Transfer Learning

Introduction

What is Transfer Learning?

Introduction

What is Transfer Learning?

- Transfer learning is a powerful technique used in Deep Learning
    - By harnessing the ability to **reuse existing models and their knowledge** on new problems, transfer learning has opened doors to training deep neural networks even with limited data.
- This breakthrough is especially significant in data science, where practical scenarios often need more labeled data.

Introduction

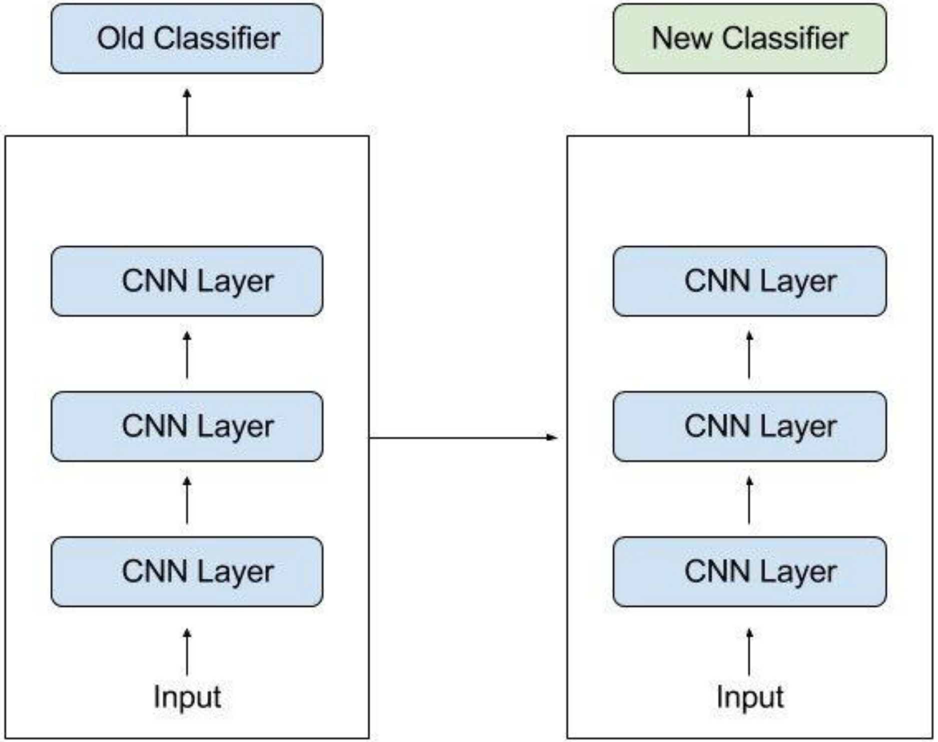
What is Transfer Learning?

- The **reuse** of a pre-trained model on a new problem is known as transfer learning
  - A machine uses the knowledge learned from a prior assignment to increase prediction about a new task.

Introduction

What is Transfer Learning?

- The knowledge of an already trained machine learning model is **transferred** to a **different but closely linked problem** throughout transfer learning.
    - *For example, if you trained a simple classifier to predict whether an image contains a backpack, you could use the model's training knowledge to identify other objects such as sunglasses.*

