# Transfer Learning with Pretrained Models

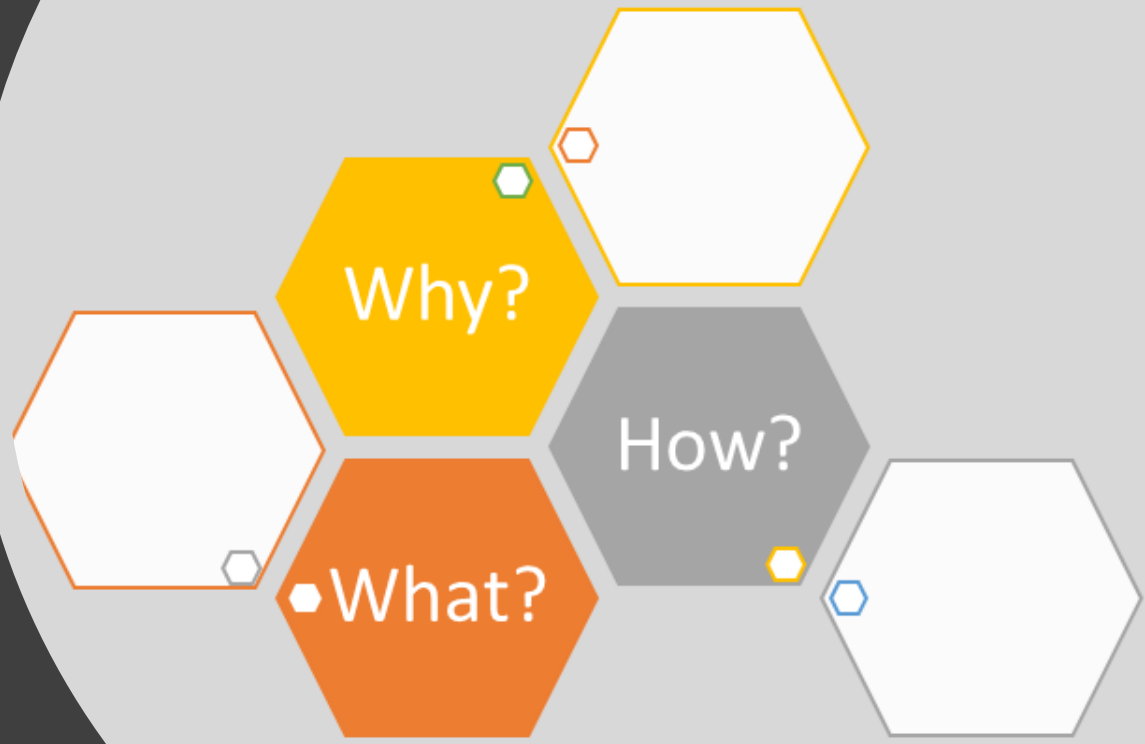
Suvarna Kadam

PhD Scholar, Department of Technology SPPU

Asst. Professor at DYPCOE Akurdi

# Outline

- Motivation (Why to transfer knowledge?)
- Existing Approaches of transfer learning
- Transfer Learning in Deep Models
    1. Train a (source) deep network or find a pretrained model
    2. Repurpose the pretrained model for your needs
    3. Predict with new target model
- Case study for image classification with pretrained model



# Traditional Machine Learning

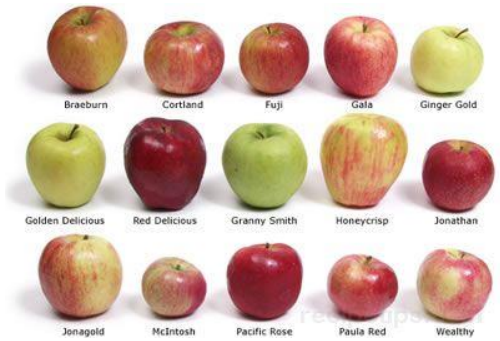
Learning is hard!

Lot of data and compute!

