Exercise for MA-INF 2213 Computer Vision SS21
26.04.2021
Submission deadline: 09.05.2021

Important: Use Python 3.8 for your solutions. You are not allowed to use any additional python modules beyond the ones imported in the templates. Otherwise you won't get any points. Code with runtime errors or returning obviously rubbish results (e.g. nans, inf, meaningless visual output in future exercises) will give you at most half of the points. Points are also assigned to questions inside the coding exercises.

You can complete the exercise in groups of two, but only one submission per group is allowed. Include a *readme.txt* file with your group members into each solution. Points for solutions without readme file will only be given to the uploader.

1 Regression

We consider the problem of object pose regression. The world variables $y = [y_0, y_1]$ given by rotating an object along two coordinate axes. Also given are observed 510 dimensional $x = [x_0, x_1, ...x_{509}]$ PHoG[1] features per image. Appropriate *regression*.txt* files are provided for training regressing parameters and evaluating them. The header is arranged as [`numImages rowLength worldDimension`] and each row further on holds the concatenation $[y_i, x_i]_{1\times 512}$ for image I_i.
Maximum Likelihood rule will be used for learning. To evaluate the performance, take the maximum likelihood parameters to predict the values $\hat{y}_i \in 0, 1$ on the val and test sets. As a performance metric, report the MSE relative to variance of y: $\frac{MSE(\hat{y}_i, y_i)}{Var(y_i)}$ for $i \in 0, 1$

1. **Linear Regression**: Learn a linear regressor using the training data for both world variables $y = [y_0, y_1]$ independently and evaluate its performance on the test data. (*3 Points*)

2. **Non Linear Regression**: Learn a non-linear regressor for both variables independently using RBF kernels. The centers for RBF kernels will be learnt by reducing the observed features into codebooks. Do not use regularization, i.e. simply perform linear regression in the feature space you map $\mathbf{x}$ to. Split the training data into a train and a val split in a ratio of 1:1. Estimate the optimal number of clusters and σ on the val set. Also evaluate its performance on the test data. (*2 Points*)

3. **Dual Model Regression**: Learn a dual-model regressor for both variables $y = [y_0, y_1]$ independently using RBF function as the kernel. Do not use regularization, i.e. simply perform linear regression in the feature space you map $\mathbf{x}$ to. Estimate the right value for σ (standard deviation in the RBF kernel) using the val set (from 1.2). Also evaluate regressor's performance on the test data. What do you think about the val set proposed in the template? What does the regressor become equivalent to if σ approaches 0? What happens to the regression if σ approaches infinity? (*5 Points*)

2 Classification

We consider the problem of binary classification (bottles and horses). Given are the 510 dimensional PHoG features for each class, separated for training and testing. The data is arranged in a similar manner as above.

1. **Logistic Regression**: Using the bottles as positive and horses as negative examples, learn a linear classifier based on logistic regression. You may choose a simple gradient descent or the Newton's method for optimization.

   Train your classifier for 10000 iterations printing the loss and the accuracy every 1000 iterations. You should reach 75% accuracy on the test set. Can your model get stuck in a local minimum and why?

   $$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

   where, TP and FN stand for true positive and false negative respectively. (*6 Points*)

2. **Derivatives of loss function**: Given the loss function for logistic linear regression parameterized by ϕ (Eq. 9.6 in [2])

   $$L = \sum_{i=1}^{I} w_i \log\left[\frac{1}{1+\exp[-\phi^T \mathrm{x}_i]}\right] + \sum_{i=1}^{I}(1-w_i)\log\left[\frac{\exp\left[-\phi^T \mathrm{x}_i\right]}{1+\exp[-\phi^T \mathrm{x}_i]}\right]$$

   Show the gradient to be:

   $$\frac{\partial L}{\partial \phi} = -\sum_{i=1}^{I}\left(\frac{1}{1+\exp[-\phi^T \mathrm{x}_i]} - w_i\right)\mathrm{x}_i \tag{2}$$

   (*4 Points*)

References

[1] *A. Bosch*, *A. Zisserman*, and *X. Munoz*, Representing shape with a spatial pyramid kernel, In Proceedings of the International Conference on Image and Video Retrieval, pp. 401-408, 2007.

[2] *C. Bishop*, Pattern Recognition and Machine Learning, Springer 2006