| COMP 642 — Machine Learning | Mar 24, 2022 |
| --- | --- |

# Lecture 19: Self-Supervised Learning

*Lecturer: Anshumali Shrivastava* *Scribe By: Danny Lin, Shih-Yao Wang*

**Disclaimer:** *These lecture notes are intended to develop the thought process and intuition in machine learning. The materials are not thoroughly reviewed and can contain errors.*

## 1 Types of Machine Learning[1, 2]

There are different types of learning in the field of machine learning and so far we have covered supervised learning, unsupervised learning, and semi-supervised learning in the previous classes. In this lecture, we first recap the concept of learning mentioned before. Then, we dive into the field of self-supervised learning.

### 1.1 Supervised Learning

In supervised learning, the training data is labelled. As shown in Figure 1, the dataset is composed of inputs along with their corresponding labels. The model is trained on the training dataset and is used to make prediction on unseen data. There are two types of supervised learning: Regression and Classification.

- Regression: predict continuous numerical value.
- Classification: predict discrete class label.

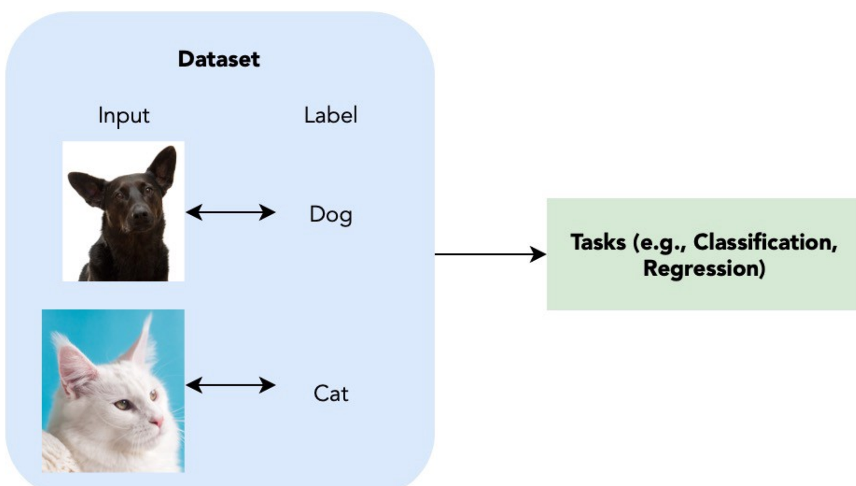


Figure 1: Illustration of Supervised Learning (Image source: reference: [1])

### 1.2 Unsupervised Learning

In contrast to supervised learning, unsupervised learning trains the model on inputs without labels (see Figure 2). The goal is to let model learn the underlying pattern within the dataset. One of the most popular task of unsupervised learning is clustering, which we had discussed in the previous lectures.

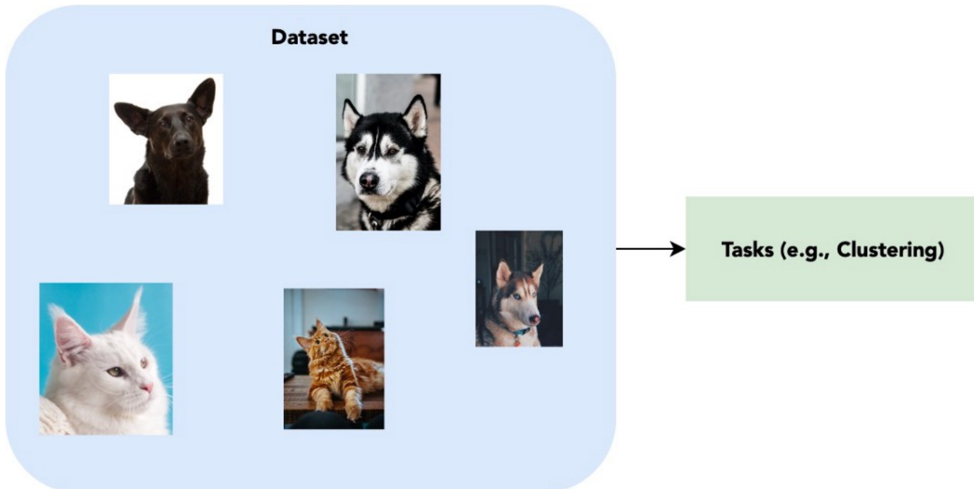


Figure 2: Illustration of Unsupervised Learning (Image source: reference: [1])

### 1.3 Semi-Supervised learning

Semi-supervised learning is the combination of supervised and unsupervised learning. In Figure 3, the dataset is divided into labelled data and unlabelled data. The aim is to use a small amount of labelled data with the majority of data being unlabelled to train our model. One technique called pseudo-labelling comes into play in this scenario. Taking image classification as example, we can follow the steps below to train our model [3].