Catastrophic Forgetting

# Why learning is hard for machines?

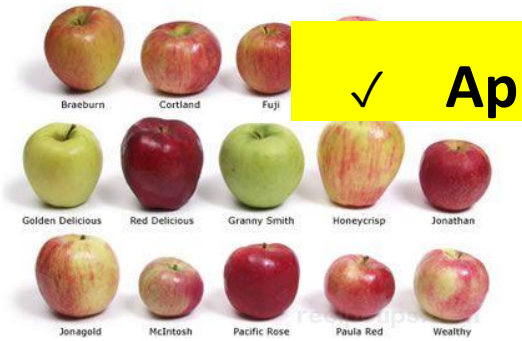

# Learning from observations



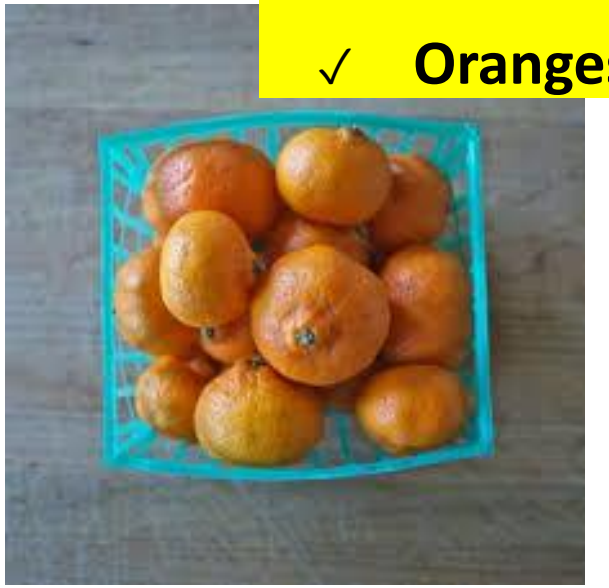

✓ **Apples**

✓ **Oranges**

# Learning from observations




✓ **Apples**



✓ **Oranges**

**Apple/Orange?**

Training

Testing

Traditional Machine Learning

Learning is hard! ✓

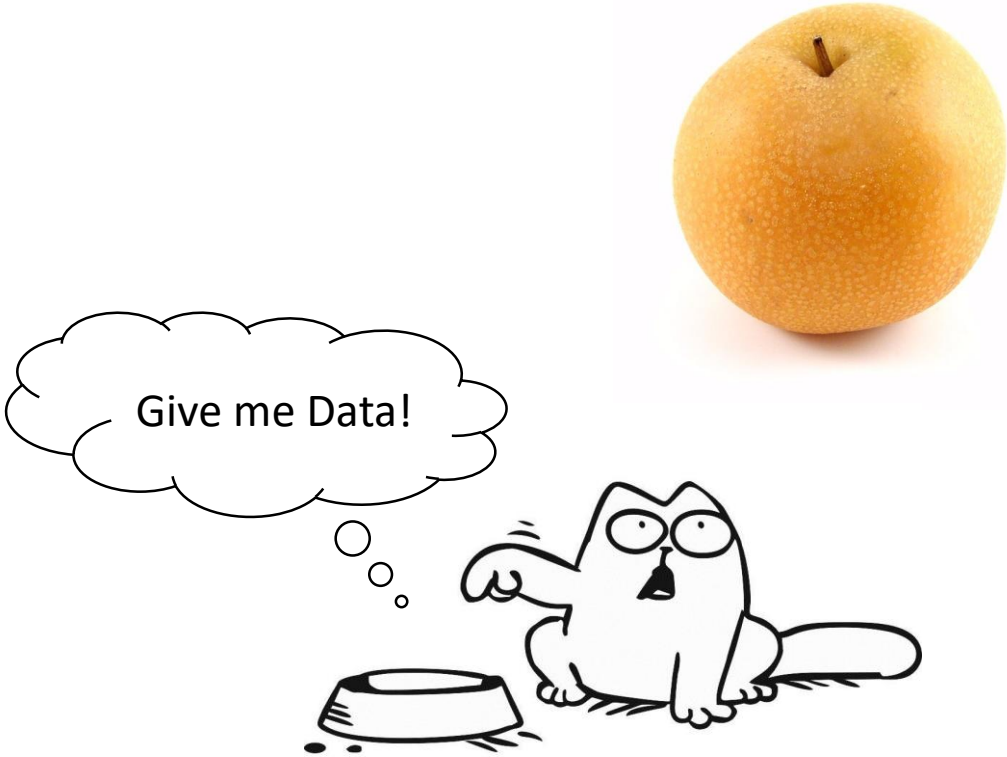
Lot of data and compute!

Catastrophic Forgetting

# Finding Pattern

Purpose of Machine Learning - Find Patterns in Data!

But the problem is ML need **lots of data & Compute** to learn patterns effectively.



Traditional Machine Learning

Learning is hard! ✓

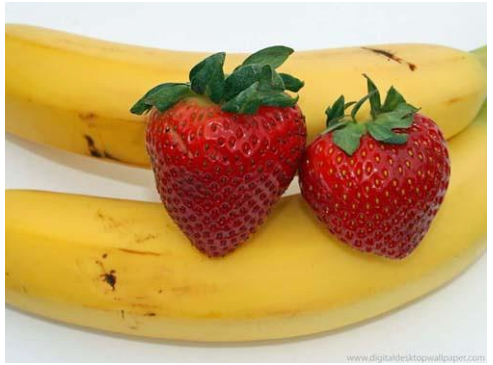
Lot of data and compute! ✓

Catastrophic Forgetting

# Machines have amnesia!

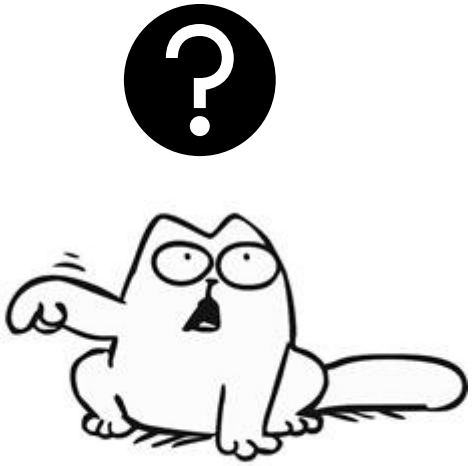
Machines can learn one task at a time.

**Apples ?**
**Orange?**



**Banana ?**
**Straberry?**

But the problem is catashropic forgetting where ML models forget knowledge learned from first task while learning new task!



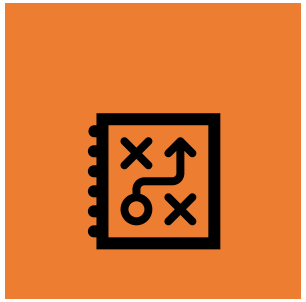
Traditional Machine Learning

Learning is hard! ✓

Lot of data and compute! ✓

Catastrophic Forgetting ✓

# Deep Learning has additional challenges…

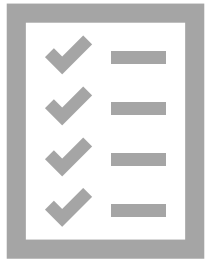


**Training Deep models is hard**

Need lot of labelled data and compute to train

Need  to experiment with architecture

Need to experiment with hyper parameters(learning rate, batch size, momentum, and weight decay)

