



# Foundations of Machine Learning - Exercise (SS 25)

## Assignment 7: Decision Trees

Arvinth Arunbabu

[arvinth.arun@ki.uni-stuttgart.de](mailto:arvinth.arun@ki.uni-stuttgart.de)

Akram Sadat Hosseini

[Akram.Hosseini@ki.uni-stuttgart.de](mailto:Akram.Hosseini@ki.uni-stuttgart.de)

Jiaxin Pan

[jiaxin.pan@ki.uni-stuttgart.de](mailto:jiaxin.pan@ki.uni-stuttgart.de)

Daniel Frank

[daniel.frank@ki.uni-stuttgart.de](mailto:daniel.frank@ki.uni-stuttgart.de)

Nadeen Fathallah

[Nadeen.Fathallah@ki.uni-stuttgart.de](mailto:Nadeen.Fathallah@ki.uni-stuttgart.de)

Farane Jalali

[farane.jalali-farahani@ki.uni-stuttgart.de](mailto:farane.jalali-farahani@ki.uni-stuttgart.de)

Tim Schneider

[tim.schneider@ki.uni-stuttgart.de](mailto:tim.schneider@ki.uni-stuttgart.de)

Cosimo Gregucci

[cosimo.gregucci@ki.uni-stuttgart.de](mailto:cosimo.gregucci@ki.uni-stuttgart.de)

Osama Mohammed

[osama.mohammed@ki.uni-stuttgart.de](mailto:osama.mohammed@ki.uni-stuttgart.de)

Jingcheng Wu

[jingcheng.wu@ki.uni-stuttgart.de](mailto:jingcheng.wu@ki.uni-stuttgart.de)

Submit your theoretical solution in ILIAS as a single PDF file.<sup>1</sup> Make sure to list the full names of all participants, matriculation number, study program, and B.Sc. or M.Sc. on the first page. Optionally, you can *additionally* upload source files (e.g., PPTX files). Submit your programming task in ILIAS as a single Jupyter notebook. If you have any questions, feel free to ask them in the exercise forum in ILIAS.

**Submission is open until Monday, 9th of June, 12:00 noon.**

---

<sup>1</sup>Your drawing software probably allows exporting as PDF. An alternative option is to use a PDF printer. If you create multiple PDF files, use a merging tool (like [pdfarranger](#)) to combine the PDFs into a single file.



## Task 1: Decision Trees

Researchers are studying mushroom characteristics to determine whether a mushroom is **Poisonous**. Each sample is labeled with basic observed attributes.

Each observation includes:

- **Cap Color:** e.g., Brown, White, Red
- **Gill Color:** e.g., Black, Brown, White
- **Sticky Texture:** Yes, No (indicates whether the mushroom has a sticky surface)

The goal is to predict whether a mushroom is **Poisonous**.

**Table 1** Mushroom Classification Dataset

Cap Color	Gill Color	Sticky Texture	Poisonous
Brown	White	Yes	Yes
Orange	White	Yes	Yes
Red	White	No	Yes
Yellow	White	No	Yes
Brown	Yellow	No	Yes
Orange	Yellow	Yes	No
Red	Yellow	Yes	No
Yellow	White	Yes	No
Brown	White	Yes	No
Yellow	White	No	No

1. **Task** Compute the optimal first split for the corresponding decision tree. Your split decision should maximize **normalized information gain**.
2. **Task** Using the optimal first split identified in the previous step, sketch the one-level decision tree with only the root and its immediate branches. Label the root node with the selected attribute, indicate the outcome for each branch, and show the predicted class at each leaf.



## Task 2: Decision Trees for Regression

While decision trees are commonly associated with classification tasks, they can also be applied to regression problems where the target variable is continuous.

1. **Task** The standard decision-tree algorithm is tailored to categorical class labels. What aspects do we need to change to grow a decision tree for regression problems? Provide strategies to modify the algorithm such that it can solve regression problems.



## Task 3: Train and Fine-tune a Decision Tree

Follow the instructions in the jupyter notebook.