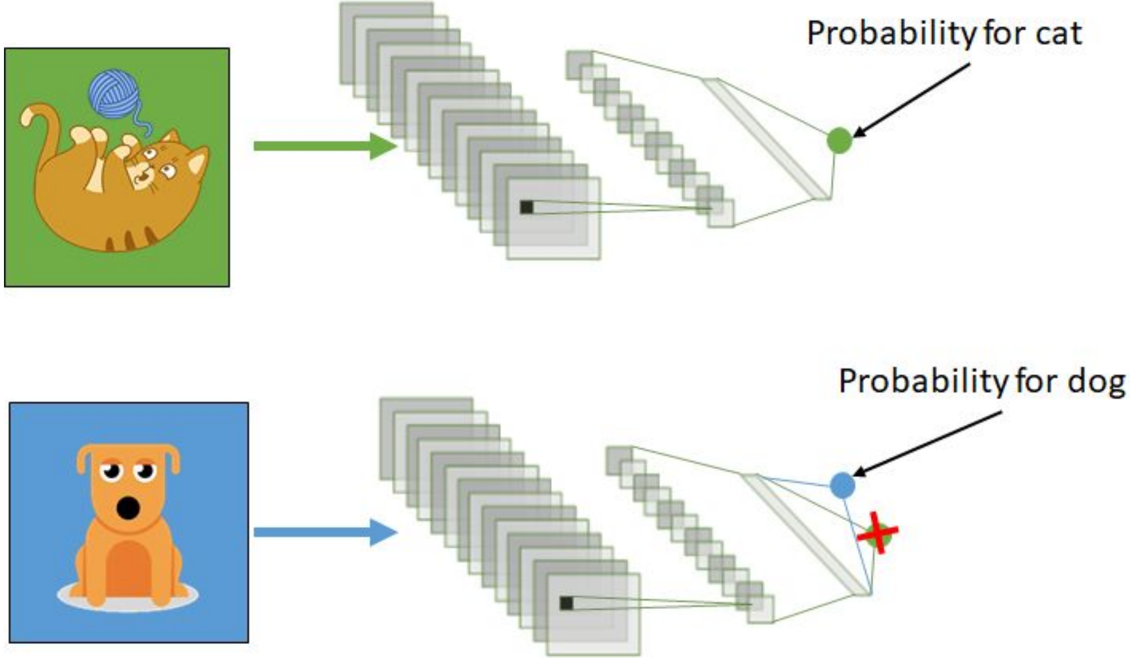

Old Classifier
New Classifier
CNN Layer
CNN Layer
CNN Layer
Input
CNN Layer
CNN Layer
CNN Layer
Input

Introduction

- With transfer learning, we basically try to use what we’ve learned in one task to understand the concepts in another.
    - **weights are being automatically being shifted to a network performing “task A” from a network that perform a new “task B.”**

How Transfer Learning Works?

- **In computer vision, neural networks typically aim to detect**
  - **edges in the first layer,**
  - **forms in the middle layer,**
  - **and task-specific features in the latter layers.**
- The early and central layers are employed in transfer learning, and the latter layers are only retrained.
- It makes use of the labelled data from the task it was trained on.


Probability for cat
Probability for dog

How Transfer Learning Works?

- Let’s return to the example of a model that has been intended to identify a backpack in an image and will now be used to detect sunglasses.
- Because the model has trained to recognise objects in the earlier levels, we will simply **retrain the subsequent layers to understand what distinguishes sunglasses from other objects.**

Why Should You Use Transfer Learning?

- Transfer learning offers a number of advantages,
  - the most important of which are reduced training time,
  - improved neural network performance (in most circumstances),
  - and the absence of a large amount of data.

- To train a neural model from scratch, a lot of data is typically needed, but access to that data isn't always possible – this is when transfer learning comes in handy.

Why Should You Use Transfer Learning?

Because the model has already been **pre-trained**, a good machine learning model can be generated with fairly **little training data using transfer learning**.

This is especially useful in natural language processing, where huge labelled datasets require a lot of expert knowledge.

Types Transfer Learning

Transductive Transfer Learning

- Transductive transfer learning involves transferring knowledge from a specific source domain to a different but related target domain, with the primary goal of the target domain.
- It is particularly useful when labeled data from the target domain is scarce.
- Transductive transfer learning requires the model to make predictions on target data using previously acquired knowledge. Since target data is mathematically similar to the source data, the model finds patterns and functions more quickly.

Types Transfer Learning

Transductive Transfer Learning

- For example, consider adapting a sentiment analysis model trained on product reviews to analyze movie reviews.
  - The source domain (product reviews) and the target domain (movie reviews) differ in context and specifics, but they share similarities in structure and language use.
  - The model quickly learns to apply its understanding of sentiment from the product domain to the film domain.

Types Transfer Learning

Inductive Transfer Learning

- Inductive transfer learning occurs when the source and target domains coincide, but the tasks the model must complete are different.
- The pre-trained model already knows the source data and trains faster for new tasks.
- An example of inductive transfer learning is in natural language processing (NLP).
    - Models are pre-trained on a large corpus of text and then fine-tuned using inductive transfer for specific tasks like sentiment analysis.
    - Similarly, computer vision models like VGG are pre-trained on large image datasets and then optimized for object detection.

Types Transfer Learning

Unsupervised Transfer Learning

- Unsupervised transfer learning uses a similar strategy to inductive transfer learning to develop new skills.
- However, this form of transfer learning is used when only unlabeled data is available in both the source and target domains.
- The model learns common features from unlabeled data to generalize more accurately when tasked with a specific activity.
- This method is useful when obtaining labeled source data is difficult or expensive.