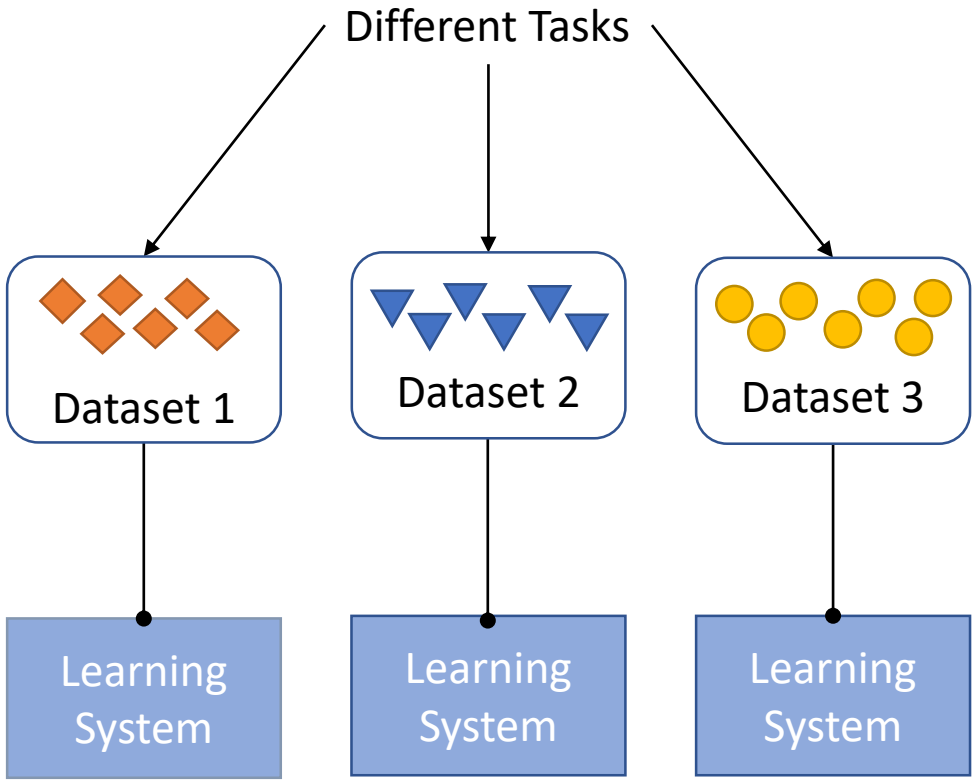


**Deep models are trained for specific task**

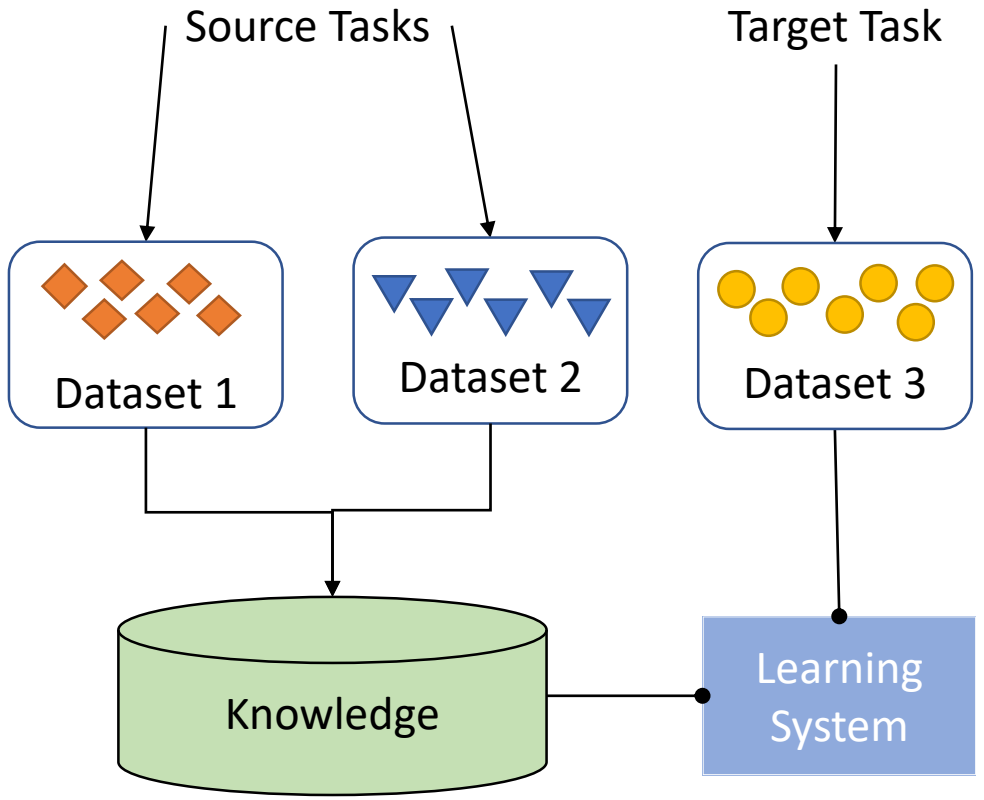
Traditional deep models can’t multi-task

To learn new task, we need to re-train them from scratch!

