Biomedical Data Science & AI

Exercise sheet 8

SoSe2021

Prof. Dr. Holger Fröhlich, Mohamed Aborageh, Vinay Bharadhwaj, Yasamin Salimi

Due date: June 22th

Questions

Exercise 1 - Ensemble Learning (8 points)

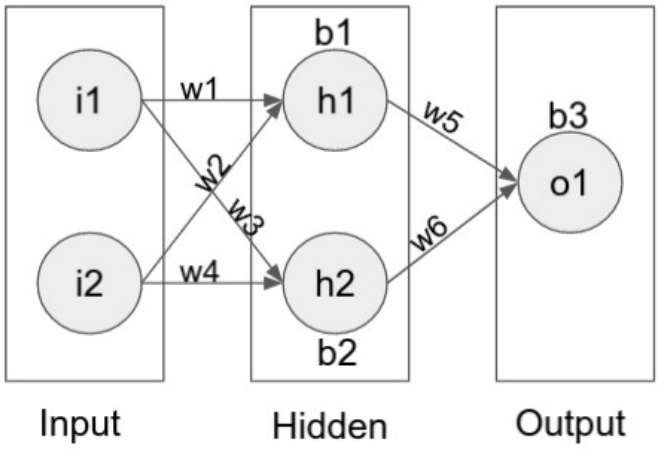
1. Inform yourself about gradient boosting, then answer the following questions in your own words: **(2 points)**
    a. What do the individual weak learners model? How does this relate to the gradient of the loss function?
    b. What is the difference between gradient boosting and random forest?
2. Which modifications make gradient boosting robust against overfitting? **(1 point)**
3. Using the *titanic_survival_dataset.csv*, train the following models using nested cross validation while optimizing a selected number of hyperparameters in the inner loop using grid search, then compute the probabilities of your targets:
    a. Random forest, optimizing the number of estimators **(1 point)**
    b. Gradient boosting, optimizing boosting steps **(1 point)**
    c. Lasso penalized logistic regression, optimizing L1 regularization strength **(1 point)**

    (Using a large parameter grid results in an extended computation time. We advise using a maximum of 5 values per hyperparameter)
4. Inform yourself about calibration curves (reliability diagrams). Use the predicted probabilities of each model from 3) to plot a calibration curve, then explain your results. **(2 points)**

Exercise 2 - NN theoretical (10 points)

1. Suppose there is a Multi-Layer Perceptron (MLP) composed of one input layer with 8 neurons, followed by one hidden layer with 30 artificial neurons, and one output layer with 3 artificial neurons. All artificial neurons use the ReLU activation function.
    a. Deduce the shape of input matrix X, hidden layer's weight vector W_h, bias vector b_h and the shape of the network's output matrix Y. **(2 points)**
    b. Write the equation that computes the network's output matrix Y as a function of X, W_h , b_h , W_o and b_o. **(2 points)**
2. What are principle and unavoidable limitations of the backpropagation (BP)? **(1 point)**

3. The shown figure is a 3 layer neural network.
    a. Compute h1, h2, o1, and total error using ReLU units. **Note:** b1, b2 and b3 represent the biases added to their respective units. **(2 points)**
    b. Calculate the updates of the network weights w1, …, w6 and bias terms b1, b2, b3 using backpropagation. Assume a learning rate of 1 for the sake of simplicity. **Note:** Remember that a bias term is equivalent to a weighted constant input 1. **(3 points)**



i1=0.5, i2=0.8
w1=0.15, w2=0.2, w3=0.25, w4=0.3, w5=0.4, w6=0.55
b1=0.4, b2=0.3, b3=0.6

Activation function for h1 and h2 is ReLU
Expected output=1

Exercise 3 - NN Programming (7 points)

1. Familiarize yourself with tensorflow and train a neural network with 2 hidden layers (10 and 8 units respectively) and predict the label feature using the *titanic_survival_dataset.csv* dataset. **(2 points)**
2. Evaluate the performance of the neural network for the same dataset in a nested cross validation by optimizing the number of units in the 2nd hidden layer in the inner cross validation. **(3 points)**
3. How does the neural network perform in comparison to the models in the calibration curve from the previous task and plot the results alongside the other models in the calibration plot? **(2 points)**