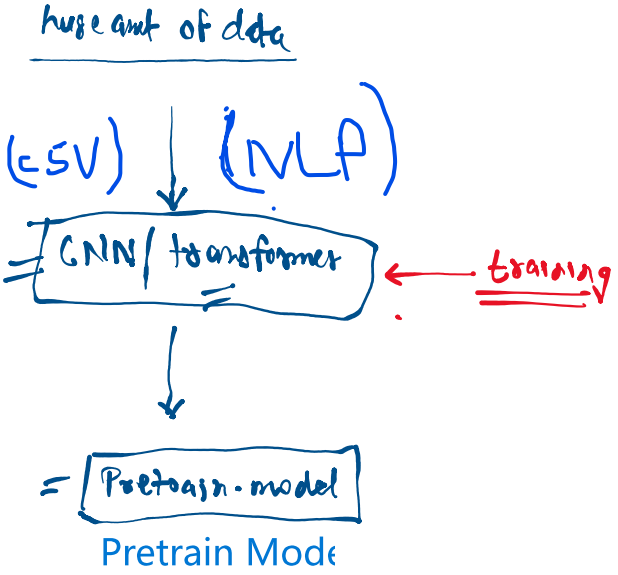
What is Transfer Learning and Finetuning?

Monday, April 28, 2025 4:04 PM

- **"Transfer learning means taking experience from one problem and using it to solve another related problem."**

Real Life Example: "It's like learning to ride a bicycle and then using that skill to learn how to ride a motorcycle faster."

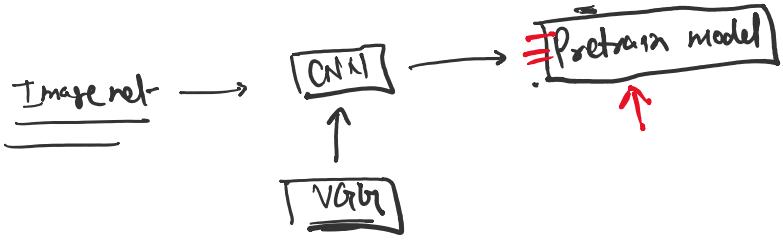
Computer Scientist took this simple idea and integrated into the deep learning.



- **ILSVRC (ImageNet Large Scale Visual Recognition Challenge)** version mein:
  - **Training Images**: ~1.2 million
  - **Validation Images**: 50,000
  - **Test Images**: 100,000 (labels hidden, competition ke liye)
  - **Classes**: 1000 (dog, cat, car, airplane, etc.)
- **Full ImageNet Dataset** (original, jo Stanford maintain karta hai):
  - **14+ million** images
  - **21,841** categories (WordNet synsets ke hisaab se)

This is example of those 1000 classes

Cat
Dog
Vehicle
Aircraft
Bird
Fruit
Vehicle
Sports Object
Musical Instrument
Humans



Now suppose your specific class is not included among the 1000 classes available in the ImageNet dataset.

Golden retriever dog
Sunny
Tata Nao

Now the ways of doing this Transfer Learning:

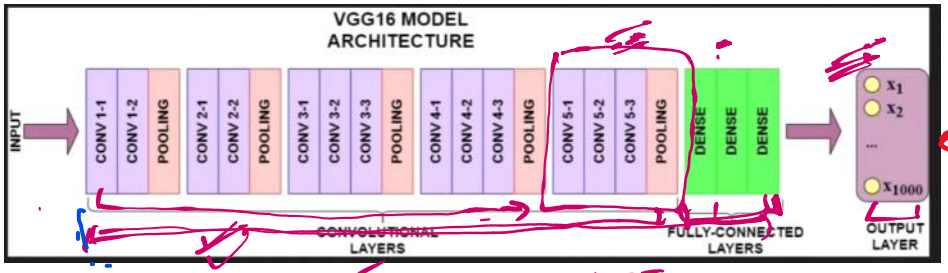
In CNN based Architecture: # 3 layer step

- **Replacing the output layer of pretrain model**
- Or don't change weights means keep freezing all the weights just retrain the Neural Network Weights
- Or **unfreezing some layers** of the network and **fine-tuning** it on the new, task-specific dataset

Now Lets take a very simple example:

First understand it using the CNN based Architecture.

First understand it using the CNN based Architecture.



