Analysis and classification of leaves.

[Your name / group name], [Date]

*The template contains some text describing points you may want to discuss under each heading. These are suggested elements to cover; feel free to drop what you find irrelevant. The only requirement for the project description is that it aligns with the ML lifecycle sketched in the course. In other words, the project description follows the structure reflected in the below section headings. Delete all the instructions (in italics) before submission.*

# **DESCRIBE THE PROBLEM**

## **SCOPE**

* *“Business objective”: Describe the project’s goal and “business impact.”*
* *How will the solution produced in the project be used? What similar solutions already exist / how is the problem solved today? How would you do the task manually without using machine learning?*

The objective of this project is to accurately identify the types of leaves, given features such as, the shape, lengths, and other identifiers. One of the potential uses of the results, would be to create a mobile application which makes it possible for people to take a picture of a leaf, and get the corresponding tree species, as well as information about it.

This could be one aspect of the application, and further development could include the possibility of classifying plant species in general.

There already exists multiple applications which provides the same features, however in this project, the focus will be on leaves, which may provide a more in depth and accurate solutions than more general-purpose applications.

* *How will the performance be measured via “business metrics”?*
* *If your machine learning model will be part of a more extensive “pipeline” or system, describe the system’s components. Consider how changes in one part of the system may impact other parts.*
* *Describe the stakeholders of the project*
* *Describe a tentative timeline for the project. Include milestones.*
* *Define what resources, for example, computational resources and personnel, will be required to complete the project.*

The primary resource, and expense, of the project will be man-hours spent, both in terms of development, as well as time spent researching. Additionally, there will be computational resources spent, especially during the training of the models, which may be a limiting factor when it comes to the accuracy of the timeline of the project.

## **METRICS**

* *Describe the minimal “business metric” performance for the project to be considered a success*
* *Describe machine learning and software metrics or objectives that will be used to measure whether the system/solution is working. Here, you should use metrics that can be easily calculated. Accuracy, mean squared error, latency, and throughput are examples. Describe how they are connected to the “business objective” you described earlier.*

# **DATA**

*Describe what data and labels you will use. What kind of data is it? Where will you obtain data / how can you collect data? How much data is already available, and how much do you estimate is needed? If the problem is to be attacked using “supervised learning” methods, describe how you will get the “ground truth” labels. How will you ensure that the labels are sufficiently accurate? That they are consistent?*

*Describe any privacy issues or other relevant ethical considerations.*

*How will the data be represented for the machine learning models? Describe potential needs for data cleaning, feature engineering, scaling, and the like.*

# **MODELING**

*Describe which machine learning models you will explore. Describe how you plan to estimate baseline performance and baseline behavior. Remember that your first models should be simple. Baseline performance can typically be estimated using simple models or even non-machine learning-based solutions. You can also search for results obtained by others on the same or a similar task. You can also estimate “human-level performance” if relevant. Describe how you plan to investigate prediction mistakes and “feature importance” and how this will be used to improve your results.*

# **DEPLOYMENT**

*How will the model(s) be deployed? How will the predictions be used? What are your plans for monitoring and maintaining the machine learning system? If relevant, how do you plan to improve the system after deployment?*

# **REFERENCES**

*List sources you’ve used during the planning of the project. The list of references should indicate the feasibility of your project.*