1. kolokvij (dodatni termin)

Predmet: *Strojno učenje*
Datum: *11. december 2023*

First midterm exam

1. Linear regression
   - What are the linear regression assumptions about residuals?
   - How do we check if the assumptions hold? Describe or draw an example of when the assumptions are true and when they are not true.
   - Can linear regression be used even if between the target variable and attributes the relationships are not linear?
   - If you think you can, describe how? If you think not, state why?
2. We want to predict house prices, we have 200 examples and 80 attributes. We decide to use linear regression with lasso regularisation for modelling.
   - Why is this model suitable or not for solving the given problem?
   - Write down the loss function for linear regression with lasso regularisation.
   - What is the regularisation parameter and how does it affect the model?
   - We would like to standardise the data before learning the model. How do proceed?
3. Logistic regression
   - Write down and describe the loss function for logistic regression (for binary classification).
   - Add L2 regularisation to the loss function for logistic regression. Describe how you would determine the coefficients of this model with gradient descent? Describe the procedure and write down the equations.
   - How do we approach multiclass classification problems with logistic regression?
4. Similarly to Task 2, we predict the price of an apartment, but this time we decide to use a decision tree:
   - Why is a decision tree an inappropriate model for solving the problem? Why is it a bad model in general?
   - Would using a random forests model solve the problems that arise when solving it with a decision tree?

- Describe how random forests and gradient boosting trees deal with the bias variance trade-off?
- Which of the two ensemble models is subject to overfitting? Why?

5. We want to solve a classification problem (see image **??**). We predict two classes (two colours of the circles in the figure).
   - Which model would you use? Describe it.
   - What hyperparameters do you have to set and how would you set them?
   - Draw the boundaries that the chosen classifier would create.

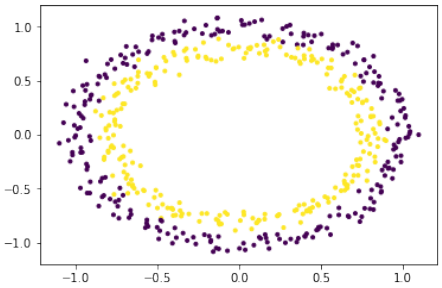

Figure 1: Slika naloge 5.