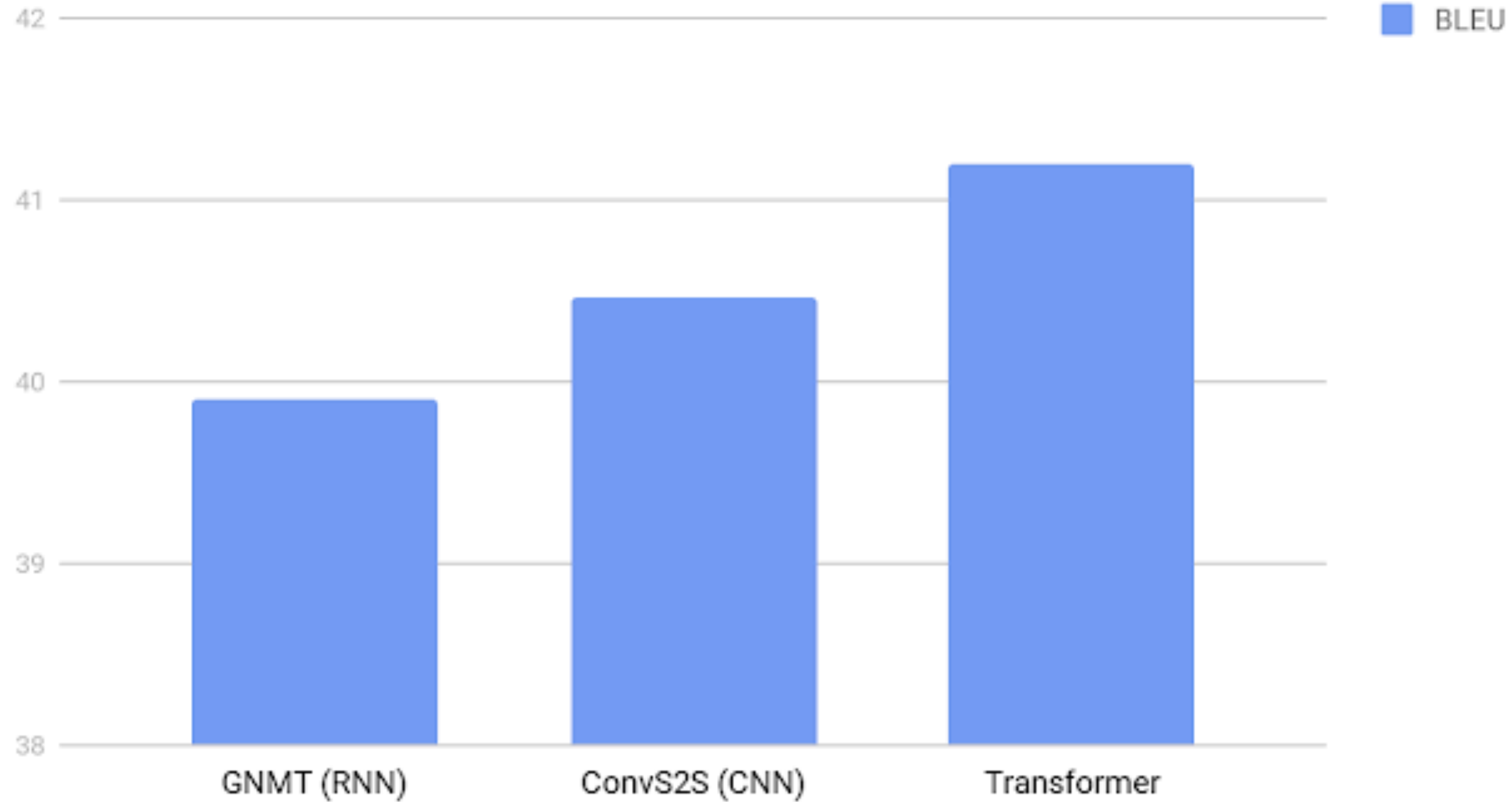


Pics taken from : <https://jalammar.github.io/illustrated-transformer/>

Neural Machine Translation Example

Neural Machine Translation Example

English French Translation Quality

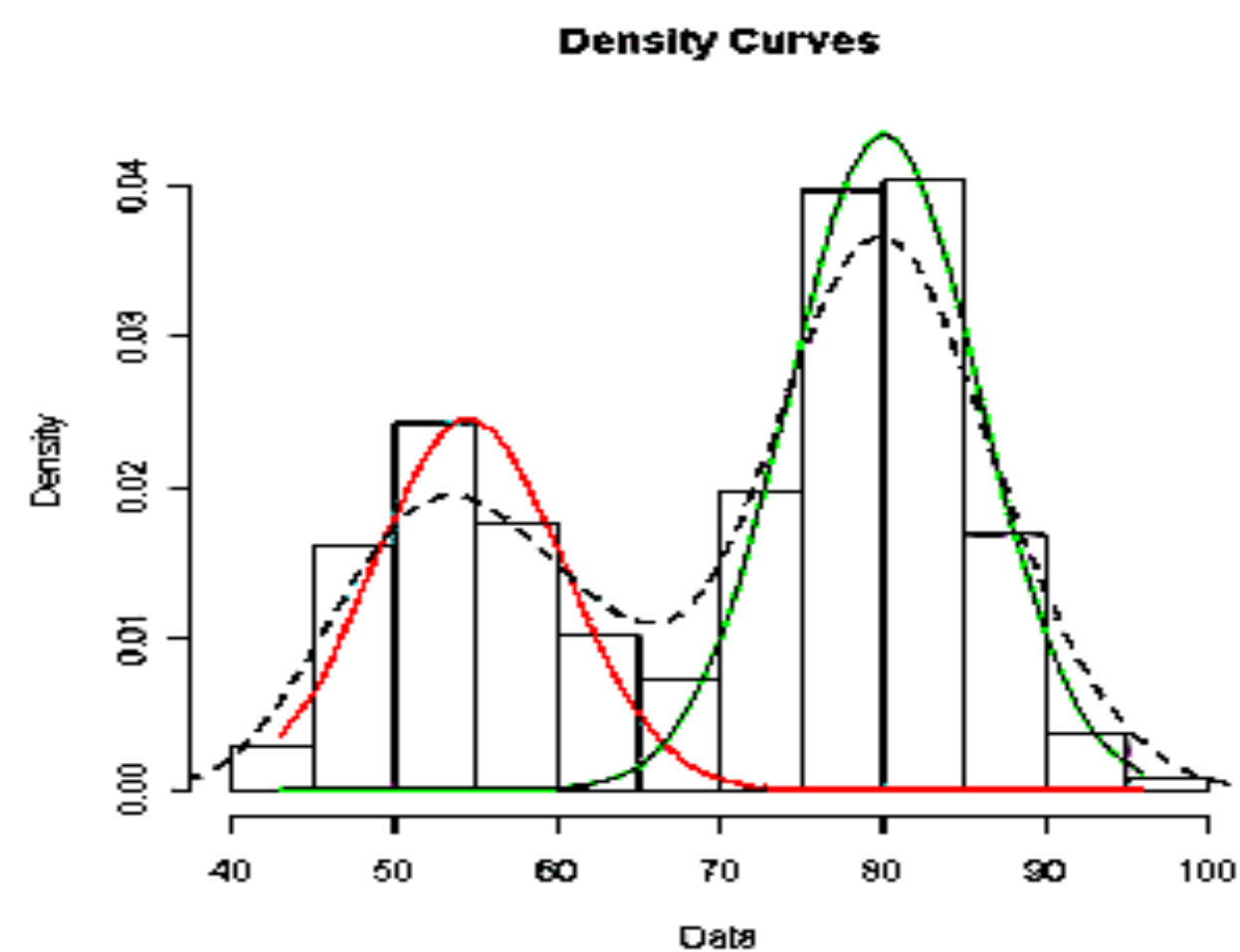


