CSE 6363 - *Machine Learning*

Homework 1- Spring 2019

Due Date: Feb. 8 2019, 11:59 pm

MLE and MAP

1. In class we covered the derivation of basic learning algorithms to derive a model for a coin flip task. Consider a similar problems where we monitor the time of the occurrence of a severe computer failure (which requires a system reboot) and which occurs according to a Poisson process (i.e. it is equally likely to happen at any point in time with an arrival rate of λ). For a Poisson process the probability of the first event to occur at time x after a restart is described by an exponential distribution:

$$p_\lambda(x) = \lambda e^{-\lambda x}$$

We are assuming here that the different data points we measured are independent, i.e. nothing changes between reboots.

a) Derive the performance function and the optimization result for analytic MLE optimization for a model learning algorithm that returns the MLE for the parameter λ of the model given a data set $D = \{k_1, ...k_n\}$. Make sure you show your steps.

b) Apply the learning algorithm from a) to the following dataset:

$$D = \{1.5, 3, 2.5, 2.75, 2.9, 3\} .$$

c) Derive the optimization for a MAP approach using the conjugate prior, the Gamma distribution. The Gamma distribution is:

$$p_{\alpha,\beta}(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}$$

Note that α and β are constants and that there still is only one parameter, λ, to be learned. Show your derivation and the result for the data in part $b)$ and values for α and β of 5 and 10, respectively.

K Nearest Neighbor

2. Consider the problem where we want to predict the gender of a person from a set of input parameters, namely height, weight, and age. Assume our training data is given as follows:

$$
\begin{aligned}
D = \{ \quad & ((170, 57, 32), \quad W), \\
& ((192, 95, 28), \quad M), \\
& ((150, 45, 30), \quad W), \\
& ((170, 65, 29), \quad M), \\
& ((175, 78, 35), \quad M), \\
& ((185, 90, 32), \quad M), \\
& ((170, 65, 28), \quad W), \\
& ((155, 48, 31), \quad W), \\
& ((160, 55, 30), \quad W), \\
& ((182, 80, 30), \quad M), \\
& ((175, 69, 28), \quad W), \\
& ((180, 80, 27), \quad M), \\
& ((160, 50, 31), \quad W), \\
& ((175, 72, 30), \quad M), \quad \}
\end{aligned}
$$

a) Using Cartesian distance as the similarity measurements show the results of the gender prediction for the following data items for values of K of 1, 3, and 5. Include the intermedia steps (i.e. distance calculation, neighbor selection, prediction).

$$(155, 40, 35), (170, 70, 32), (175, 70, 35), (180, 90, 20)$$

b) Implement the KNN algorithm for this problem. Your implementation should work with different training data sets and allow to input a data point for the prediction.

c) Repeat the prediction using KNN when the age data is removed. Try to determine (using multiple target values) which data gives you better predictions. Show your intermediate results.

Gaussian Naïve Bayes Classification

3. Using the data from Problem 2, build a Gaussian Naïve Bayes classifier for this problem. For this you have to learn Gaussian distribution parameters for each input data feature, i.e. for $p(height|W), p(height|M)$, $p(weight|W), p(weight|M), p(age|W), p(age|M)$.

   a) Learn/derive the parameters for the Gaussian Naïve Bayes Classifier and apply them to the same target as in problem $2b$). Show your intermediate steps.

   b) Implement the Gaussian Naïve Bayes Classifier for this problem.

   c) Repeat the experiment in part $2c$) with the Gaussian Naïve Bayes Classifier.

   d) Compare the results of the two classifiers and discuss reasons why one might perform better than the other.