Convolutional Neural Network

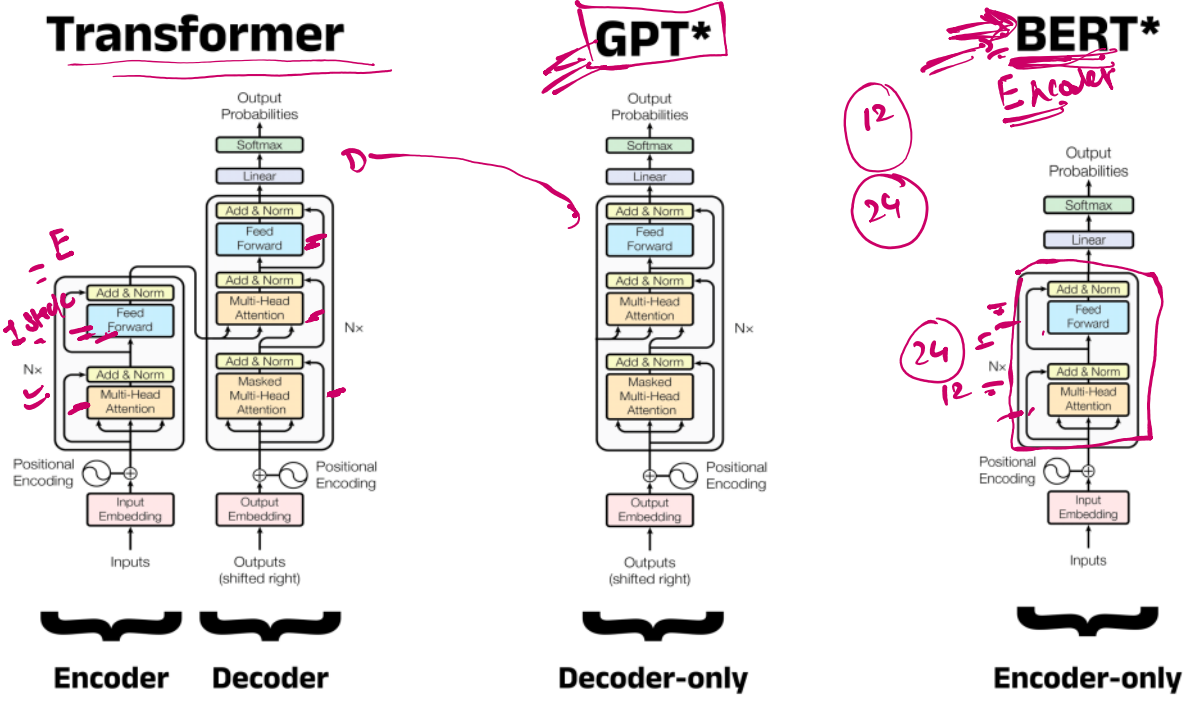
why this transfer learning works

We train a model on a **huge dataset** initially, allowing it to **learn primitive features**(like edges, textures, and simple shapes).
When applying it to a **different dataset**, we don't retrain the entire model — we only retrain certain layers if the new data is **very different**.
Otherwise, we can directly reuse the model's knowledge.

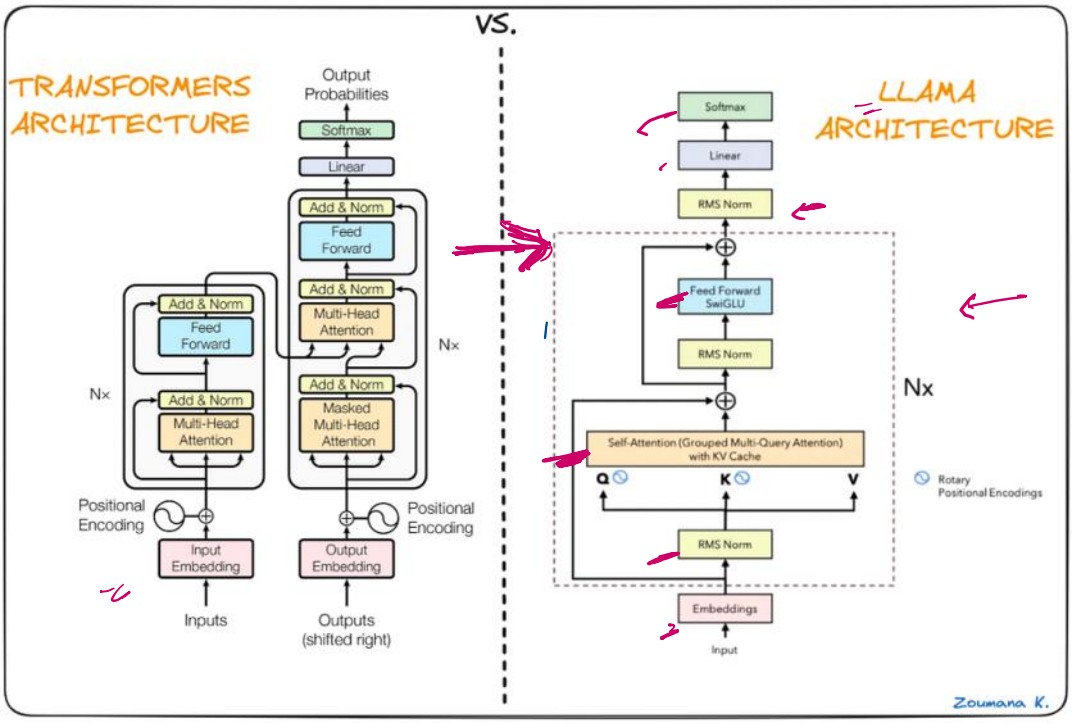
Transfer Learning in Generative AI (Using LLMs)