- Train the classification model with the labelled dataset.
- Use the model to predict/label the unlabelled data. The data labelled with high confidence will be included in the dataset for future training.
- Repeat the above processes until all the data has been used.

One thing we need to pay careful attention is that the initial labelled data may introduce biases during training. These biases could affect pseudo-labelling and thus affect the end result of our model.

### 1.4 Self-Supervised Learning

Self-supervised learning can be viewed as a subset of unsupervised learning. Even though both learning techniques focus on the training of unlabelled data, self-supervised learning solves the problem in a supervised manner (see Figure 4). In other words, the unsupervised learning problems are transformed into supervised learning problems.

## 2 Introduction to Self-Supervised Learning

Figure 5 gives us a great intuition on what self-supervised learning problems are and how we can construct them. Essentially, we could predict any part of the input data from any other part [4]. Thus, we

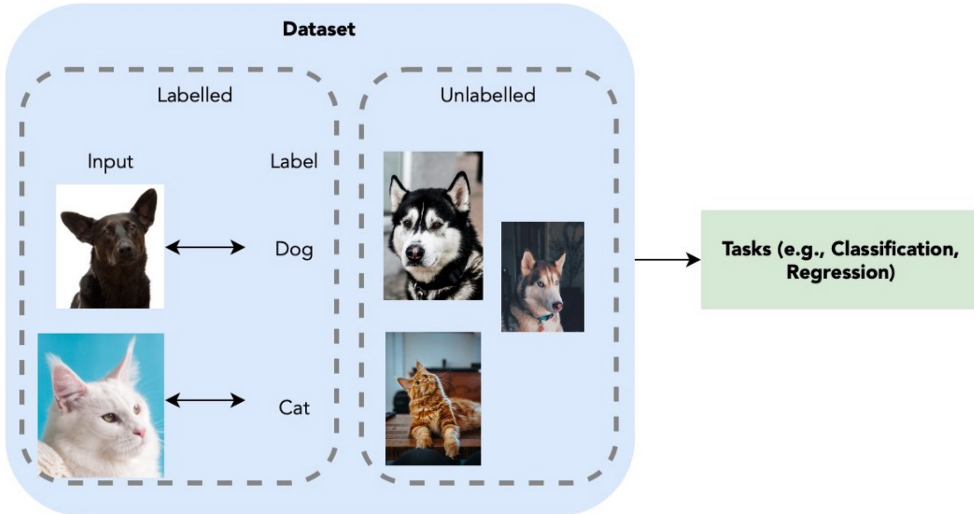


Figure 3: Illustration of Semi-Supervised Learning (Image source: reference: [1])

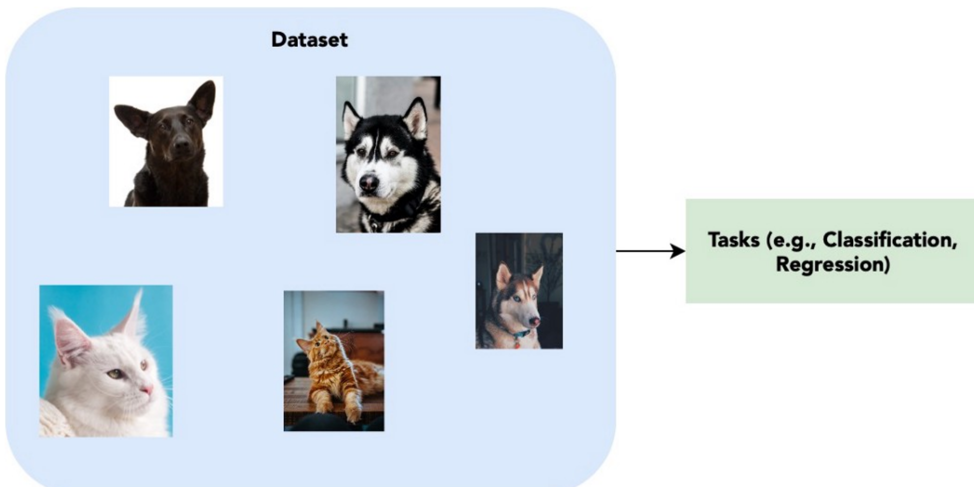


Figure 4: Illustration of Self-Supervised Learning (Image source: reference: [1])

can pretend to predict a part of the input, which we actually have no idea about, with our unlabelled data. Several examples were discussed in the class and we show some of them here. You can refer to reference [4] for more details.

