University of Stuttgart | ANALYTIC COMPUTING

Machine Learning (SS 24)

Assignment 7: Decision Trees

Yi Wang
yi.wang@ki.uni-stuttgart.de

Akram Sadat Hosseini
Akram.Hosseini@ki.uni-stuttgart.de

Nadeen Fathallah
Nadeen.Fathallah@ki.uni-stuttgart.de

Tim Schneider
tim.schneider@ki.uni-stuttgart.de

Osama Mohammed
osama.mohammed@ki.uni-stuttgart.de

Daniel Frank
daniel.frank@ki.uni-stuttgart.de

This assignment consists of 3 pages and the `07-Decision-Trees.ipynb` notebook. with the following 2 tasks:

- Task 1: Entropy, Cross Entropy and Kullback-Leibler (KL) divergence (40 Points) 2
- Task 2: Decision Tree for Regression (60 Points) 3

Submit your theoretical solution in ILIAS as a single PDF file.[1] 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, 10th of June 2024, 12:00 noon.

[1] 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.

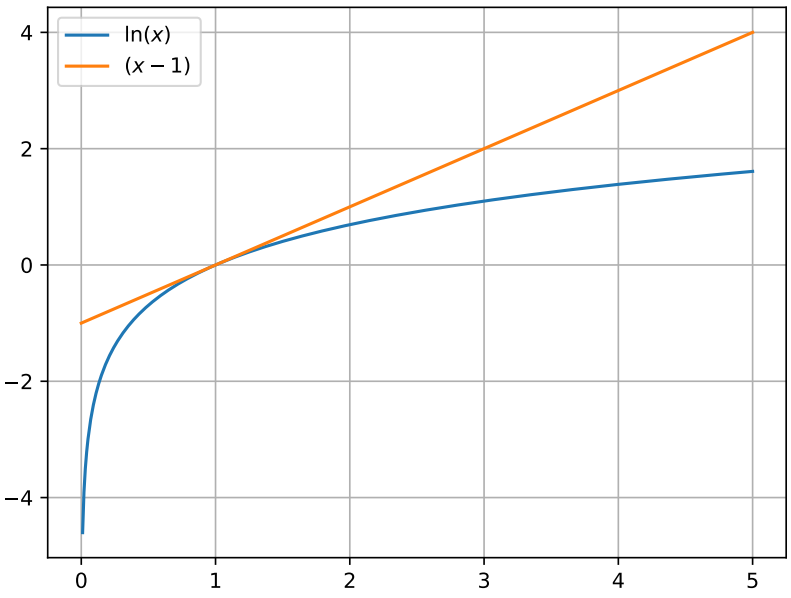

Figure 1 A helpful inequality on the natural logarithm.

Task 1: Entropy, Cross Entropy and KL divergence (40 Points)

Assume you are given two discrete distributions, on the same probability space and of the same size $P = [p_1, \ldots, p_n]$, and $Q = [q_1, \ldots, q_n]$. Show that the following properties that were introduced in the lecture hold:

1. **Task (3 Points)** The *Cross-Entropy* of the same distribution is the *Entropy*, $H(P, P) = H(P)$.
2. **Task (10 Points)** The *Cross-Entropy* is the sum of the *Entropy* and the *KL divergence*, $H(P, Q) = H(P) + D_{\mathrm{KL}}(P||Q)$.
3. **Task (5 Points)** The *KL divergence* $D_{\mathrm{KL}}(P||Q)$ is in general not symmetric. Show this with an example for the $P = [p_1, p_2]$, and $Q = [q_1, q_2]$.
4. **Task (7 Points)** Explain the relation between *Maximum Likelihood estimation* and minimum *Cross-Entropy*, think about a machine learning example.
5. **Task (15 Points)** *Optional for B.Sc. Data Science students*

   The *KL divergence* of the same distribution is zero, $D_{\mathrm{KL}}(P||Q) = 0 \iff P = Q$.

   Hint: First, show that KL divergence *is not negative with the inequality* $\ln(x) \leq (x - 1)$ *and figure out where this inequality is actually tight, see also Figure 1. It holds that* $\log_2(x) = \frac{\ln x}{\ln 2}$.

Task 2: Decision Tree for Regression (60 Points)

1. **Task (60 Points)** Follow the instructions in the *Jupyter* notebook `07-Decision-Trees.ipynb`.