

Transfer Learning

Transfer Learning is a favored approach in many deep learning projects, especially useful when data is scarce or when collecting and labeling data requires significant time and resources.

The core idea of transfer learning is to take a model that has been pre-trained on a large dataset (e.g., ImageNet) and further fine-tune it for a new, specific task. Through this process, the model can acquire knowledge for the new task more quickly and effectively.

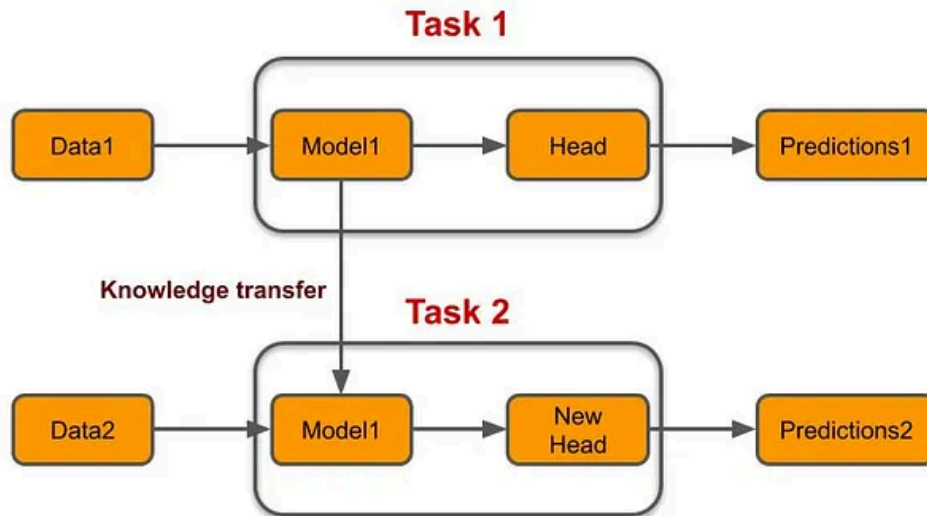
The process of transfer learning typically involves the following steps:

- **Selecting a Pre-trained Model:** Choose a model that has been pre-trained on a large dataset. This model has already learned a lot of features, which can be reused as useful knowledge for a new task.
- **Using as a Feature Extractor:** Remove the output layer of the pre-trained model and use the rest of the network as a feature extractor for the new dataset. The features extracted in this way are used to train a classifier for the new task.
- **Fine-tuning:** Perform additional learning on some or all parts of the pre-trained model to adjust it better for the new task. During this process, a low learning rate is typically used to ensure the model retains the useful information it has already learned while adjusting to the new data.

The advantages of transfer learning include:

- **Reduced Data Requirements:** Using a pre-trained model reduces the amount of data needed for the new task.
- **Shorter Learning Time:** Since the model already possesses a significant amount of knowledge, the learning time for the new task is reduced. Improved Performance: Pre-trained models on large datasets have already learned various features, which often leads to better performance on new tasks.

Transfer Learning



Feature Extraction Process

- **Using Convolutional and Pooling Layers:** The convolutional and pooling layers of a pre-trained model are employed to extract high-level features from images. These layers have already learned to identify various patterns, textures, shapes, etc., making them capable of extracting useful features from new image data as well.
- **Replacing and Training the Data Classifier:** The last fully connected layers of the pre-trained model are removed and replaced with new fully connected layers that suit the new task. Only these layers are trained on new data, allowing the model to classify images based on the extracted features.