Types Transfer Learning

Unsupervised Transfer Learning

- For example, consider the task of identifying different types of motorcycles in traffic images.
    - Initially, the model is trained on a large set of unlabeled vehicle images. In this case, the model independently determines similarities and distinctive features among different types of vehicles like cars, buses, and motorcycles.
    - Subsequently, the model is introduced to a small, specific set of motorcycle images.
    - The model's performance significantly improves compared to before.

Steps to Use Transfer Learning

- When:
  - we don’t have enough annotated data to train our model with and there is a pre-trained model that has been trained on similar data and tasks.
- How:
  - If you have the original model, you might simply restore it and retrain some layers for your job.

Steps to Use Transfer Learning

- When:
  - we don’t have enough annotated data to train our model with and there is a pre-trained model that has been trained on similar data and tasks.
- How:
  - If you have the original model, you might simply restore it and retrain some layers for your job.

- Transfer learning only works if the features learnt in the first task are general, meaning they can be applied to another activity. Furthermore, the model’s input must be the same size as it was when it was first trained.

Steps to Use Transfer Learning

There are three main steps to fine-tune a machine learning model for a new task:

1. Selection of a Pre-trained Model
2. Configuration of Pre-trained Models
3. Introducing New Layers

Steps to Use Transfer Learning

Selection of a Pre-trained Model

- Firstly, select a pre-trained model with previous knowledge or skills for a related task.
- A useful context for choosing a suitable model involves determining the original task of each model.
- Understanding the original tasks performed by the model can help find one that transfers most effectively to a new task.

Steps to Use Transfer Learning

Configuration of Pre-trained Models

- After selecting the source model, configure it to transfer knowledge to a model to complete the related task. There are two main methods to do this:

    1. **Freezing Pre-trained Layers**

    - Layers are the building blocks of neural networks. Each layer consists of a set of neurons and performs specific transformations on input data.
    - Weights are the parameters used by the network for decision-making.
    - By freezing the weights of pre-trained layers, you keep them fixed, preserving the knowledge the deep learning model has gained from the source task.

Steps to Use Transfer Learning

Configuration of Pre-trained Models

- After selecting the source model, configure it to transfer knowledge to a model to complete the related task. There are two main methods to do this:

    2. **Removing the Last Layer**

    - In some use cases, it's possible to remove the last layers of the pre-trained model.
    - In most ML architectures, the last layers are task-specific.
        - Removing these final layers allows reconfiguring the model for the new task requirements.

Steps to Use Transfer Learning

Introducing New Layers

- Introducing new layers in addition to the pre-trained model helps adapt to the specialized nature of the new task.
- The new layers fine-tune the model to the nuances and functions of the new requirement.

Steps to Use Transfer Learning

Introducing New Layers

- Introducing new layers in addition to the pre-trained model helps adapt to the specialized nature of the new task.
- The new layers fine-tune the model to the nuances and functions of the new requirement.
- ....
- And now you just need to train your NN

Steps to Use Transfer Learning

Introducing New Layers

- Introducing new layers in addition to the pre-trained model helps adapt to the specialized nature of the new task.
- The new layers fine-tune the model to the nuances and functions of the new requirement.
- ….
- And now you just need to train your NN

Knowledge Distillation

Introduction

What is Distillation Learning?

Introduction

What is Distillation Learning and Why do we need this?

- Neural models in recent years have been successful in almost every field including extremely complex problem statements.
- However, **these models are huge in size**, with **millions** (and **billions**) of **parameters**, and thus **cannot be deployed** on **embedded or mobile devices with limited computational resources**.
- Knowledge distillation was introduced by the Godfather of AI, Geoffrey Hinton, and his two co-workers at Google, Oriol Vinyals and Jeff Dean in 2015.

Introduction

What is Distillation Learning?

It is also called:

- **Teacher-student (T-S) learning**
- **Model Compression learning**

Introduction

What is Distillation Learning?

