

Exercise 1:

In logistic regression, we estimate the probability $\pi(\mathbf{x}) = \mathbb{P}(y = 1 \mid \mathbf{x})$. To decide if \hat{y} is 0 or 1, we follow:

$$\hat{y} = 1 \Leftrightarrow \hat{\pi}(\mathbf{x}) \geq a$$

- a) What happens if you are choosing $a = 0.5$? More precisely, from which value of $\theta^T x$ do you predict $\hat{y} = 1$ rather than $\hat{y} = 0$?
- b) Explain (using words) why $a = 0.5$ is a sensible threshold.

Exercise 2:

Choose some of the classifiers already introduced in the lecture and visualize their decision boundaries for relevant hyperparameters. Use `mlbench::mlbench.spirals` to generate data and use `plot_learner_prediction` for visualization. To refresh your knowledge about `mlr3` you can take a look at <https://mlr3book.ml-org.com/basics.html>.

Exercise 3:

- a) What is the relationship between softmax $\pi_k(x) = \frac{\exp(\theta_k^T x)}{\sum_{j=1}^g \exp(\theta_j^T x)}$ and the logistic function $\pi(\mathbf{x}) = \frac{1}{1 + \exp(\theta^T x)}$ for $g = 2$ (binary classification)?
- b) Derive the loss function of the softmax regression. Suppose there are n instances and g features. (Hint: Apply the same steps as we used in the lecture for the logistic regression) How does this loss function relate to the negative log likelihood?
- c) Explain how the predictions of softmax regression (multiclass classification) looks like (probabilities and classes) and define the parameter space.

Exercise 4:

You are given the following table with the target variable **Banana**:

ID	Color	Form	Origin	Banana ?
1	yellow	oblong	imported	yes
2	yellow	round	domestic	no
3	yellow	oblong	imported	no
4	brown	oblong	imported	yes
5	brown	round	domestic	no
6	green	round	imported	yes
7	green	oblong	domestic	no
8	red	round	imported	no

- a) We want to use a naive Bayes classifier to predict whether a new fruit is a Banana or not. Calculate the posterior probability $\pi(x)$ for a new observation (yellow, round, imported). How would you classify the object?

- b) Assume you have an additional feature "Length", which measures the length in cm. Describe in 1-2 sentences how you would handle this numeric feature with Naive Bayes.