

Apprentissage Statistique Automatique II

Lucas Reding, Thibault Defourneau

TP - Part 2

Project - Food item classification for automated retail sorting

A large food retailer wants to automate its warehouse sorting system. Workers currently receive mixed boxes of food items (fruit, vegetables, packaged foods), and misplacing an item in the wrong category causes:

- Operational costs (wrong stock levels)
- Delays in distribution
- Waste (perishable items stored in wrong conditions)

To reduce errors and speed up logistics, the company wants to build an image-based classifier that assigns a food item to the correct category as soon as a photo is taken on a conveyor belt.

Students are hired as data scientists to create the first prototype.

Goal: Train a Deep learning model that takes an image of a food item and predicts its correct category

Resources: Here are some useful dataset to reach your goal

- Fruits 360 dataset from **Kaggle**
- Fruits and Vegetables Classification from **Kaggle**
- Food-101 from **Kaggle**

Second session: Model improvements & analysis

1. Design an improvement strategy based on Session 1 results (data augmentation tuning, unfreezing layers, regularization, optimizer choice).
2. Implement the selected improvements and retrain the model.
3. Compare performance with the original baseline models (CNN from scratch and pre-trained model).
4. Explore a way to use a Large Language Model to reach your goal, and compare this approach against the previous approach.
5. Analyze remaining errors and discuss operational impact for the retailer (risk of mis-sorting items).
6. Package the final model and prepare a short report summarizing methodology, results, and recommendations.