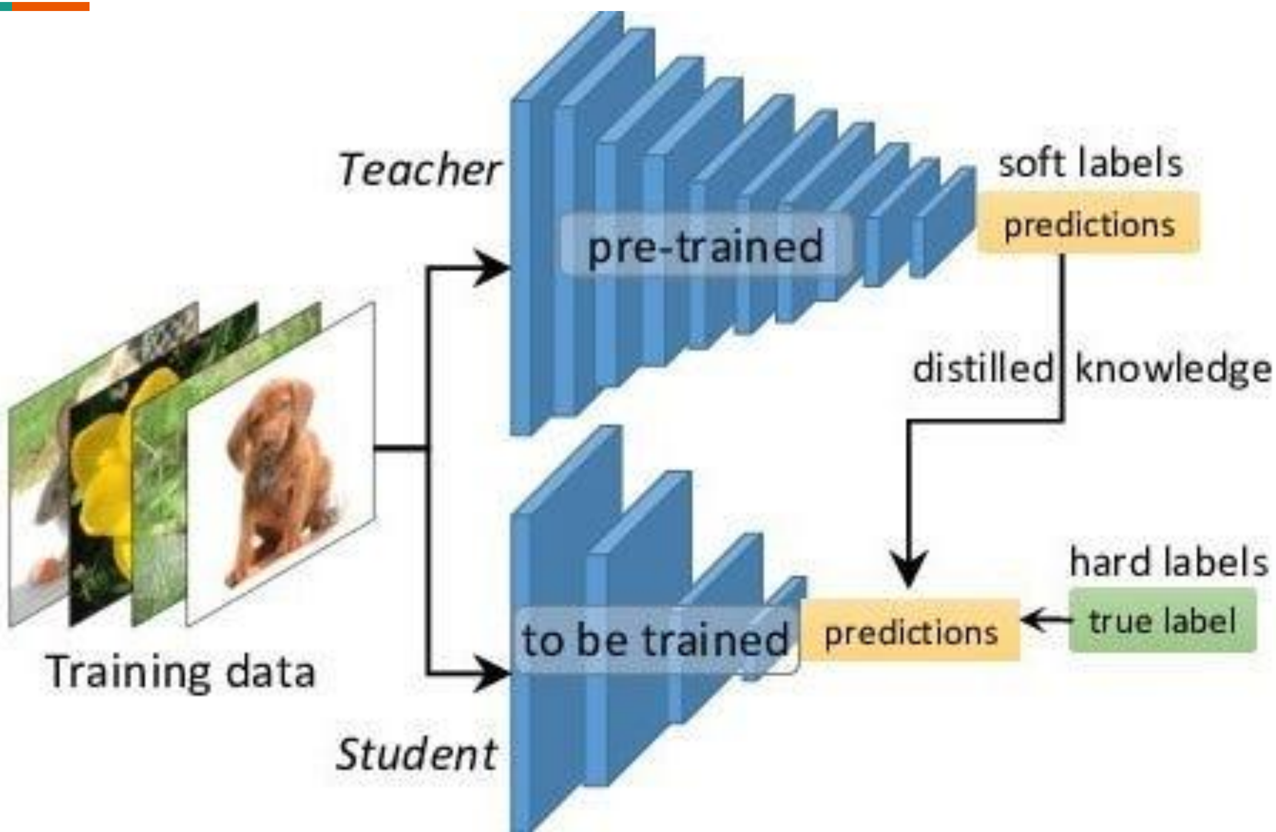
- Distillation learning is a type of machine learning approach that involves **using a small, efficient model** (the "**student**" model) to **learn from** a **larger**, more **complex model** (the "**teacher**" model).
    - The student model is trained to replicate the output of the teacher model as closely as possible, using the output of the teacher model as a supervision signal.

Introduction

- **The goal of distillation learning** is to **transfer knowledge** from the teacher model to the student model, so that the student model can perform the same task as the teacher model but with a smaller size and faster inference speed.

