

Project proposal : Automated Classification of Fruit Growth Patterns

Lucia Korbel'ová

The objective of this machine learning project is to develop a model capable of categorizing different types of fruits based on their growth patterns. Fruit have three different ways how they can grow: on trees, on bushes or on vines. Given the inherent imbalance in the distribution of fruit growth patterns—because the majority of fruit types grow on trees—I have opted for a binary classification approach, focusing on whether a fruit grows on a tree or not. This project falls within the scope of supervised binary classification, utilizing images of fruits as input data and producing classifications based on growth patterns.

For this task, I plan to use the Fruits-360 dataset from Kaggle, initially designed for fruit recognition based on fruit names. This dataset contains diverse fruits and vegetables, each captured in 100x100 pixel images surrounded by a white background. For each fruit and vegetable type the dataset contains images of the object captured in various angles. To adapt the dataset to my binary classification task, I will organize fruits into three directories representing their growth patterns (tree, bush, vines) to make the labelling process easier and exclude irrelevant items such as vegetables and nuts. Additionally, I will refine the dataset by removing certain fruit types from the training set to assess the model's ability to generalize across previously unseen shapes. The testing dataset will also include real-life images of fruits, introducing complexity with multiple fruits per image and natural growth settings.

With a dataset comprising over 10,000 images, including more than 2,000 for training, I plan to employ a pretrained neural network as the foundation for my image classifier neural network. Since I will use a pretrained neural network, the size of my dataset is sufficient to achieve my task. To optimize the neural network's hyperparameters, a systematic random search of the hyperparameter space will be conducted. Furthermore, various image preprocessing techniques will be explored, and their effectiveness will be compared to determine the most impactful approach. The culmination of these efforts will result in the creation of the final model for automated fruit growth pattern classification.