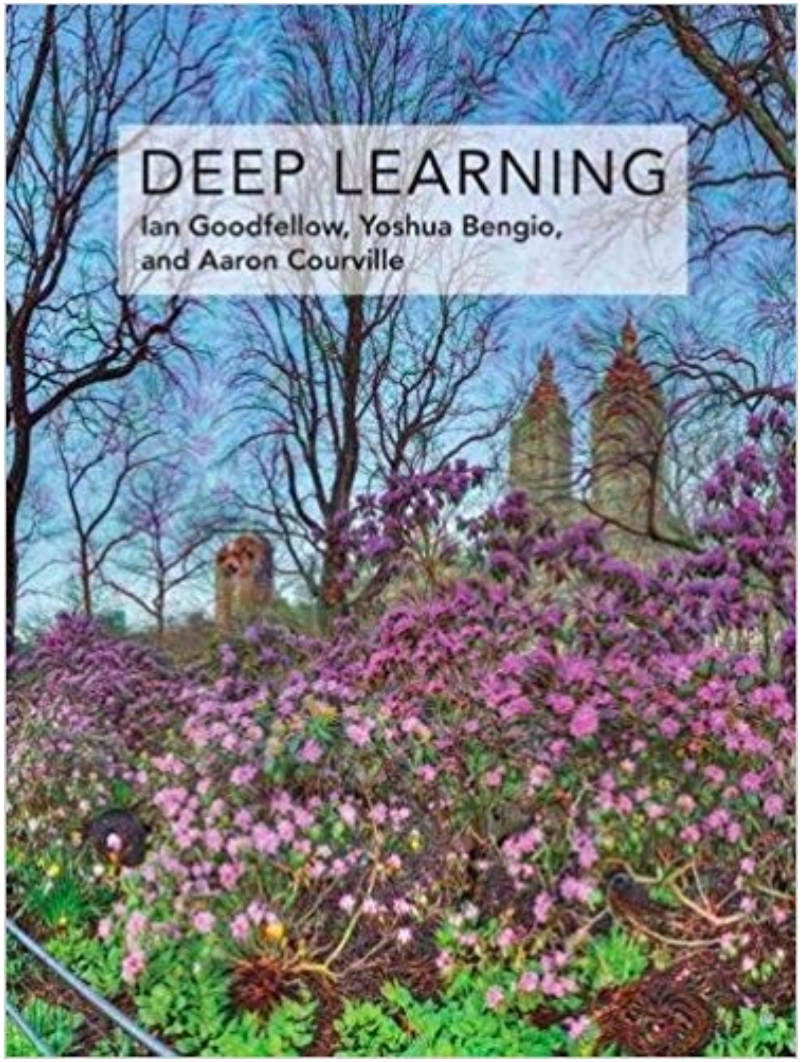
國立陽明交通大學
NATIONAL YANG MING CHIAO TUNG UNIVERSITY

Deep Learning
深度學習
Fall 2023

Transfer Learning
(Chapter 15)

Prof. Chia-Han Lee
李佳翰 教授



- Figure source: Textbook and Internet
- You are encouraged to buy the textbook.
- Please respect the copyright of the textbook. Do not distribute the materials to other people.

Transfer learning and domain adaptation

- Transfer learning and domain adaptation refer to the situation where what has been learned in one setting (e.g., distribution P_1) is exploited to improve generalization in another setting (say, distribution P_2).

- In transfer learning, we assume that many of the factors that explain the variations in P_1 are relevant to the variations that need to be captured for learning P_2.

- This is typically understood in a supervised learning context, where the input is the same but the target may be of a different nature. For example, we may learn about one set of visual categories, such as cats and dogs, in the first setting, then learn about a different set of visual categories, such as ants and wasps, in the second setting.

Transfer learning and domain adaptation

- If there is significantly more data in the first setting (sampled from P_1), then that may help to learn representations that are useful to quickly generalize from only very few examples drawn from P_2.

- Many visual categories share low-level notions of edges and visual shapes, the effects of geometric changes, changes in lighting, and so on.

- In general, transfer learning, multitask learning, and domain adaptation can be achieved via representation learning when there exist features that are useful for the different settings or tasks, corresponding to underlying factors that appear in more than one setting.

Transfer learning and domain adaptation

- Sometimes, however, what is shared among the different tasks is not the semantics of the input but the semantics of the output.

- For example, a speech recognition system needs to produce valid sentences at the output layer, but the earlier layers near the input may need to recognize very different versions of the same phonemes or subphonemic vocalizations depending on which person is speaking.

- In cases like these, it makes more sense to share the upper layers (near the output) of the neural network and have a task-specific preprocessing.

Transfer learning and domain adaptation

- In domain adaptation, the task (and the optimal input-to-output mapping) remains the same between each setting, but the input distribution is slightly different.

- Consider the task of sentiment analysis, which consists of determining whether a comment expresses positive or negative sentiment. Comments come from many categories. A domain adaptation scenario can arise when a sentiment predictor trained on customer reviews of media content, such as books, videos and music, is later used to analyze comments about consumer electronics, such as televisions or smartphones.

- There is an underlying function that tells whether any statement is positive, neutral, or negative, but the vocabulary/style may vary from one domain to another.

Transfer learning and domain adaptation

- A related problem is that of concept drift, which we can view as a form of transfer learning due to gradual changes in the data distribution over time.
- Both concept drift and transfer learning can be viewed as particular forms of multitask learning.
- The objective of multitask learning is to take advantage of data from the first setting to extract information that may be useful when learning or even when directly making predictions in the second setting.
- Using the same representation in both settings allows the representation to benefit from the training data that is available for both tasks.

Transfer learning and domain adaptation

- Two extreme forms of transfer learning are one-shot learning and zero-shot learning, also called zero-data learning. Only one labeled example of the transfer task is given for one-shot learning, while no labeled examples are given at all for the zero-shot learning task.

- One-shot learning is possible because the representation learns to cleanly separate the underlying classes during the first stage. During the transfer learning stage, only one labeled example is needed to infer the label of many possible test examples that all cluster around the same point in representation space.

Transfer learning and domain adaptation

- For zero-shot learning, consider the problem of having a learner read a large collection of text and then solve object recognition problems.

- It may be possible to recognize a specific object class even without having seen an image of that object if the text describes the object well enough. For example, having read that a cat has four legs and pointy ears, the learner might be able to guess that an image is a cat without having seen a cat before.

- Zero-data learning and zero-shot learning are only possible because additional information has been exploited during training.

Transfer learning and domain adaptation

- In machine translation, even though we may not have labeled examples translating word A in language X to word B in language Y, we can generalize and guess a translation for word A because we have learned a distributed representation for words in language X and a distributed representation for words in language Y, then created a link (possibly two-way) relating the two spaces, via training examples consisting of matched pairs of sentences in both languages.

- This transfer will be most successful if the two representations and the relations between them are learned jointly.