Unsupervised Learning

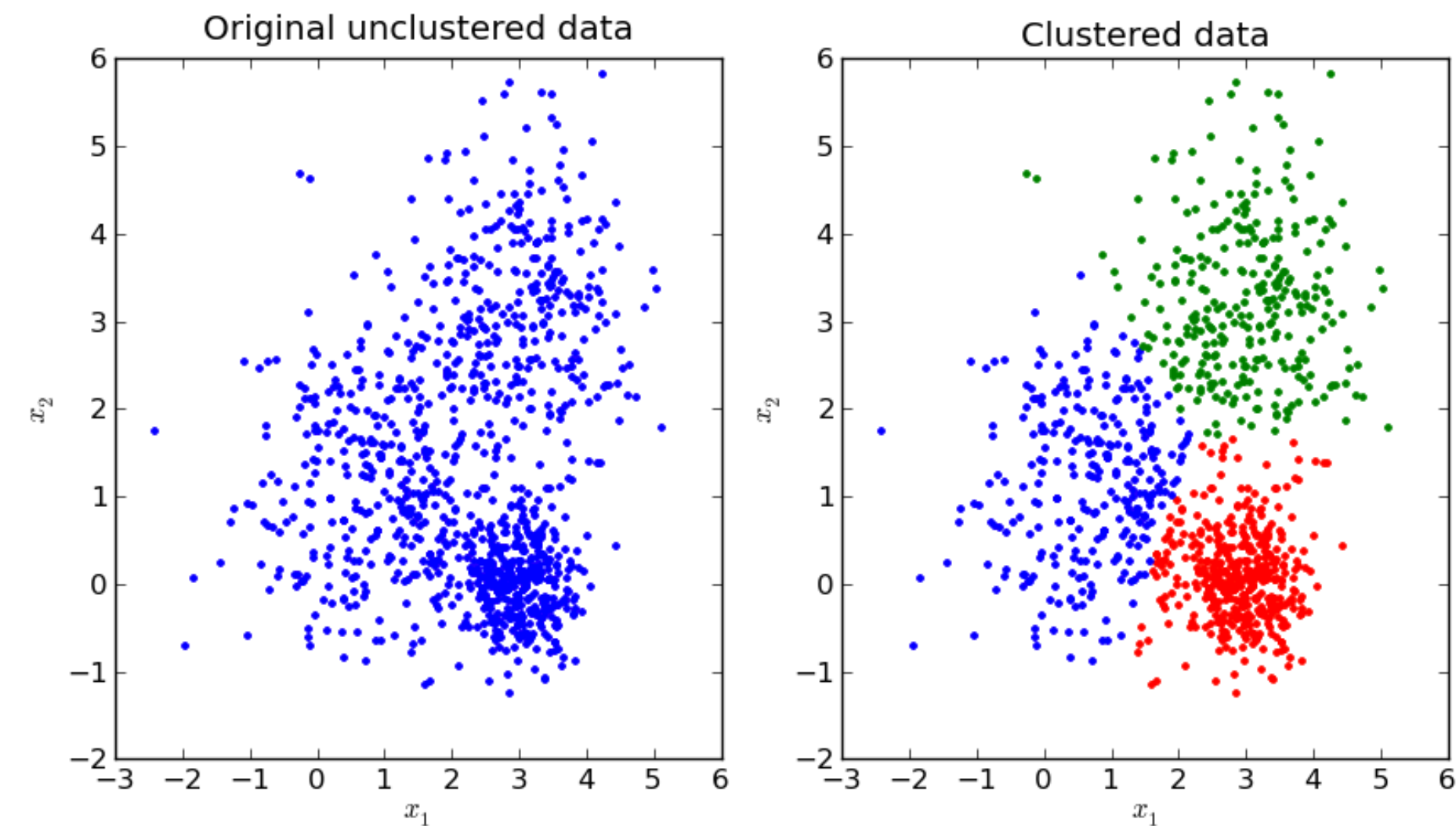
Unsupervised Learning

- Developing models that do not need labels
- May model the generation of data.
- May allow generation of new data samples
- Broad strategies for unsupervised learning

