	INDIAN INSTITUTE OF TECHNOLOGY **KHARAGPUR**	**Stamp/Signature of the Invigilator**

EXAMINATION (End Semester)	**SEMESTER (Autumn)**

Roll Number											**Name**	
Subject Number	**C**	**S**	**6**	**0**	**0**	**5**	**0**		**Subject Name**		***Machine Learning***	
Department/Centre of the Student											**Additional Sheets**	

Instructions and Guidelines to Students appearing in the Examination

1. Ensure that you have occupied the seat as per the examination schedule.
2. Ensure that you do not have a mobile phone or a similar gadget with you even in switched off mode. Note that loose papers, notes, books should not be in your possession, even if those are irrelevant to the paper you are writing.
3. Date book, codes or any other materials are allowed only under the instruction from the paper-setter.
4. Use the instrument box, pencil box and non-programmable calculator is allowed during the examination. However, exchange of these items is not permitted.
5. Additional sheets, graph papers and relevant tables will be provided on request.
6. Write on both sides of the answer script and do not tear off any page. Report to the invigilator if the answer script has torn page(s).
7. Show the admit card / identity card whenever asked for by the invigilator. It is your responsibility to ensure that your attendance is recorded by the invigilator.
8. You may leave the examination hall for wash room or for drinking water. Record your absence from the examination hall in the register provided. Smoking and consumption of any kind of beverages is not allowed inside the examination hall.
9. After the completion of the examination do not leave the seat until invigilator collects the answer script.
10. During the examination, either inside the examination hall or outside the examination hall, gathering information from any kind of sources or any such attempts, exchange or helping in exchange of information with others or any such attempts will be treated as adopting ‘unfair means’. Do not adopt ‘unfair means’ and do not indulge in unseemly behavior as well.

Violation of any of the instructions may lead to disciplinary action.

Signature of the Student

To be filled in by the Examiner

Question Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Total**
Marks obtained											

Marks obtained (in words)	**Signature of the Examiner**	**Signature of the Scrutinizer**

Instructions: ***Answer all FIVE questions in the spaces (box) provided. Do Not Issue Separate Answerscript.***

Time: 3 hrs, Total Marks: 5x20=100, Provide explanations wherever required. The question paper has 12 pages

---

Rough Work

1.i. State the optimization problem solved to train the soft margin support vector machine. Define each term.

ii. A linear SVM with 2 dimensional input was trained using the following training set: Class y = +1: {(X_1: 1, 1), (X_2: 1, 0), (X_3: 0, 1)}, Class y = -1: {(X_4: 0, 0.5)}. The following Lagrange multipliers were obtained after solving the QP problem: $\lambda_1 = 0$, $\lambda_2 = 2$, $\lambda_3 = 0$, $\lambda_4 = 2$. Find the equation of the optimal separating hyperplane.

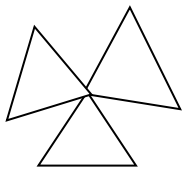
iii. The quadratic kernel is defined to be: $k(X, Y) = (X^T Y + c)^2$, where $X, Y \in \mathbb{R}^n$, and $c \geq 0$. Find the feature mapping function $\Phi(X)$ corresponding to the quadratic kernel. **[5 + 5 + 10 = 20]**

Rough Work

2.i. Consider the following *distance matrix* between six points. Draw the dendogram of these points obtained by the *complete linkage* clustering algorithm. **[14 + 6 = 20]**

	P1	**P2**	**P3**	**P4**	**P5**	**P6**
P1	0.00	0.24	0.22	0.37	0.34	0.23
P2		0.00	0.15	0.20	0.14	0.25
P3			0.00	0.15	0.28	0.11
P4				0.00	0.29	0.22
P5					0.00	0.39
P6						0.00

ii. Approximately sketch the clusters obtained by clustering algorithms for the data distributions given below. The darkness of a region denotes its density.

Clustering Algorithm	K-means (K=2)	Single linkage (2 clusters)	DBSCAN (suitable epsilon and MinPts values)
			
			

Rough Work

3. We have the following perceptron like model for predicting y from x: $z = w_0 + w_1 x + w_2 x^2$, $y = 1 + e^z$ Error is measured as $E = \frac{1}{2}(\log y - \log t)^2$. Here, the input x and target output t are both scalars. We want to train the scalar model weights w_0, w_1, w_2, using gradient descent on the error function E. Derive the weight update rules of the gradient descent algorithm. **[20]**

Rough Work

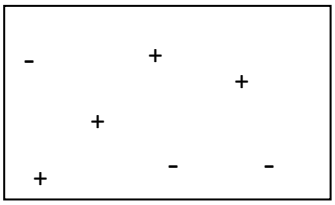
4. We are given the following set of training examples. The features F1, F2, F3, are Boolean-valued. Consider a very weak learner within the AdaBoost algorithm. This learner simply chooses for its classification the *lowest-numbered feature that has not yet been used*. Its only intelligence is that it decides whether or not to *negate* this feature, depending on which option works best, i.e., the first time called it will return either *F1* or *NOT(F1)* as its model. Obtain the classifier in each iteration of Adaboost. Stop training if all features have been chosen.

[20]

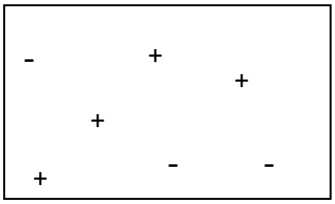
F1	*F2*	*F3*	*Class*
T	T	F	+
F	T	T	+
T	F	T	-
F	T	F	-
F	F	T	-

Rough Work

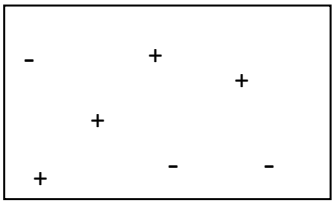
5. We want to classify the following labelled examples. Roughly, show on the diagrams how each learning algorithm *might* partition feature space into classes +/- based on these examples. Qualitatively explain, to the right of each diagram, why this algorithm might partition the data in this manner. **[5 x 4 = 20]**



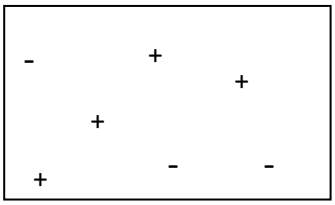
Naive Bayes



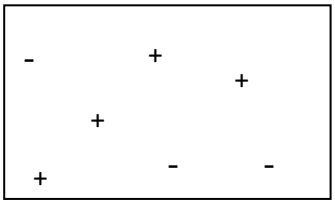
1-Nearest Neighbor



Decision Tree



Perceptron Trained with delta rule



Neural Network with 2 Hidden Units Trained with Backpropagation

--- END ----

Rough Work