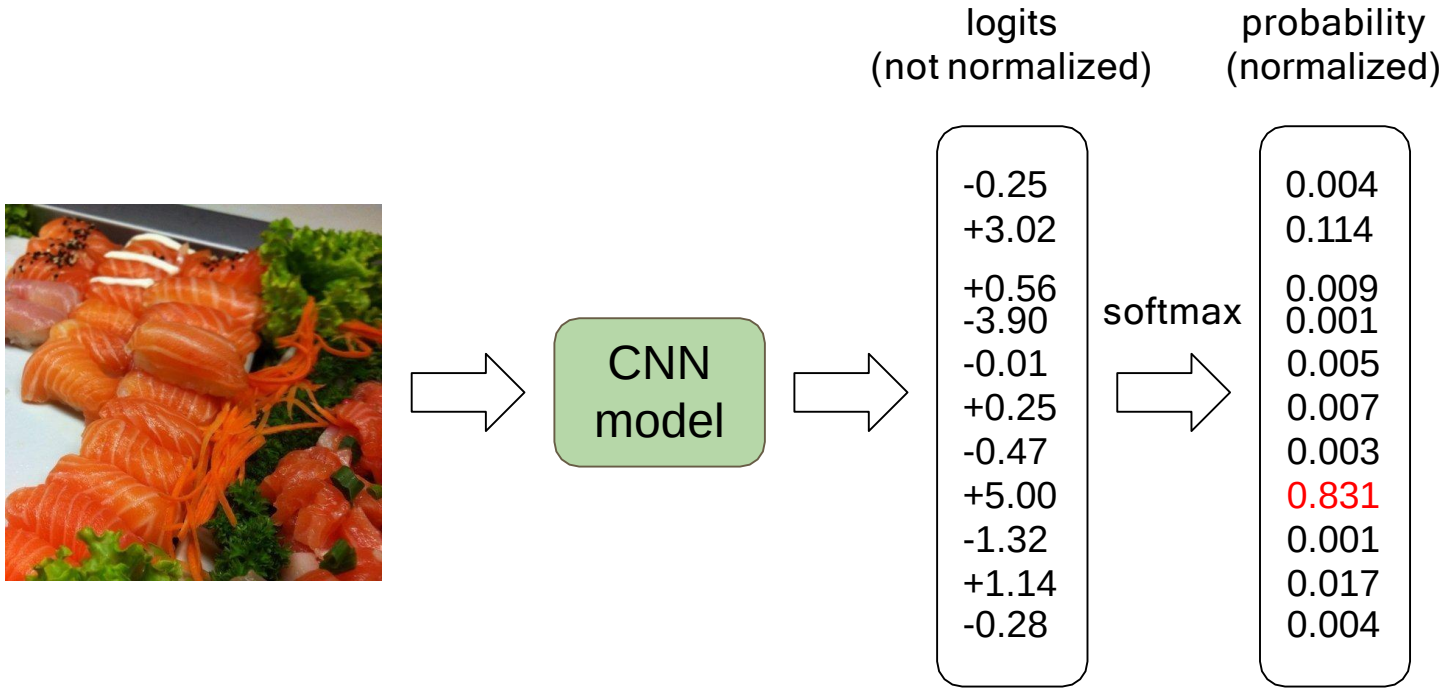
Task-Food Classification

Task-Food Classification

- The images are collected from the food-11 dataset classified into 11 classes.
- The dataset here is slightly modified:
- Training set: 270 * 11 **labeled** images + 6786 **unlabeled** images
- Validation set: 60 * 11 **labeled** images
- Testing set: 3347 images
- **DO NOT** utilize the original dataset or labels.
    - This is cheating.