## 2.1 Example: Predicting Image Rotation[5]

In this example, images can be rotated by a multiple of $90°$ ($0°$, $90°$, $180°$, or $270°$). The model is trained to predict the rotation of the input images. Although the images are unlabelled, we know the degree of rotation applied to them. Thus, we can view this problem as a 4-class classification problem and the labels are the 4 possible degree of rotation (see Figure 6).

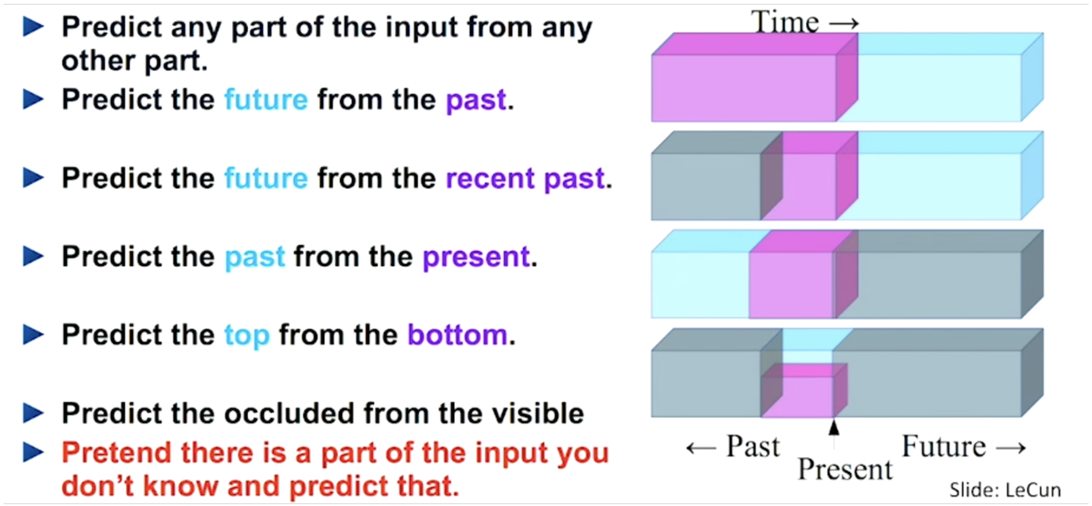


Figure 5: Summary of Self-Supervised Learning (Image source: reference: [4])

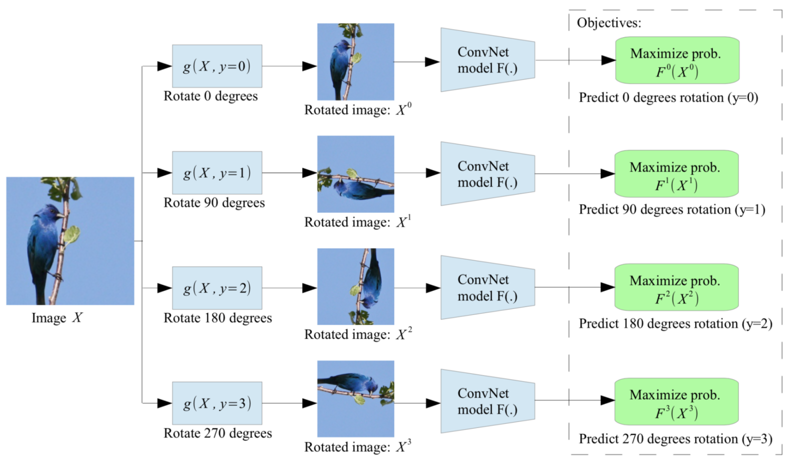


Figure 6: Illustration of self-supervised learning for rotated images (Image source: reference: [5])

## 2.2 Example: Representation Learning for Rotated/Projected Image

Supposed we have a dataset that contains images of faces. These faces are either smiley or angry and they are put in a straight position without any rotation or projection. If we feed a rotated face to the fitted model, the model might not be able to give us the correct label (happy or angry). This is because the model had never seen a rotated face before so that it has no idea what to do. We can solve this problem by using an autoencoder to learn the representation of the training dataset. The bottleneck of the autoencoder contains the representation of the input data and we use this representation to further train our classifier. The encoder can extract features from rotated/projected images that are similar to the features within the training dataset. Therefore, the fitted model will be able to classify the images correctly.

# 3 GPT-3

## 3.1 About GPT-3

Generative Pre-trained Transformer 3 (GPT-3), created by the silicon valley tech company OpenAI, is a language model that leverages deep learning techniques and generates human-like output(text). Pre-trained with 499 billion tokens, including text from websites, typical crawl, books, news, Wikipedia, etc., the model delivers excellent ability in a vast range of tasks. Such tasks include generating articles, text summarization, question answering, translating text, and even generating code [6, 7, 8].

