Submission due date: 10th January 2024

1 Gradient boosted decision trees

1. Derive the Jacobian and Hessian of the logistic loss (show all steps!)

2. Given the following data of table tennis matches of the ML team against a team of Reinforcement Learning Engineers (RLE). x_1 denotes the length of the racket used by the ML team, x_2 its weight. y encodes the outcome of the match: $y = 1$ means the game was lost, $y = 0$ means the game was won.

x_1	x_2	y
1	2	1
1	3	1
4	1	1
2	1	0
3	2	0
4	3	0

(a) Plot the data.

(b) Learn a gradient boosted decision tree to help the ML team pick the right racket the next time. Use $\gamma = 1$ and $\lambda = 1$ and learn only one split for each tree (i.e. a decision stump). Learn 2 trees in total. Explain the result.

(c) Use `sklearn.ensemble.GradientBoostingClassifier` to check how many boosting stages (n_estimators) are required for the right prediction.

(d) Why do we need γ, λ? Explain their usage and what they stand for to a Machine Learning outsider in the context to this task, i.e. what do they mean in terms of size / weight of the racket, number of recorded games in the past, etc.