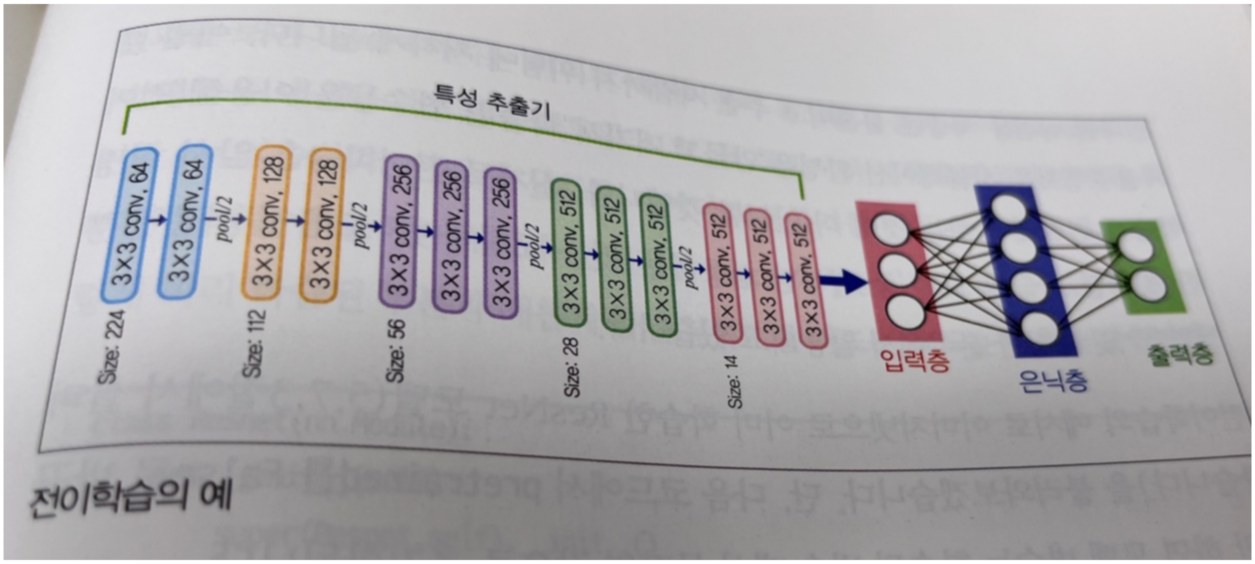
Transfer Learning



Written by: 류채은 Chaeeun Ryu (superbunny38 *at* gmail *dot* com)

< Transfer Learning >

In transfer learning, you can leverage knowledge (features, weights etc) from previously trained models for training newer models and even tackle problems like having less data for the newer task. Transfer learning should enable us to utilize knowledge from previously learned tasks and apply them to newer, related ones. If we have significantly more data for task ***T1***, we may utilize its learning, and generalize this knowledge (features, weights) for task ***T2*** (which has significantly less data). In the case of problems in the computer vision domain, certain low-level features, such as edges, shapes, corners and intensity, can be shared across tasks, and thus enable knowledge transfer among tasks. Also, as we have depicted in the earlier figure, knowledge from an existing task acts as an additional input when learning a new target task.

- **Inductive Transfer learning:** In this scenario, the source and target domains are the same, yet the source and target tasks are different from each other. The algorithms try to utilize the inductive biases of the source domain to help improve the target task. Depending upon whether the source domain contains labeled data or not, this can be further divided into two subcategories, similar to multitask learning and self-taught learning, respectively.

- **Unsupervised Transfer Learning:**This setting is similar to inductive transfer itself, with a focus on unsupervised tasks in the target domain. The source and target domains are similar, but the tasks are different. In this scenario, labeled data is unavailable in either of the domains.
- **Transductive Transfer Learning:**In this scenario, there are similarities between the source and target tasks, but the corresponding domains are different. In this setting, the source domain has a lot of labeled data, while the target domain has none. This can be further classified into subcategories, referring to settings where either the feature spaces are different or the marginal probabilities.

Domain Adaptation

Domain adaption is usually referred to in scenarios where the marginal probabilities between the source and target domains are different, such as $P(X_s) \neq P(X_t)$. There is an inherent shift or drift in the data distribution of the source and target domains that requires tweaks to transfer the learning. For instance, a corpus of movie reviews labeled as positive or negative would be different from a corpus of product-review sentiments. A classifier trained on movie-review sentiment would see a different distribution if utilized to classify product reviews. Thus, domain adaptation techniques are utilized in transfer learning in these scenarios.

Feature Extractor: The part trained in advance.

Advantages:

- Beneficial when faced with shortage of datasets
- Reduced training time