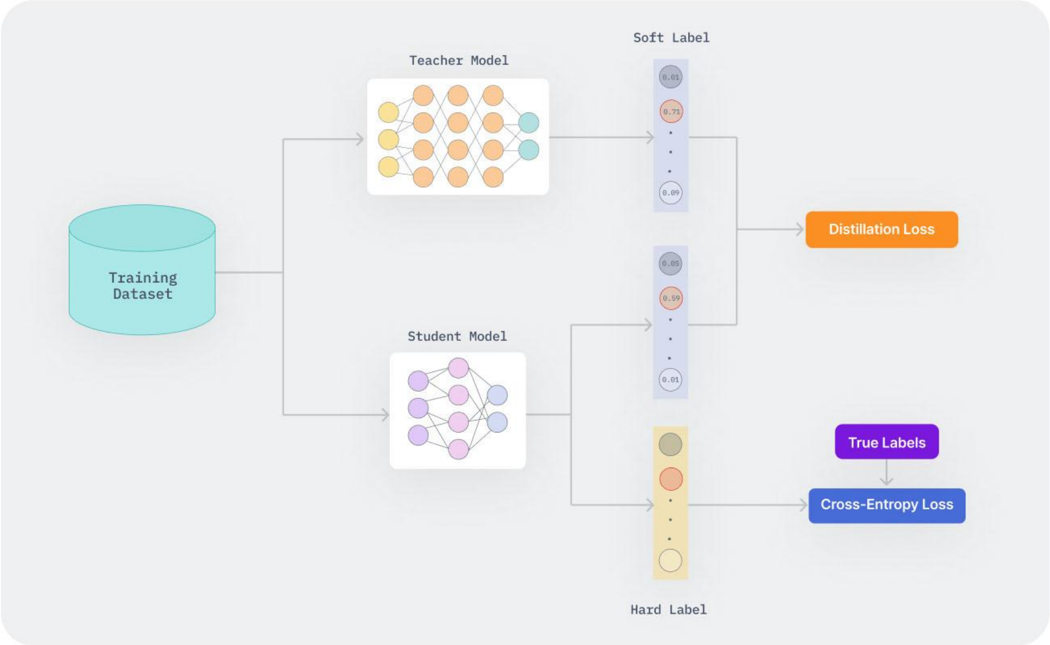
What is the Teacher Model?

- It is typically a neural network that has been trained on a large dataset and has achieved good performance on a particular task (e.g., image classification).
- The teacher model is used to guide the training of the student model by providing a supervision signal.


Teacher
pre-trained
soft labels
predictions
distilled knowledge
Training data
to be trained
predictions
hard labels
true label
Student

What is the Student Model?

- The student model is trained to mimic the output of the teacher model on a given dataset, rather than training the student model to directly optimize its parameters on the dataset.
- The student model is typically trained using a supervised learning approach, with the output of the teacher model as the target labels. The goal is to improve the performance of the student model while reducing its computational cost and memory requirements.


Soft Label
Teacher Model
0.01
0.71
0.09
Training Dataset
Distillation Loss
0.05
0.59
0.01
Student Model
True Labels
Cross-Entropy Loss
Hard Label

Types of Distillation Learning

1. **Model distillation**: This is the most common type of distillation learning, and it involves using a small neural network (the student model) to learn from a larger neural network (the teacher model). The student model is trained to replicate the output of the teacher model on a particular task (e.g., image classification), using the output of the teacher model as a supervision signal.
    a. **Softmax distillation:** This is the most commonly used method of knowledge distillation, which trains the student model to replicate the softmax output of the teacher model. The student model is trained to match the probability distribution generated by the teacher model, rather than the one-hot encoded target labels.

Types of Distillation Learning

1. **Model distillation**: This is the most common type of distillation learning, and it involves using a small neural network (the student model) to learn from a larger neural network (the teacher model). The student model is trained to replicate the output of the teacher model on a particular task (e.g., image classification), using the output of the teacher model as a supervision signal.

    b. **Hard label distillation**: This method trains the student model to replicate the hard labels (i.e., one-hot encoded target labels) predicted by the teacher model. This is less commonly used than softmax distillation because it is more sensitive to errors in the teacher model.

Types of Distillation Learning

2. **Knowledge distillation**: This type of distillation learning involves transferring knowledge from a teacher model to a student model in a more general sense, rather than just replicating the output of the teacher model. For example, a student model might be trained to mimic the internal representation or decision-making process of the teacher model, rather than just its output.

Types of Distillation Learning

3. **Multi-task distillation**: In this type of distillation learning, a student model is trained to perform multiple tasks simultaneously, using a teacher model that has been trained on each task individually. This allows the student model to learn from the teacher model and improve its performance on multiple tasks at once.

Types of Distillation Learning

4. **Transfer distillation:** This type of distillation learning involves transferring knowledge from a teacher model that has been trained on one task or dataset to a student model that is being trained on a different task or dataset. This can be useful for transferring knowledge from a large, pre-trained model to a new task or domain.

Types of Distillation Learning

5. **FitNet:** This method trains the student model to replicate the intermediate activations of the teacher model, in addition to the softmax output. This can help to preserve more detailed information from the teacher model and improve the performance of the student model.