# Traditional Vs. Transfer Learning



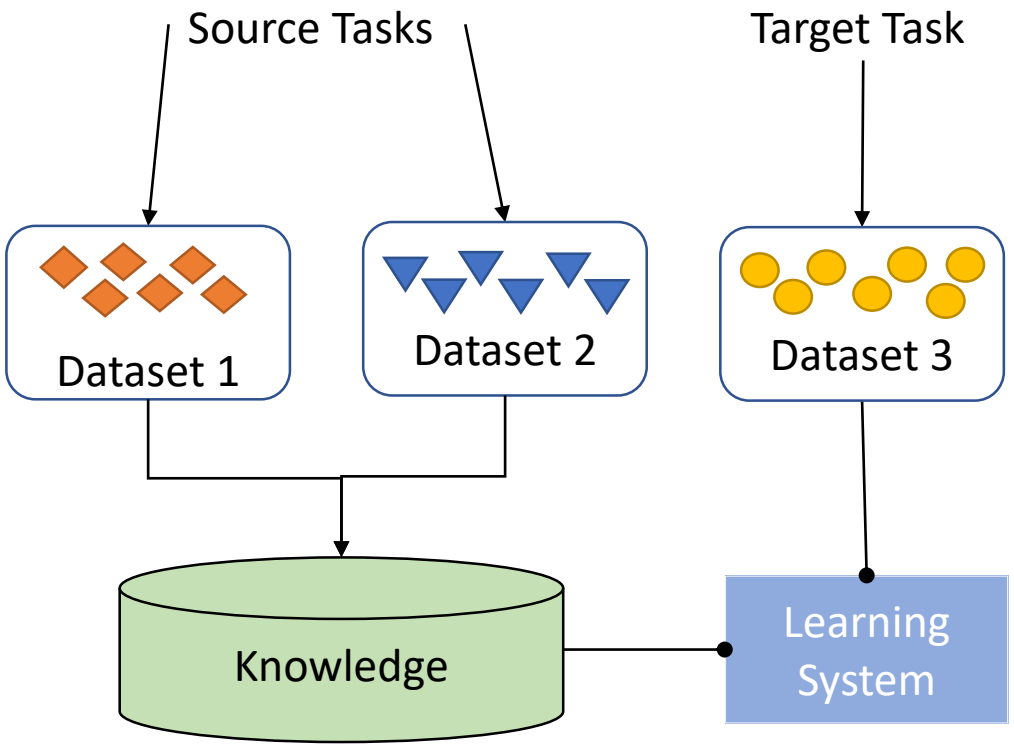

Traditional ML




Transfer Learning

# Transfer Learning

- Utilize the knowledge learned with source dataset(s) to perform a task with target dataset.
- Store Knowledge
  - Knowledge is present in the form of weights of trained model.
  - Weights represent features learned about the (source) tasks
- Transfer Knowledge
  - The learned features are reused when performing a (target) new task.
  - The model need not be trained from scratch




Transfer Learning

# Two Approaches for Deep Transfer Learning

## Train-Source-Models-from-scratch

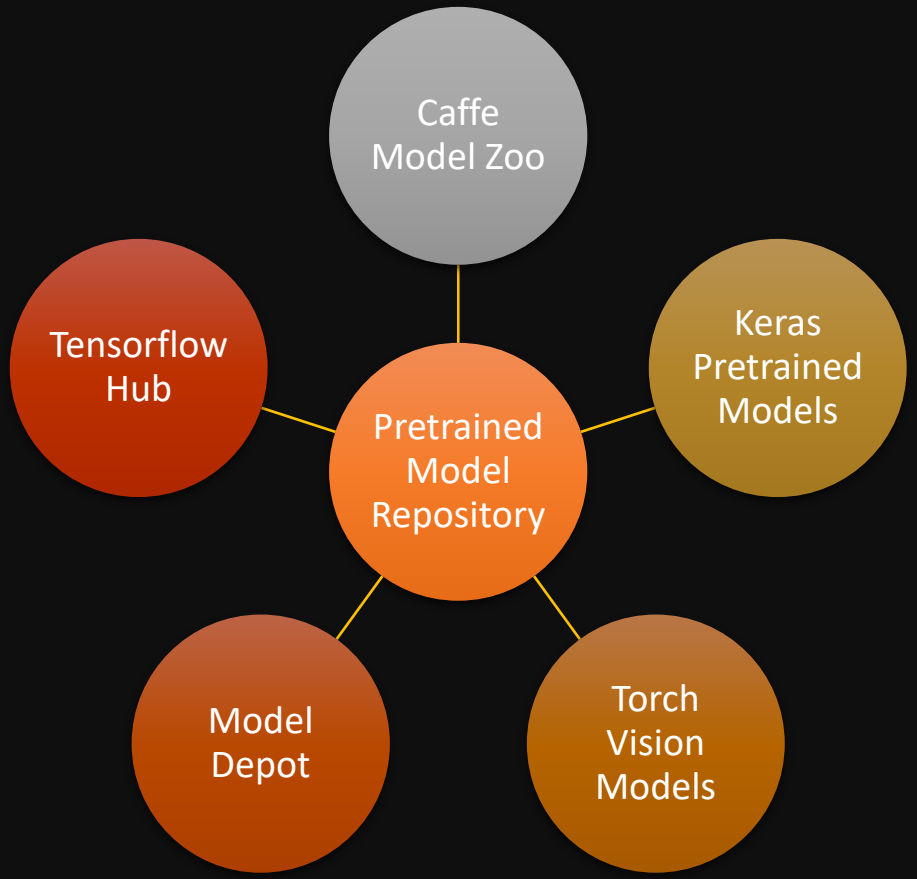
1. **Select Source Task**. Select Source task with an abundance of data
2. **Train Source Model(s)**. Train model rigorously to ensure that features are learned
3. **Reuse Model**. The trained model on the source task is used as the starting point for a model on the target task.
4. **Tune Model**. Optionally, the source model may need to be adapted or refined for target task.

## Pre-trained Model Approach

1. **Select Source Model**. A pre-trained source model is chosen from available models.
2. **Reuse Model**. The model pre-trained model can then be used as the starting point for a model on the second task of interest. This may involve using all or parts of the model, depending on the modeling technique used.
3. **Tune Model**. Optionally, the model may need to be adapted or refined on the input-output pair data available for the task of interest.