Learning the distribution of the data

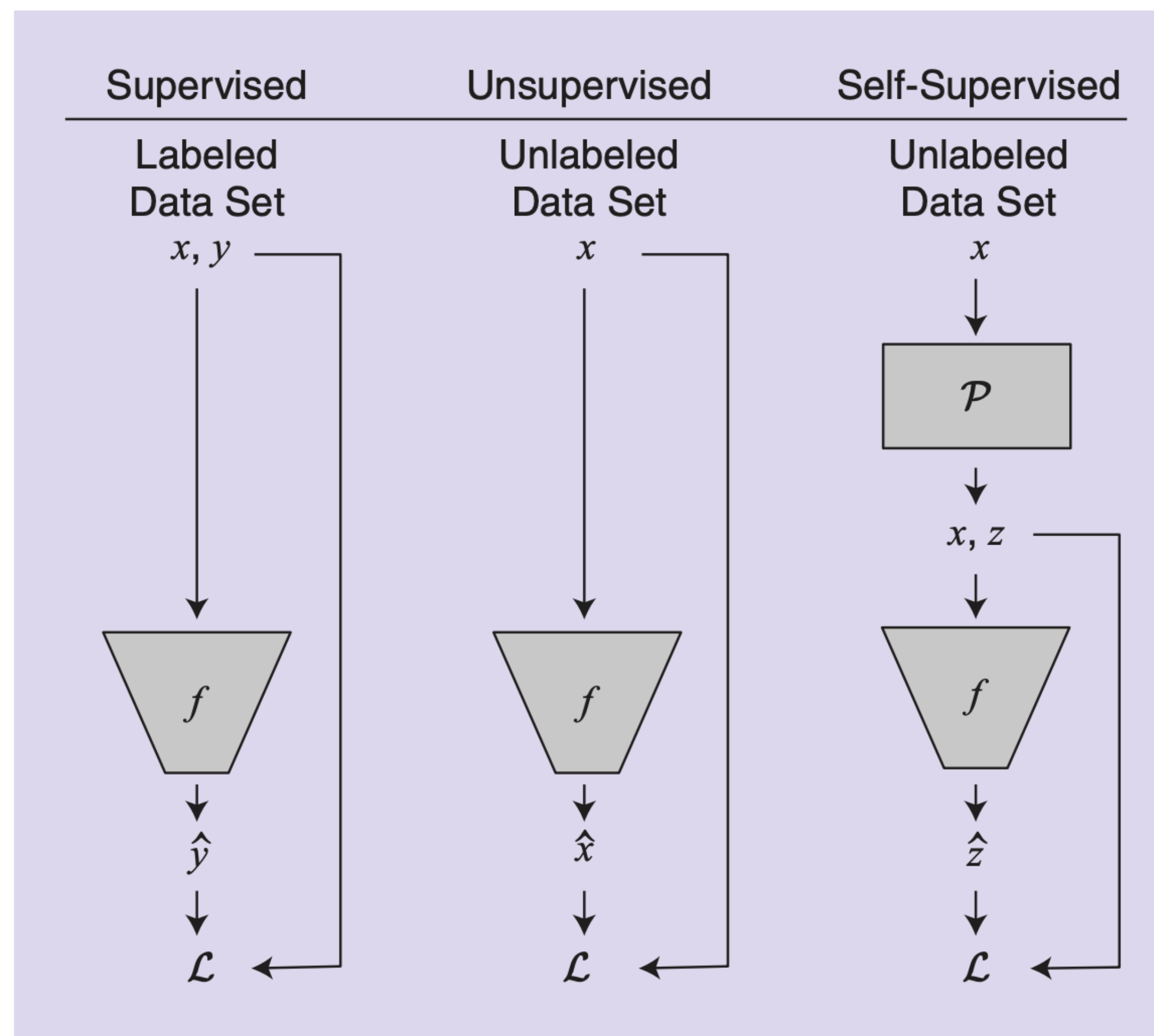


Detecting clusters in the data



Self supervision

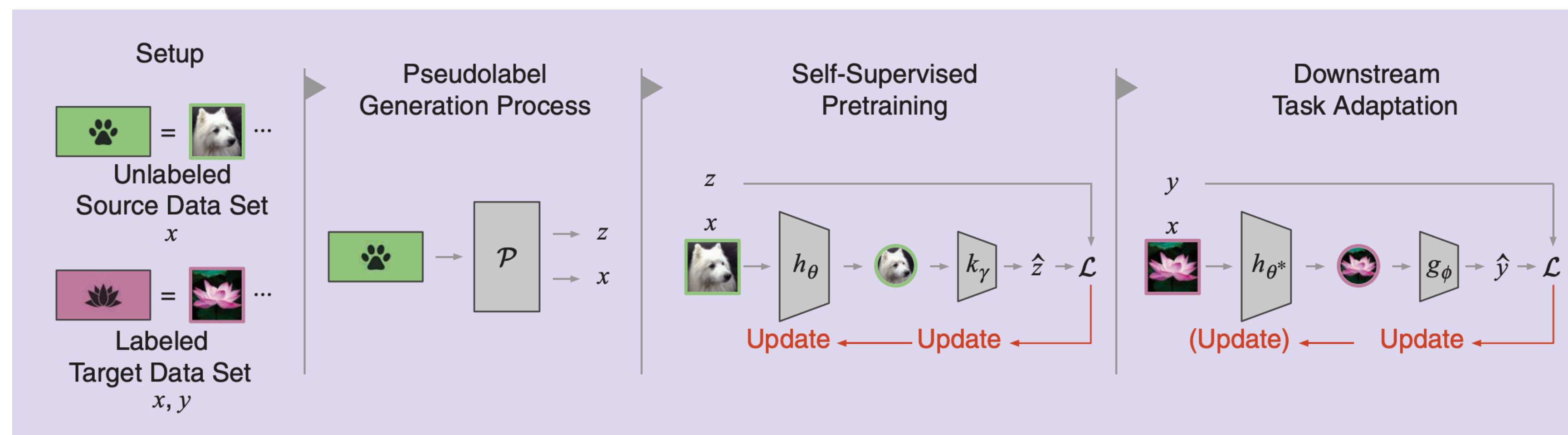
- ◆ Different from supervised and unsupervised learning
- * Does not perform distribution learning or reconstruction
- * Uses a pretext task
- * Performing contrastive or predictive learning
- ◆ Using large volumes of unsupervised data



Ericsson, Linus, et al. "Self-supervised representation learning: Introduction, advances, and challenges." *IEEE Signal Processing Magazine* 39.3 (2022): 42-62.

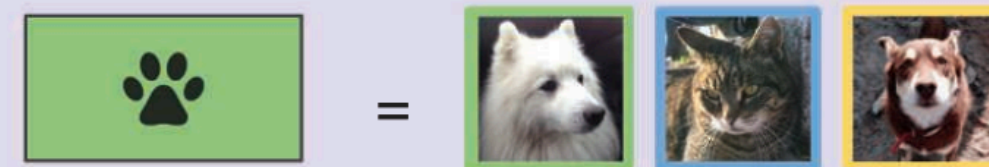
Self supervision - principle

- ◆ Two levels of modeling with unsupervised data
 - ❖ Generating a pseudo-label
 - ❖ Learning the upstream model
- ◆ Downstream task performs fine-tuning of the SSL model.