In Transformer based Architecture:

- **Replacing the output layer of pretrain model**
- Or **unfreezing some layers** of the network and **fine-tuning** it on the new, task-specific dataset



*Illustrative example, exact model architecture may vary slightly



BiT

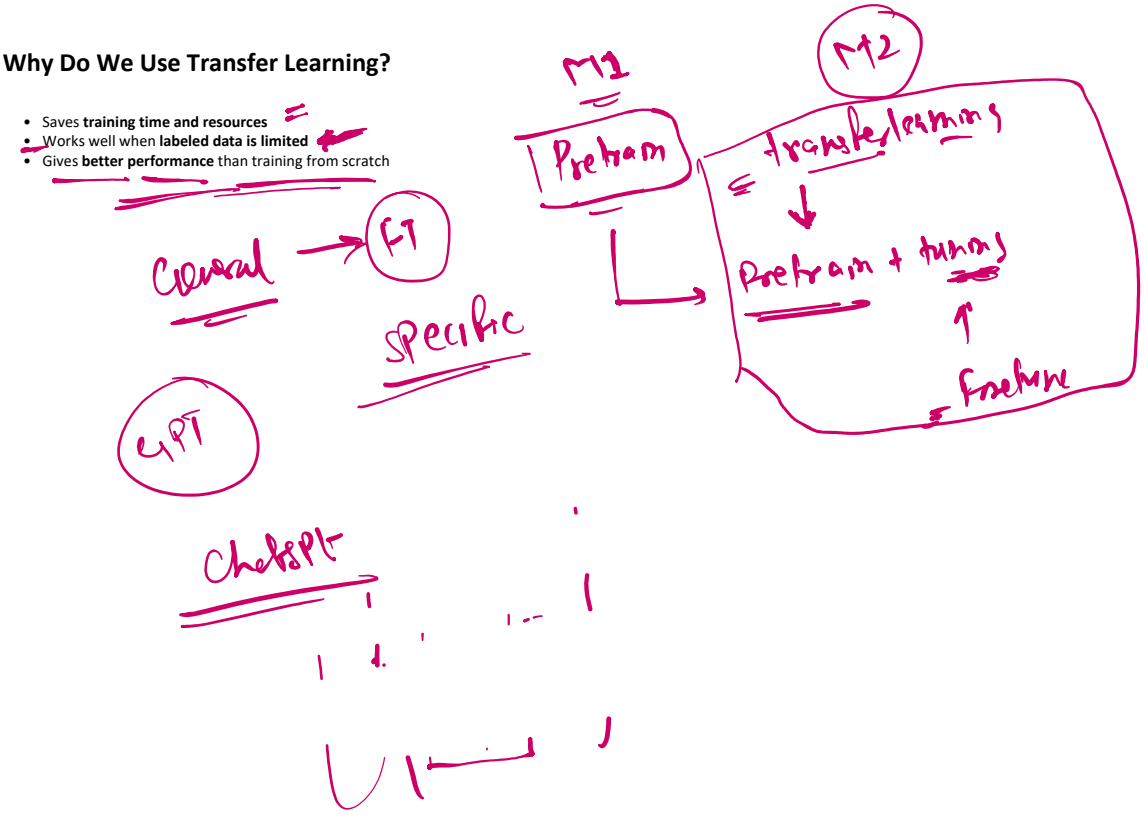
PEFT

SUBSET

Quantized

Why Do We Use Transfer Learning?

- Saves **training time and resources**
- Works well when **labeled data is limited**
- Gives **better performance** than training from scratch



CNN Finetuning

Wednesday, April 23, 2025 6:54 PM

Correct: https://colab.research.google.com/drive/1Dtipjnr4WXnI8ErDeNzzOjzGtgmV-F4r?usp=sharing

https://colab.research.google.com/drive/1U57ZndGQ6WnPdtM2rTdxPARNnTsfRds7?usp=sharing

Bert Finetuning

Monday, April 28, 2025 4:04 PM

https://colab.research.google.com/drive/1Dtipjnr4WXnI8ErDeNzzOjzGtgmV-F4r?usp=sharing

https://colab.research.google.com/drive/1U57ZndGQ6WnPdtM2rTdxPARNnTsfRds7?usp=sharing