Test: Chap 6 - Dimension Reduction

Max marks: 30 Max time: 45 minutes

1. Forward feature selection uses ——- to select feature (f_i) to be included, while backward feature elimination uses —— feature (f_j) to select features to be discarded. The approach is called —- approach. (1)

   (a) $argmin_i E(F - f_i)$, $argmin_j E(F \cup f_j)$, floating search

   (b) $argmin_i E(F - f_i)$, $argmin_j E(F - f_j)$, wrapper

   (c) $argmin_i E(F \cup f_i)$, $argmin_j E(F - f_j)$, wrapper

   (d) $argmin_i E(F \cup f_i)$, $argmin_j E(F - f_j)$, floating search

2. Let X be a ten dimensional data set in R^{10} with 1000 instances and let Σ be its variance-covariance matrix. Let w_1, w_2, w_3, w_4 be the first four principal components. Write down the constraints for maximizing the variance for the fifth principal component (w_5), with complete notation. Scan your work and upload. (10)

3. The table given below shows the variance explained by six principal components of a six dimensional data set. First —- components explain more than 80% variance and —- principal component expresses noise. (2)

w1	w2	w3	w4	w5	w6
45.286	32.514	11.259	6.236	3.986	0,719

4. For d-dimensional data set with N instances, PCA yields ——, while feature embedding yields —-. (1)

   (a) N projection vectors, d transformed coordinates

   (b) d transformed coordinates, N projection vectors

   (c) d projection vectors, N transformed coordinates

   (d) N transformed coordinates, d projection vectors

5. Consider the following 2-class, 2-dimensional training set and projection vector $< 3, 7 >$. Find the class means before and after projection, between class scatter matrix before projection, class-wise scatter of training instances after projection. Write down the formal notation for the data set and the computations. (Marks: 1*2 (before projection), 2*2 (after projection), 5 (S_B), 5 (class-wise scatter))

x1	x2	C
1	5	1
10	2	0
6	12	1
2	10	1
16	4	0