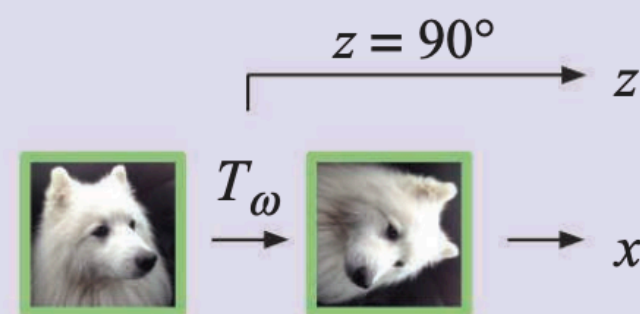


Self supervision - pre-text task

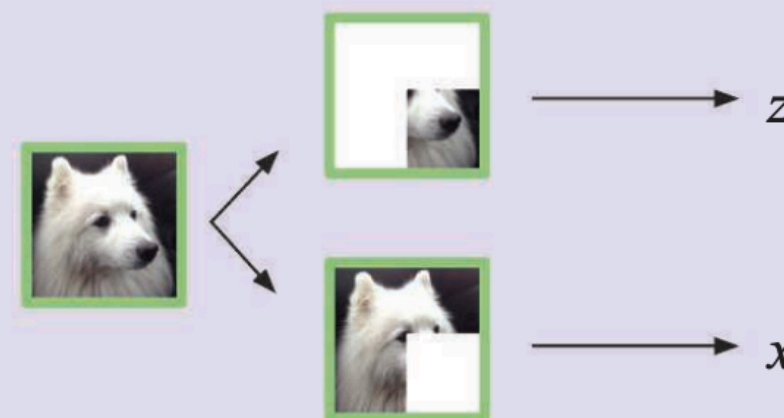
Pseudolabel Generation Processes



Transformation Prediction



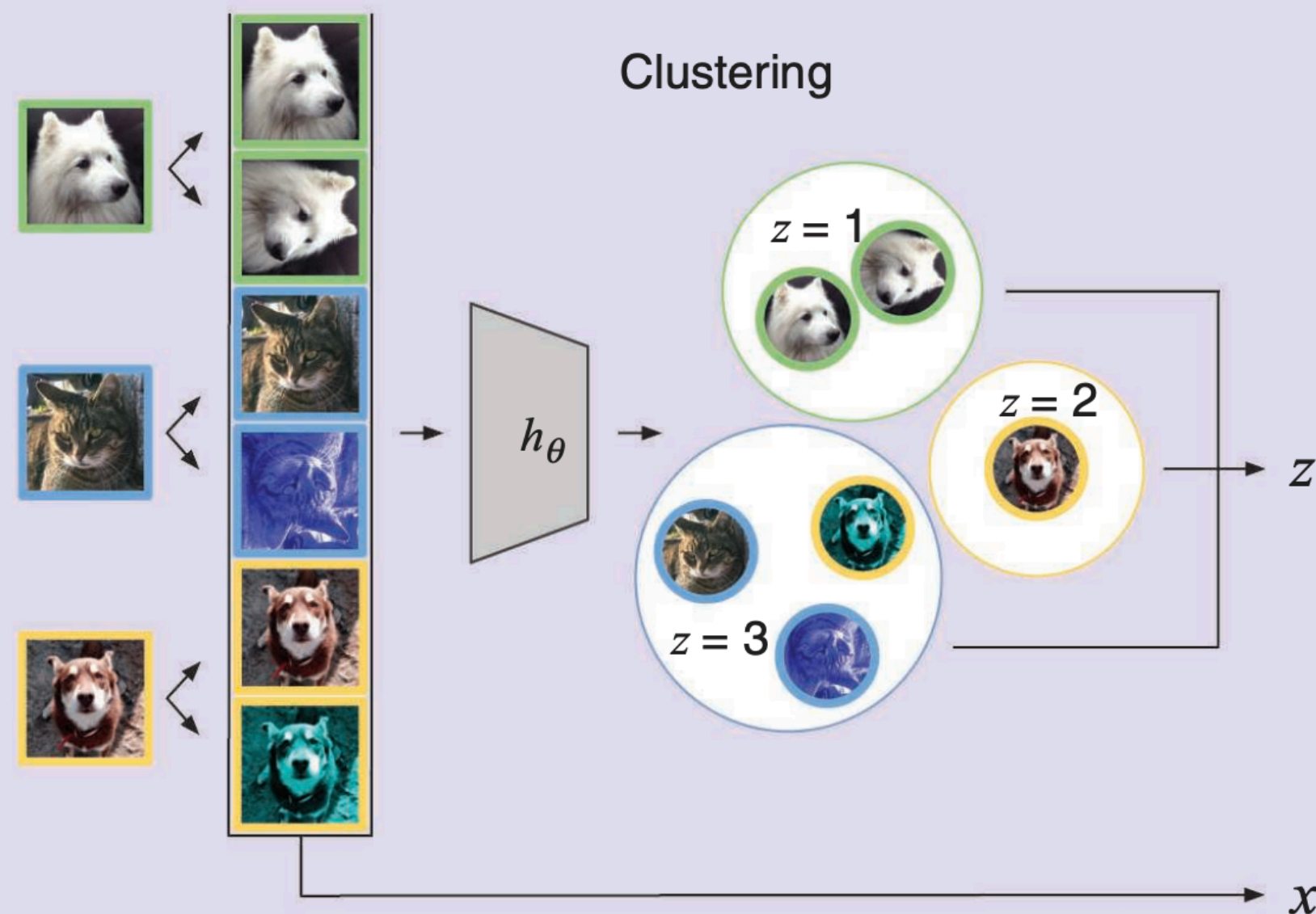
Masked Prediction



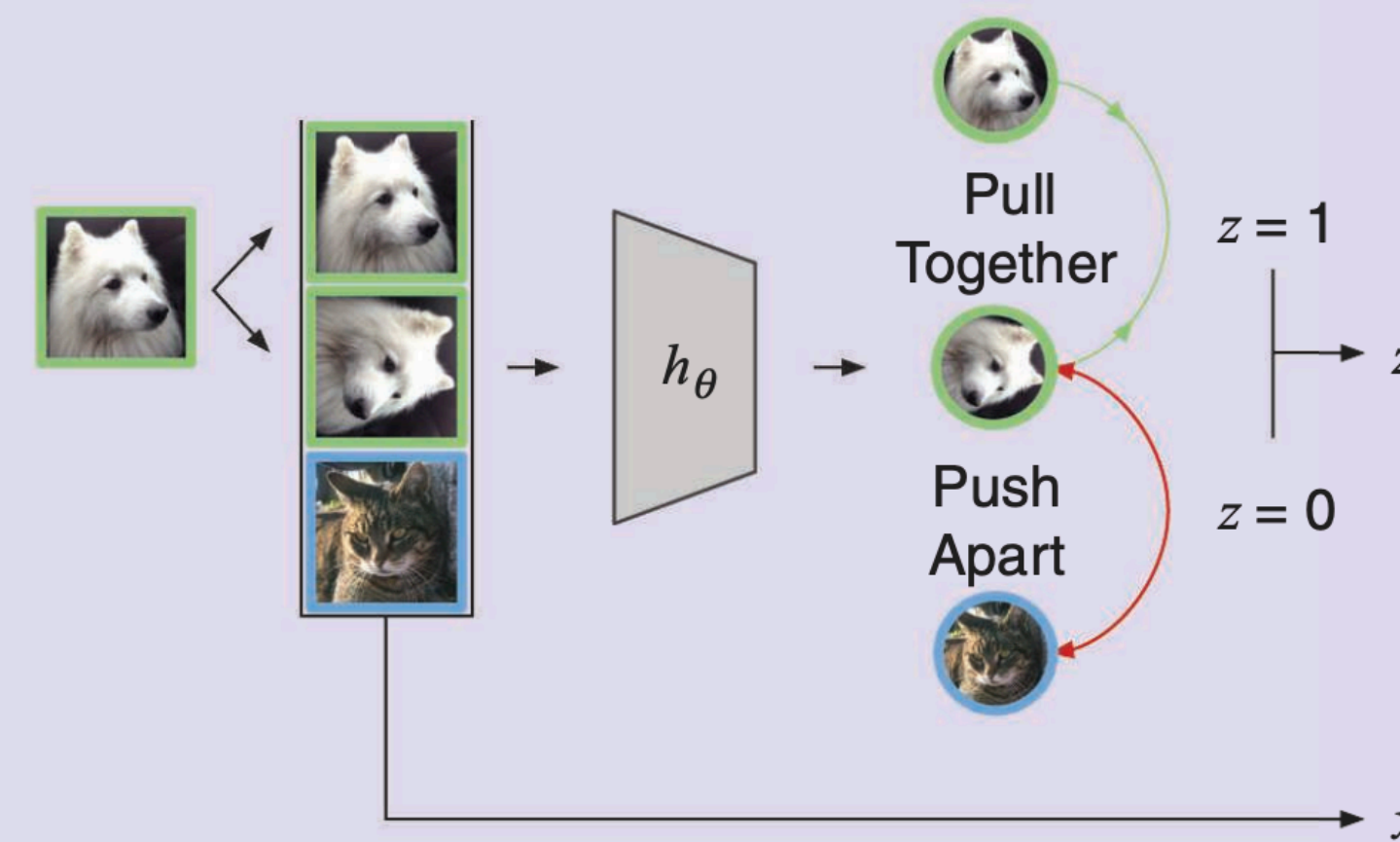
Instance Discrimination



Clustering

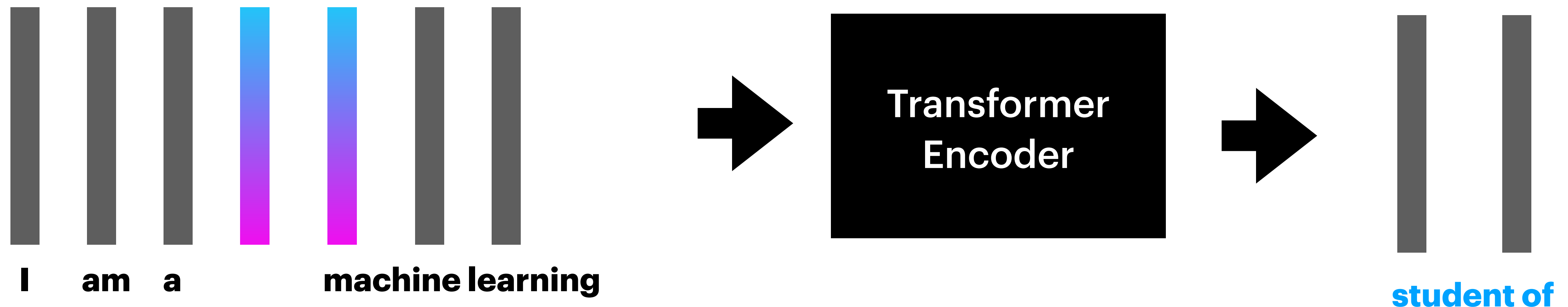


Contrastive Instance Discrimination



Self-supervision as a task

- Masking out portions of the input data
 - * Pass the rest of the embeddings (with zeros or random entries at the masked locations) to the transformer encoder
 - * Have the model predict the word tokens in the masked portions - **Masked Language Modelling (MLM)**