Task-Food Classification

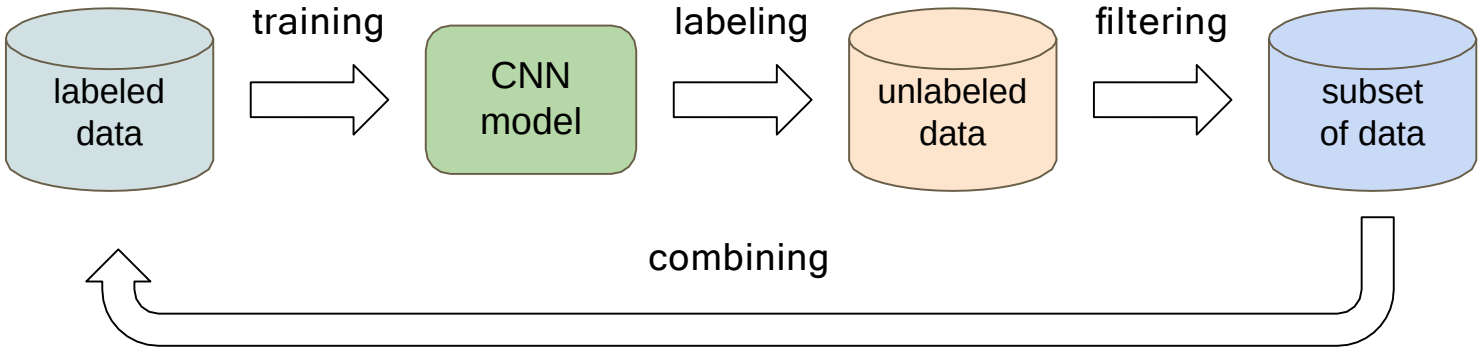


Hints

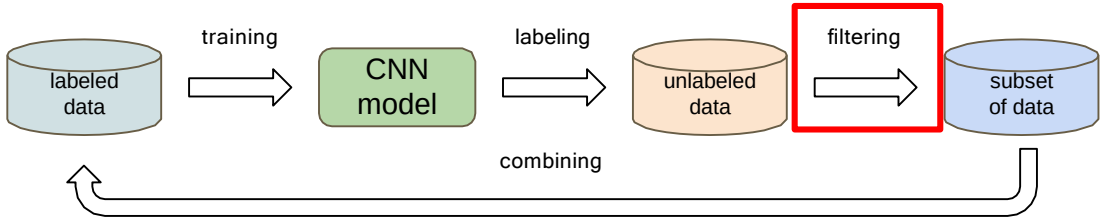
- Build a **convolutional neural network** using **labeled images** with provided codes.
- Improve the performance using **labeled images** with different model architectures or data augmentations.
- Improve the performance with **additional unlabeled images in the training set,** such as semi-supervised learning, self-supervised learning

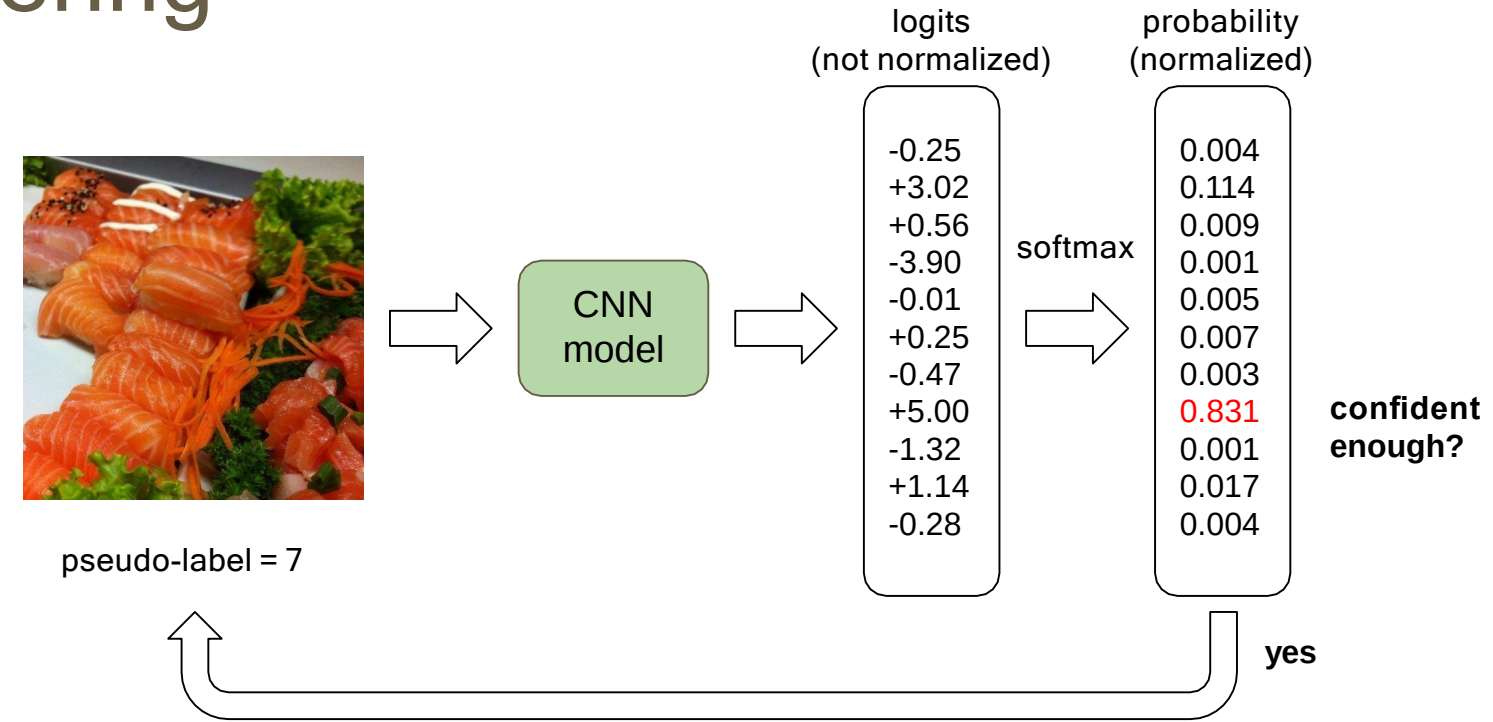
semi-supervised learning

- There are many ways to do semi-supervised learning.
- E.g., generate pseudo-labels for unlabeled data and train with them.



Pseudo-labels filtering





Kaggle Submission Format

- The submitted predictions should be in **CSV** format.
- The first row is **“Id, Category”**
- The rest of rows are “{id}, {prediction}” (e.g., 0005, 8)
- There should be (3347 + 1) rows in total.

Id	Category
0001	0
0002	9
0003	4
0004	5

Useful Resources

- Semi-supervised learning
  - MixMatch: https://arxiv.org/abs/1905.02249
  - Noisy student: https://arxiv.org/abs/1911.04252
- PyTorch
  - https://pytorch.org/
- Torchvision
  - http://pytorch.org/vision/stable/index.html