**Assumption: We have a repository of several source models.**

# Public Repositories of Pretrained Models




ModelDepot

PyTorch

Caffe2

TensorFlow Hub

# Deep Transfer Learning

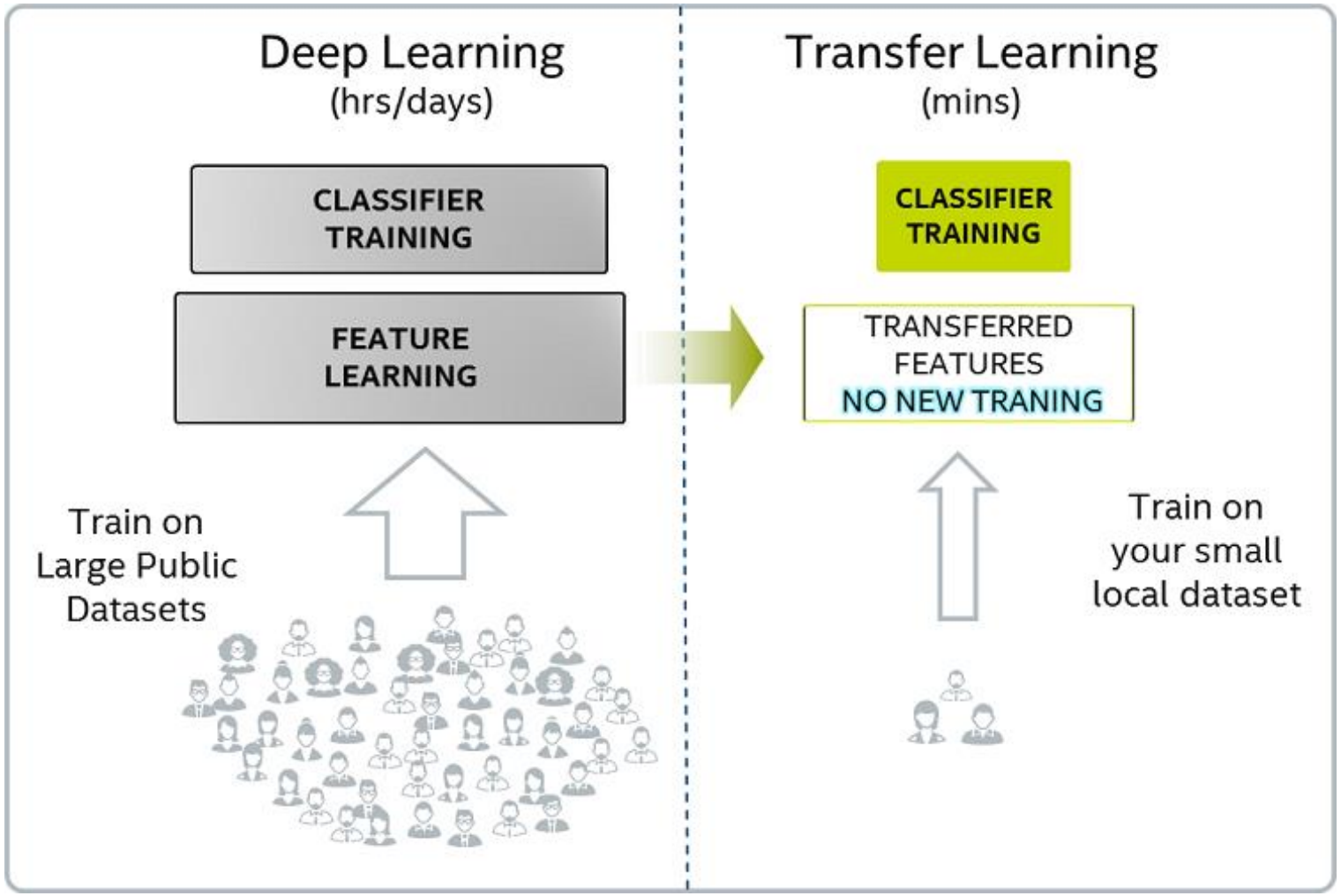



Image Credit: https://www.freecodecamp.org/news/asl-recognition-using-transfer-learning-918ba054c004/

# Repurposing Deep Model for target task

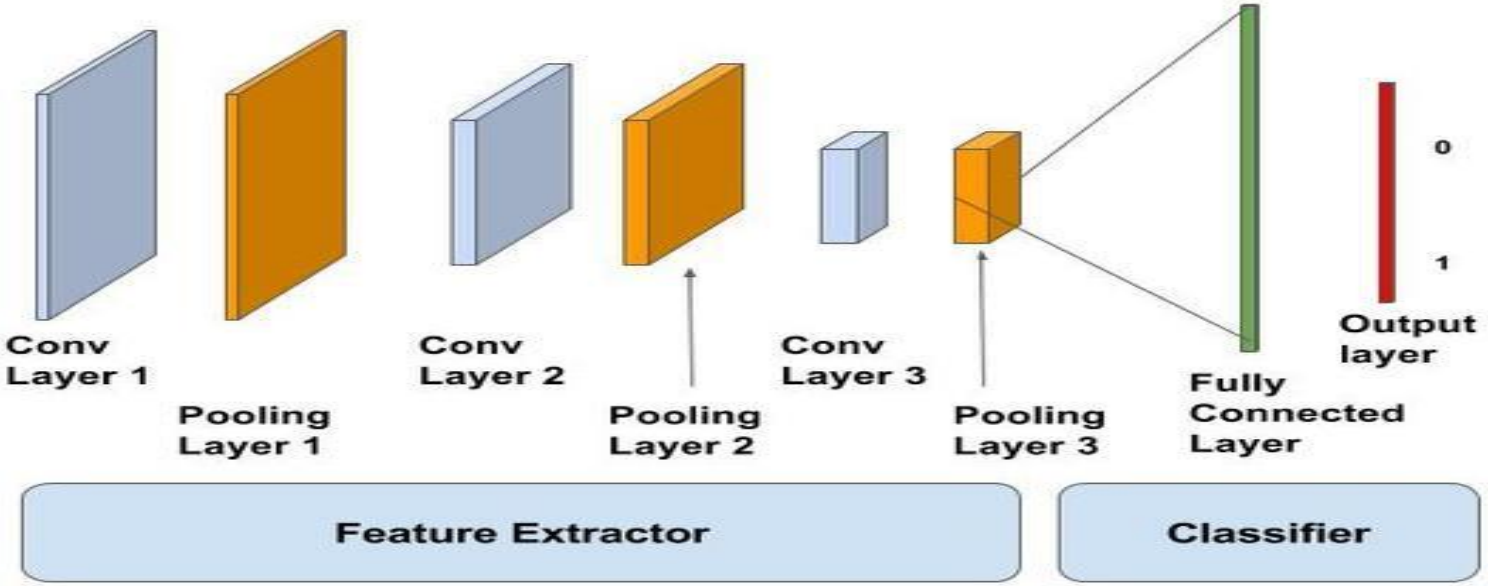
# Image classification with pretrained model