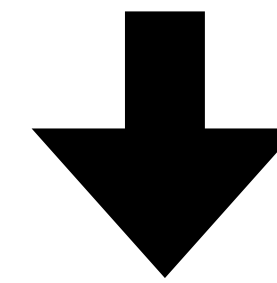
Large language models (LLMs)

- Extending the task of self-supervision
- Mine lots of text data
 - * Crawled from the web, as well as, from other resources.
- Design the **model with large capacity** (Millions → Billions of parameters)
- Pre-train the model
 - * With MLM and similar style of losses
 - * High resource of computations.
- Final trained model can be **fine-tuned for supervised tasks**
 - * Load the parameters as initialization and perform supervised learning.

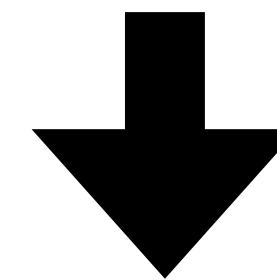
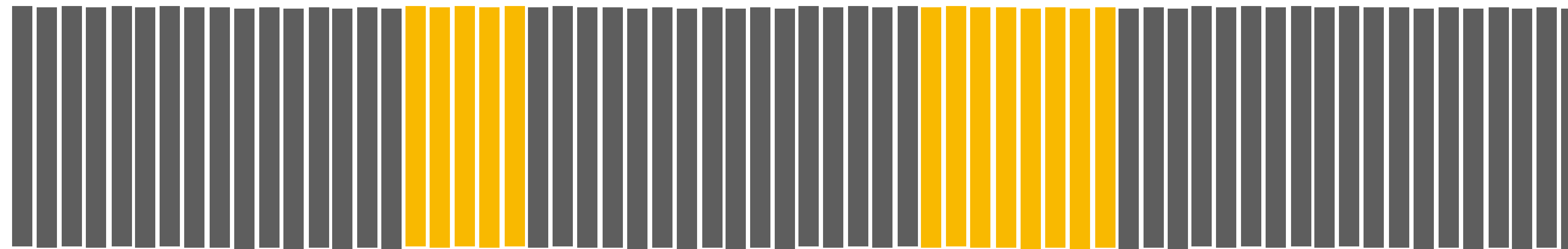
Large language models (LLMs)

- Self-supervised learning
 - * Has shown emergent abilities to generalise to wide variety of downstream tasks.
 - ✓ Tasks that the model was not trained on
 - ✓ Not seen in smaller models
 - * Enables to build reasoning capabilities in the model.
 - * **Applicable for several domains** - text, speech and images.

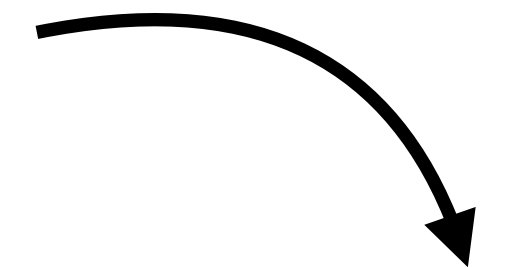
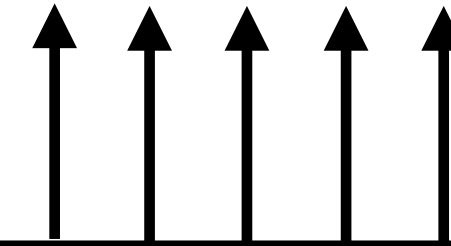
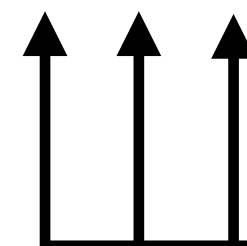
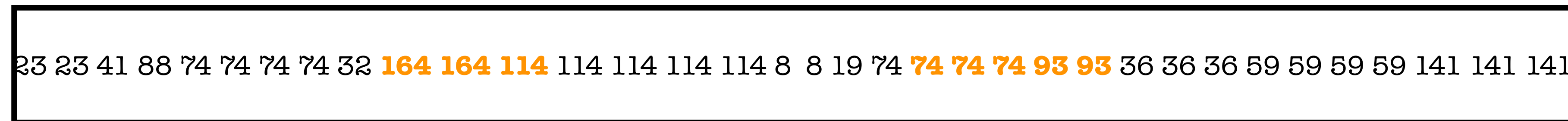
Self-supervision in audio - wav2vec



