Registration No :

Total Number of Pages : 01

M.Sc.I
FMCE905

9th Semester Regular Examination 2019-20
DATA SCIENCE
BRANCH : M.Sc.I(MC)
Time : 3 Hours
Max Marks : 70
Q.CODE : HR267
Answer Question No.1 which is compulsory and any FIVE from the rest.
The figures in the right hand margin indicate marks.

Q1		**Answer the following questions :**	**(2 x 10)**
	a)	What do you mean by bagging?	
	b)	What is logistic regression in Data Science?	
	c)	What is a Random Forest?	
	d)	What is Lift? How it is related to Association rule?	
	e)	What is the difference between Supervised and Unsupervised Learning?	
	f)	What do you mean by Bias?	
	g)	What is a Back Propagation Network?	
	h)	What is subset selection and why it is essential?	
	i)	In k-means algorithm how the centers of the clusters are initialized?	
	j)	What is ridge regression and Lasso regression?	
Q2	**a)**	What is the objective of Bootstrap method? How it is different from cross validation, Explain.	**(5)**
	b)	Write a short note on Principal Components Analysis.	**(5)**
Q3	**a)**	Discussed the different components of Artificial Neural Networks.	**(5)**
	b)	What is Linear Regression? Explain it with an example.	**(5)**
Q4	**a)**	Describe k-nearest neighbor algorithm.	**(5)**
	b)	Explain perceptorn learning algorithm.	**(5)**
Q5	**a)**	Write a short note on Ada Boost.	**(5)**
	b)	"The Higher the Bias the Lower is the variance and vice versa". Explain the statement with a suitable diagram taking X as the independent variable and Y is the dependent variable	**(5)**
Q6		Describe the McNemar's test and k-fold cross validated paired t-test to compare whether two given learning algorithms construct classifiers that have the same expected error rate.	**(10)**
Q7		What is SVM? Mention the role of SVM. Explain the linear classifier related to SVM.	**(10)**
Q8		**Write short answer on any TWO :**	**(5 x 2)**
	a)	Write a short note on following data sets	
		(i) California housing (ii) New Zealand fish	
	b)	Cluster Analysis	
	c)	Bias-Variance Tradeoff	