Target Task : Image Classification

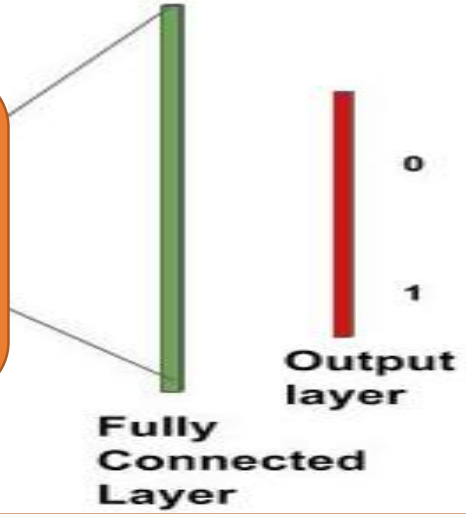
Deep Models : Convolutional Neural Networks

Source Datasets : ImageNet, Open Images,The Street View House Numbers (SVHN)

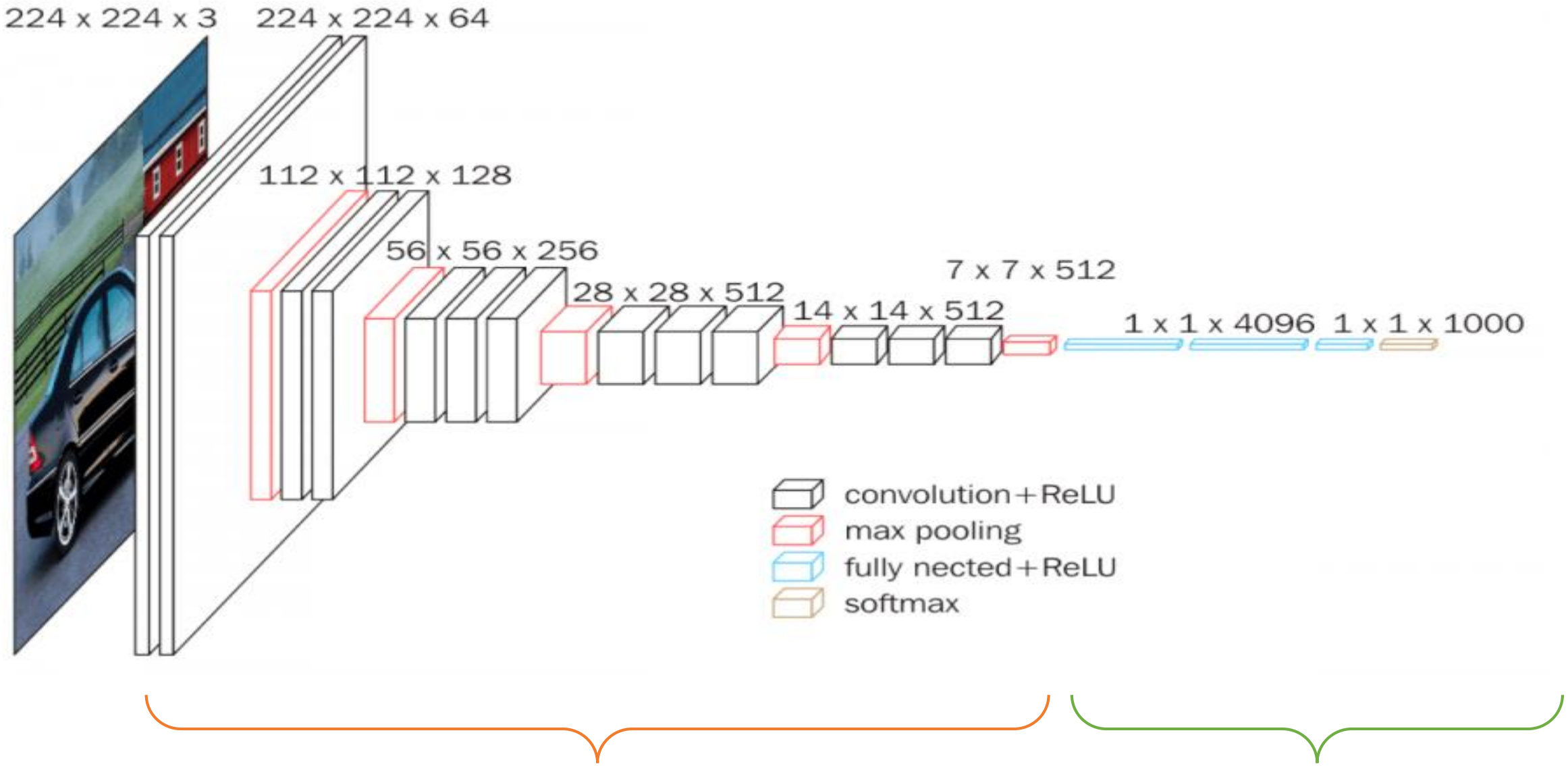
Input




Convolutional Base
With Frozen layers

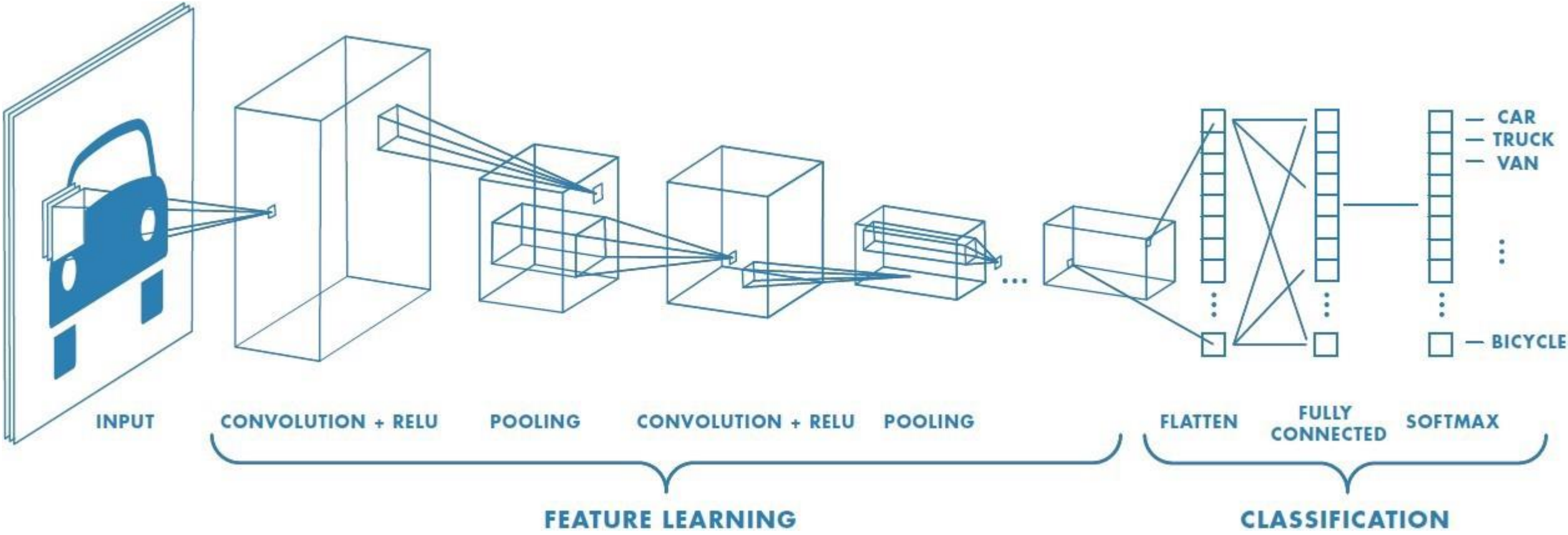



Fine Tuning for Target Task




Frozen Layers

Retrained Layers




**Frozen Layers**

**Finetuning with target dataset**

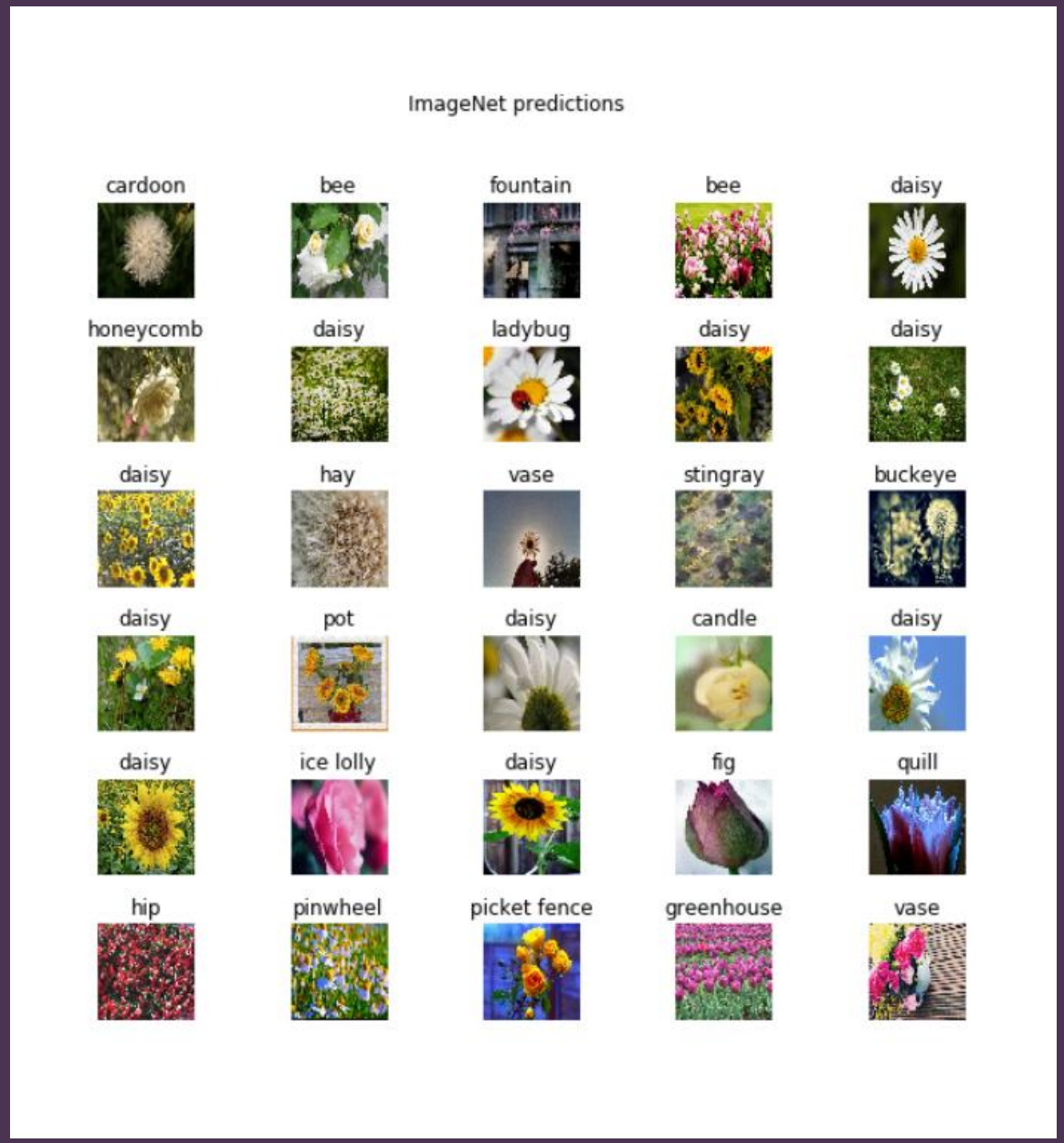
# Image Classification

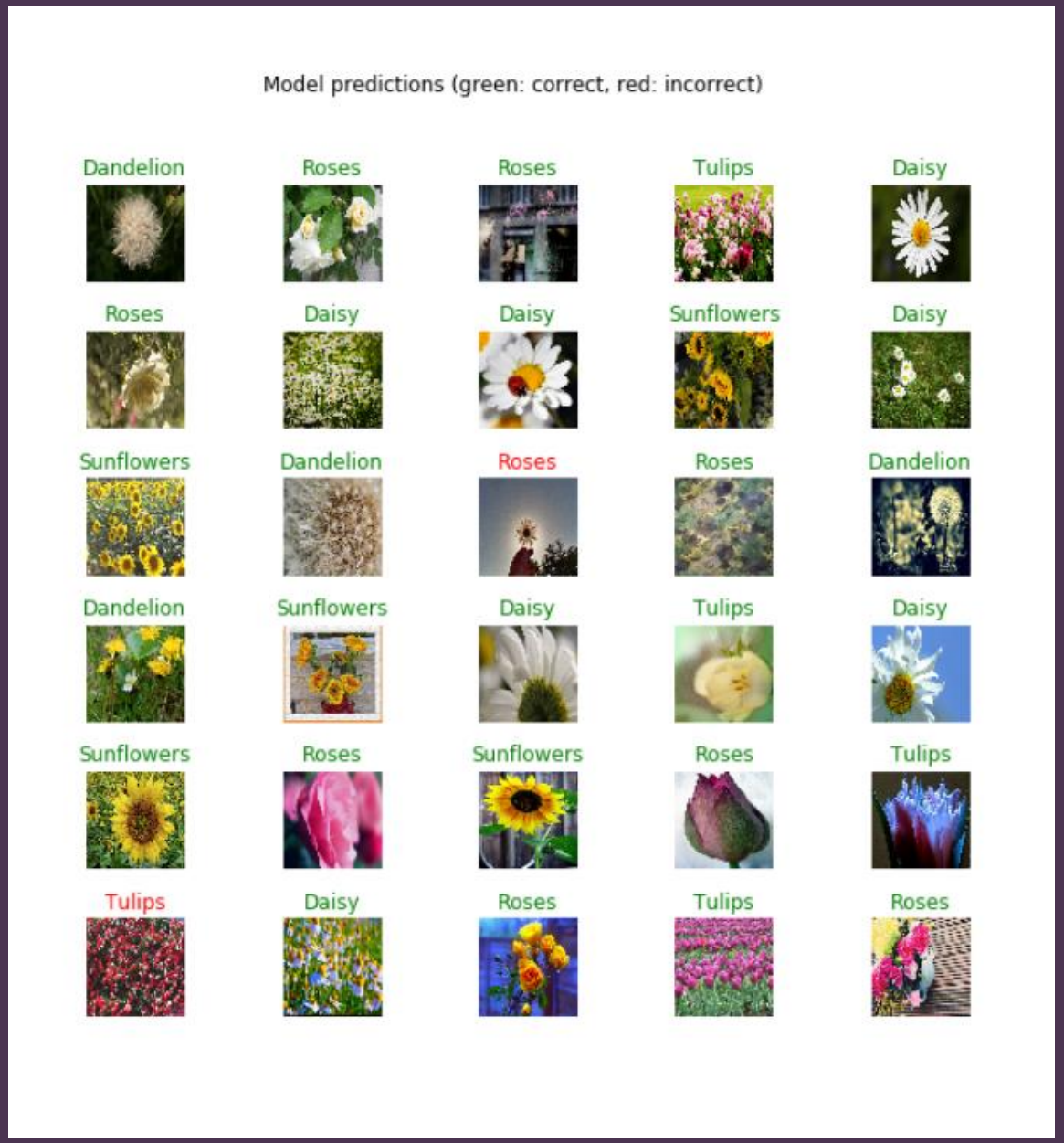
