Total marks 100
Total questions 7

1. True or False with negative marking for inccorect answer. -10 marks

a) K in K-means algorithm helps decide the correct clusters and it is supervised learning. - False
b) Backpropogation is a method of regularization in neural networks - False
c) Early stopping is a regularise in ann. -True
d) Cross validation allows comapring different algorithms on same task. - True
e) Impossible to solve unsupervised tasks using ann. -False
f) Learning rate is only important in sgd not gradient decent. - False
g) Reduction of dimensionality increases RMSE. - False
h),i), j) can't remember but were direct.

2. PCA - 10 marks

a) Draw a 1-d pca line in given 2-d graph with data points with two outliers.
b) Can we use svm instead of pca? After pca does linearity change?
c) Advantages of pca.

3. Overfitting - 10 marks

a) Given a graph marks overfitting, underfitting and just right regions.
b) Justify a).
c) Disadvantages of overfitting.
d) List methods to solve overfitting and underfitting.

4. SVM - 20 marks

Given an SVM equation with loss marked as L(w) ---- (Eq 1)
Followed by two different losses defined as equations (don't remember exactly)
L1(w) = hinge loss ---- (Eq 2)
L2(w) = squared hinge loss ---- (Eq 3)

a) Show gradient of Eq1 + Eq3 subsituted (i.e. Eq4).
b) Prove Eq4 is convex (hessian logic).
c) Importance of convex function in ml.
d) Given a graph with two cluster of data points with few outliers, can hard-margin solve it?
e) Draw two hyperplanes using SWM equations with L1(w) and L2(w).

5. Optimization and Regressio - 20 marks

Given an equation with two regularizer with a prefixed constant a belongs to [0,1] to it.

a) Discuss the effect of a.
b) Gradient of given equation.
c) Closed-form solution of b).
d) Pseudocode for SGD with regression.

6. CNN -15 marks

a) Pseudocode for MAX POOLING using given parameters.
b) Given an image of size 7*7, implement max pooling with stride 2.

c) Discuss effects of dimensionality reduction in deep neural network in order to reduce overfitting.

7. Kernel - 15 marks

a) Kernel trick and it's usage in ml.
b) Some mathematical proof using the concepts of "sum of two kernels is a kernel", "multiplication of constant with a kernel will give a kernel" and The Hadamard product.
c) What is the lowest value of n for which a kernel matix is singular?