## 3.2 How GPT-3 works

When the model's application programming interface (API) receives a piece of text, it generates specific text based on the input. Such output can be a phrase, a task, a question, or any kind of expression such as HTML syntax. For example, if we pass the following text: "The weather is" to the model, we could expect that it should be able to predict that the next word is "sunny." Figure 7 is the visualization of how GPT-3 works [9].

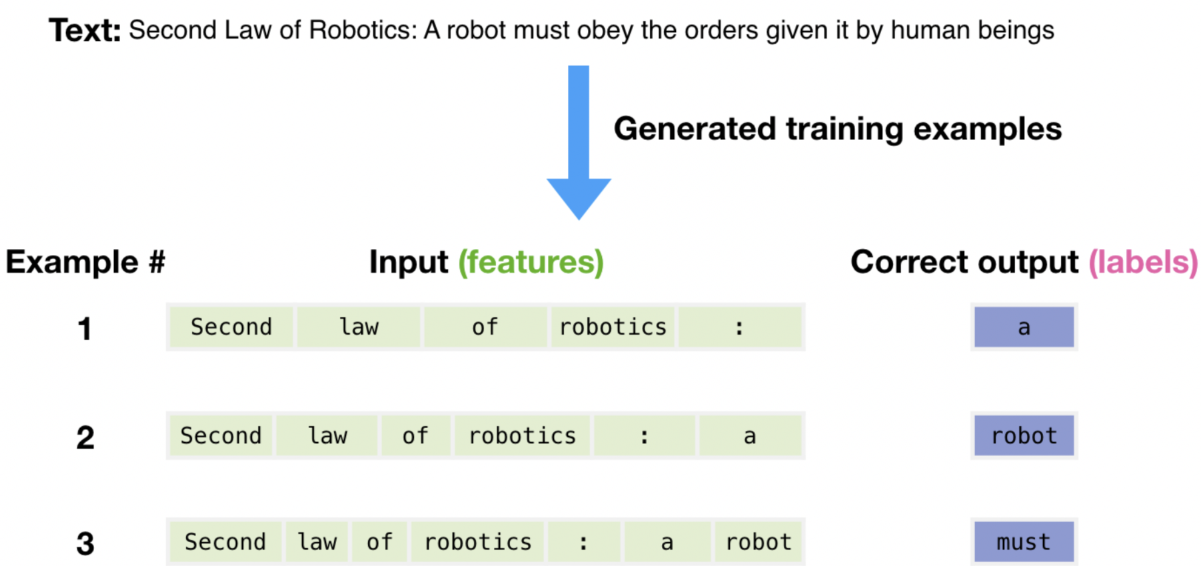


Figure 7: Visualization of how GPT-3 works (Image source: reference:[9])

# References


[1] T.-Y. Cheng, "Supervised, semi-supervised, unsupervised, and self-supervised learning," https://towardsdatascience.com/supervised-semi-supervised-unsupervised-and-self-supervised-learning-7fa79aa9247c, accessed: 2022-03-29.

[2] J. Brownlee, "14 different types of learning in machine learning," https://machinelearningmastery.com/types-of-learning-in-machine-learning//, accessed: 2022-03-29.

[3] S. Jain, “Introduction to pseudo-labelling : A semi-supervised learning technique,” https://www.analyticsvidhya.com/blog/2017/09/pseudo-labelling-semi-supervised-learning-technique/, accessed: 2022-03-29.

[4] L. Weng, “Self-supervised representation learning,” https://lilianweng.github.io/posts/2019-11-10-self-supervised/?fbclid=IwAR2ZyL1Fy9DI3vV922Rh9bf4O8SynwCDmp0OviZ7Z-9pEDryoeBqiaYvq accessed: 2022-03-29.

[5] S. Gidaris, P. Singh, and N. Komodakis, “Unsupervised representation learning by predicting image rotations,” 2018. [Online]. Available: https://arxiv.org/abs/1803.07728

[6] C. Ayuya, “Introduction to gpt3,” https://www.section.io/engineering-education/introducing-gpt3/, accessed: 2022-03-30.

[7] T. Hung, “Introduction to gpt-3,” https://developer.vonage.com/blog/20/10/05/introduction-to-gpt-3, accessed: 2022-03-30.

[8] C. Li, “Openai’s gpt-3 language model: A technical overview,” https://lambdalabs.com/blog/demystifying-gpt-3/, accessed: 2022-03-30.

[9] J. Alammar, “How gpt3 works - visualizations and animations,” https://jalammar.github.io/how-gpt3-works-visualizations-animations/, accessed: 2022-03-30.