Source Dataset: ImageNet

Target Dataset: Flowers

Task: Classification

https://colab.research.google.com/drive/1fbz7RSho0OjSVysFk0Qu1m8Bpe6MbzZk


ImageNet predictions
cardoon
bee
fountain
bee
daisy
honeycomb
daisy
ladybug
daisy
daisy
daisy
hay
vase
stingray
buckeye
daisy
pot
daisy
candle
daisy
daisy
ice lolly
daisy
fig
quill
hip
pinwheel
picket fence
greenhouse
vase


Model predictions (green: correct, red: incorrect)
Dandelion
Roses
Roses
Tulips
Daisy
Roses
Daisy
Daisy
Sunflowers
Daisy
Sunflowers
Dandelion
Roses
Roses
Dandelion
Dandelion
Sunflowers
Daisy
Tulips
Daisy
Sunflowers
Roses
Sunflowers
Roses
Tulips
Tulips
Daisy
Roses
Tulips
Roses

# Text Classification

Source Dataset: Google News

Target Dataset: imdb

Task: Text Classification (Sentiment Analysis)

https://colab.research.google.com/drive/14IzmGkNTTwCpzf9UNyaVDXK6tE3jcWqQ

# Summary

Transfer Learning with pretrained models *democratises* Deep Learning to some extent

Anyone with small dataset and low compute can benefit the heavy duty Deep learning machinery

Applications range across modalities – Image, Text and Speech

# Thank You!