Sequence of spectral vectors



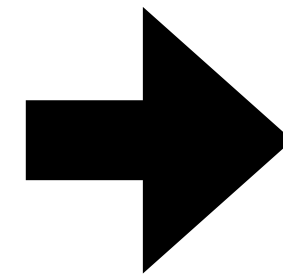
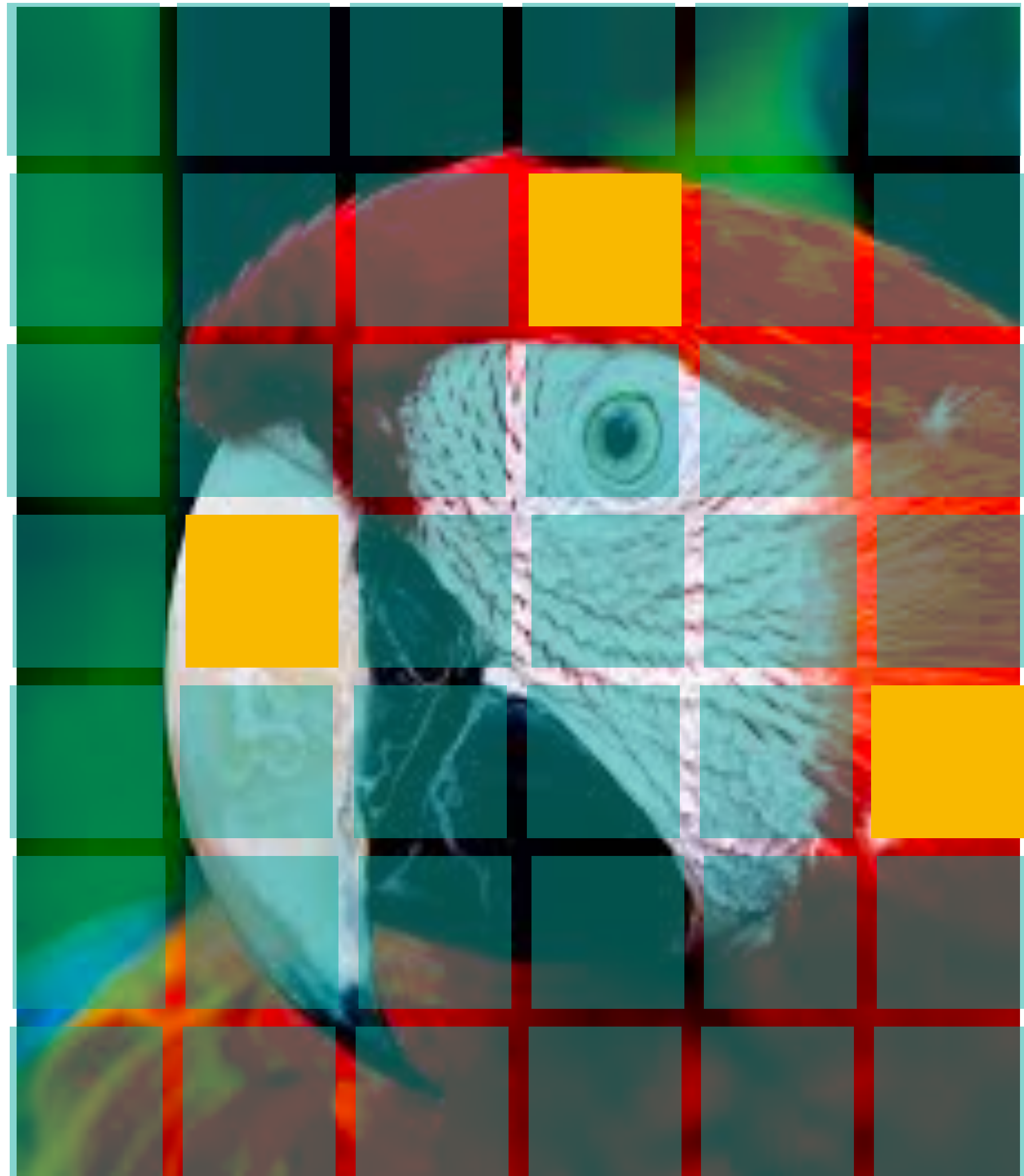
Quantization



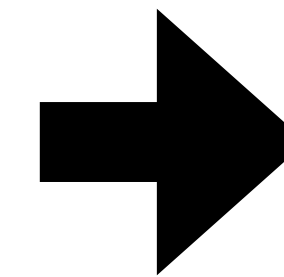
**Transformer
Encoder**



Self-supervision in images - Vision Transformer



Transformer
Encoder



Predict
Masked
Regions

LLM - Examples

- Generative Pre-trained Transformers (GPT) series

	Architecture	Data Used	Model Size
GPT-1	Transformer (12 layer, decoder only model)	Book Corpus (4.5GB)	117M
GPT-2	GPT-1 with additional normalisation layers	Web Text (40GB)	1.5B
GPT-3/3.5	GPT-2 with more layers Adding Fine-tuning tasks and human feedback	Large Web Crawl (570B)	175B
GPT-4/4o	Details Undisclosed [Trained with Text + Images]	—	—

Future works (some already underway)

- Multi-modal

- * Incorporating learning across modalities

- ✓ Create a domain specific encoder/decoder and learning the joint language model.

- Combining some labeled data with the self-supervised data to further improve the models.

- ✓ Current models like GPT use human feedback.

- Understanding the risks and vulnerabilities of these models.

Graphical Models

Graphical Models

- What are graphs
 - * What are graphical models
- Directed and Undirected graphs
- Conditional independence

THANK YOU

Sriram Ganapathy and TA team
LEAP lab, C328, EE, IISc
sriramg@iisc.ac.in