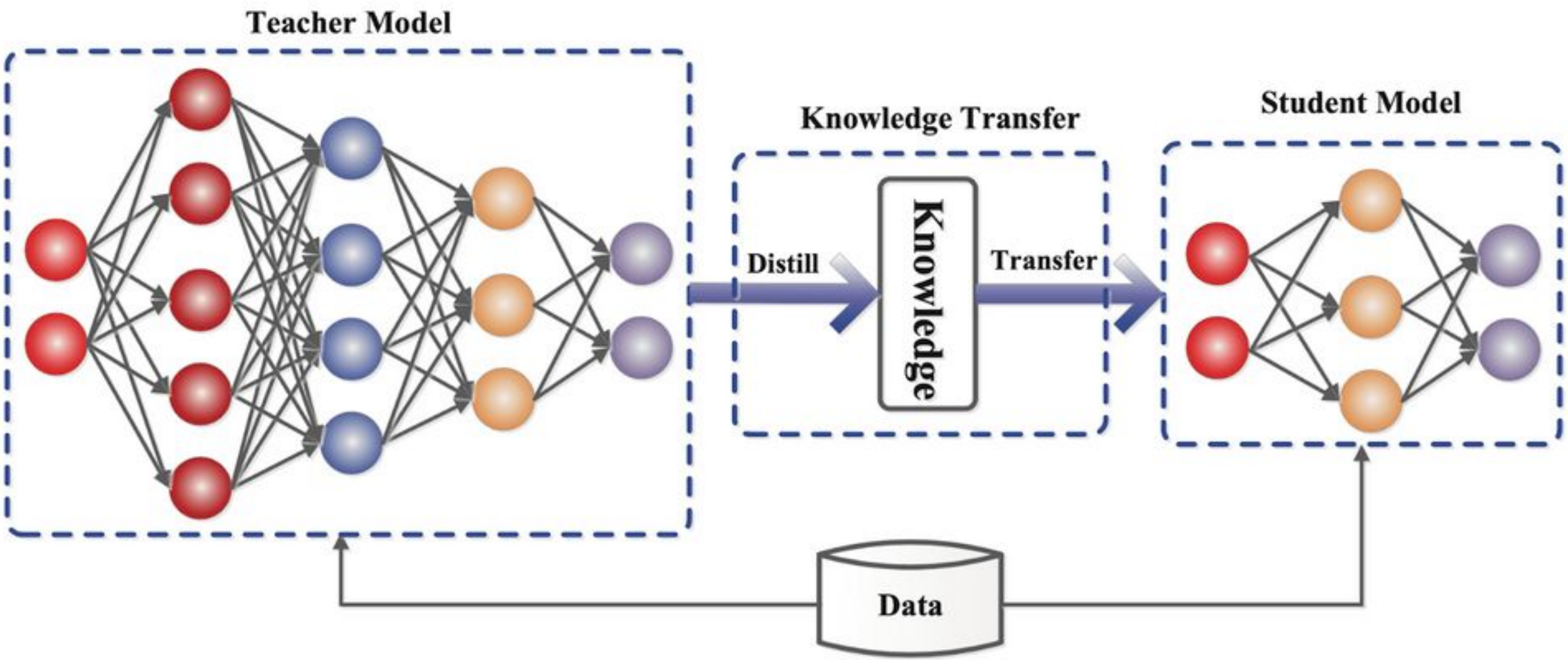
Types of Distillation Learning

6. **Attention transfer**: This method transfers the attention mechanism learned by the teacher model to the student model, which helps the student model to focus on the most important features of the input data.

Applications of Distillation Learning

- Model compression
- Transfer learning
- Ensemble learning
- Multi-task learning
- Language models
- Computer vision

- Anomaly detection
- Generative models
- Medical imaging
- Electronic health records (EHRs)
- Medical decision making
- Clinical trial


Teacher Model
Knowledge Transfer
Student Model
Distill
Knowledge
Transfer
Data

Limitations of Distillation Learning

- Limited generalization
- Dependence on the teacher model (Requires a good teacher model)
- Requires large amounts of data
- Computational cost
- Limited flexibility
- Requires fine-tuning

Credits

- https://medium.com/womenintechnology/distillation-learning-tensorflow-tutorial-478d743d2450
- https://www.analyticsvidhya.com/blog/2021/10/understanding-transfer-learning-for-deep-learning/