Machine Learning 1

1 State TRUE/FALSE (Justify if false) - 10 points

a K-means solves clustering problem

b K-means is supervised learning algorithm

c Early stopping is used for regularization of neural network

d Backpropagation is algorithm used for regularization of neural network

e Cross validation is used to compare different algorithms

f

g

h

i

j

2 Principal Component Analysis - 10 points

a Given a plot. Draw the first PC. (3 points)

b In above plot, if data is reduced to one dimension, is the given data linearly separable? (3 points)

c Advantage of applying PCA before using SVM (4 points)

3 Overfitting and Regularization - 10 points

a

b

c

4 SVM, Convex Functions and Hyperplanes - 20 points

a Given linear soft-margin SVM with squared hinge loss function. Compute the gradient of the square hinge-loss SVM. (5 points)

b Show that the above SVM is a convex problem. (5 points)

c Why it is important to know if a function is convex? (3 points)

d (Given a plot) Explain if the problem can be solved with hard margin SVM. (2 points)

e Draw two hyperplanes in above plot(one using hinge loss and other using squared hinge loss). Explain the decision. (5 points)

5 Regression, SGD - 20 points

a Given optimization problem for regression. Explain how value of alpha(co-efficient of w) affects regularization procedure. (4 points)

b Computing $_{w}J(w)$ of the above problem. (4 points)

c Compute closed form solution for the above (4 points)

d SGD for regression problem (8 points)

6 Convolutional Neural Network - 15 points

a Pseudocode for CNN with Max Pooling (6 points)

b Given Matrix M, stride = 2, apply Max pooling to get the output matrix (3 points)

c Use of Dimensionality reduction to avoid overfitting while solving binary classification problem using deep neural network (6 points)

7 Kernel - 15 points

a Kernel Trick and use in Machine Learning (3 points)

b Let k_1 and k_2 be kernels. Prove that

$$k_3(y, x) = (\alpha k_1(y, x) + \beta k_2(y, x))^n$$

Hint: Use Hadamard Theorem(Product of two positive semi definite matrices is also positive semi definite) (4 points)

c Value of n when kernel matrix K is singular (8 points)