



About ML Recognition

The Plant Doctor AI app enables you to recognize plant diseases instantly — all it takes is one photo, which is later analyzed using machine learning algorithms. While this flow usually works seamlessly and reliably, we want to ensure that our users understand the implications machine learning applications of today bear. This is especially important, given that predictions of the app could influence high-stake decisions.

Thus, we assembled some frequently asked questions. We suggest you read these before taking our app to the field. And if you still have any questions, we're happy to help you at matyas.bohacek@matsworld.io.

What is machine learning anyway?

Machine learning refers to a branch of artificial intelligence that strives to create statistical models for various tasks — such as image or speech recognition — that learn from examples. It all starts with a training dataset, where humans annotate different features, just as they would hope for the model to predict them once trained. A statistical model is then created by optimizing shallow mathematical representations to yield the best predictions on this dataset. Thanks to deep learning concepts, the models can generalize to novel inputs and scenarios outside the dataset (if everything is conducted correctly).

Long story short, it's all math, statistics, and some code — no sentient entity, no physical robot. By definition, this model will fail in some instances. While the creators of any machine learning system strive to make their models as accurate as possible, some failure rate is simply inevitable.

How was the model in your app trained, then?

Good question! We used the publicly available PlantVillage dataset, which contains 54,309 training images. When we tested the model on unseen photos, it got an accuracy of 96%. We included details about the training in an associated thesis¹ for the technical folks out there.

What does it mean for me, a farmer or a hobbyist gardener?

You don't need to worry about our math and code. We want to emphasize that our model will make errors from time to time and politely ask you to keep this in mind. Thus, always seek professional advice before taking any steps involving your plants.

¹ The thesis will be made available in May, 2023.