II SEMESTER

MBA (Business Analytics)

September

SUPPLEMENTARY EXAMINATION-EVEN 2019

PAPER CODE: MB207   TITLE OF PAPER: Predictive Modelling

Time: 3:00 Hours   Max. Marks: 60

Note: Answer any five questions out of the given six questions.
All questions carry equal marks.
Assume suitable missing data, if any.

Q.1[a] Differentiate between classification and regression algorithms in machine learning. (6)

[b] The owner of an automobile repair shop studied the waiting times for customers who arrive at the shop for an oil change. The following data with waiting times in minutes were collected over a week period.

12 10 17 11 21 13 18

i) Find Mean waiting time for the above data. (3)
ii) Compute Z score. Are there any outliers? Explain. (3)

Q.2[a] What are the difference among the mean, median and mode, and what are the advantages and disadvantages of each? (6)

[b] An Analysis of the monthly sales of two of the firms gives the following data:

	Firm A	Firm B
Number of employees	986	548
Variances	100	121
Average Sales (in INR)	52.5	47.5

i) Which firm pays out larger sales? (2)
ii) In which firm does greater variability occur? (2)
iii) What is the average sales if firm A and firm B are combined? (2)

Q.3 The following are the data for miles travelled and travel time:

Miles Travelled (in Km)	100	50	50	80	75	65	90
Travel Time (in hours)	9	5	6	7	7.5	6.5	8

i) Construct a scatter plot and interpret it. (3)

P. T. O.

ii) Use least square method to determine the regression model for the above problem. (6)

iii) Predict the travel time (in hours) when miles travelled is 85 km. (3)

Q.4[a] The residuals for 10 consecutive time periods are as follows:

Time Period	1	2	3	4	5	6	7	8	9	10
Residual		-10	-8	-6	-4	-2	+2	+4	+6	+8	+10

Compute Durbin-Watson statistic to check the autocorrelation among the residuals. (6)

[b] The following ANOVA summary table is for multiple regression model with two independent variables. Complete the table and find Adjusted R^2 : (6)

Source	Degree of freedom	Sum of squares	Mean squares	F statistic
Regression	5	80		
Error	20	120		
Total	25	200		

Q.5 Consider the following correlation matrix, relating to three variables with unities in the diagonal spaces:

		Variables	
	A	B	C
A	1.000	0.84	0.9
B	0.84	1.000	0.95
C	0.9	0.95	1.000

(left axis label: Variables)

Use the Principal Component Method of factor analysis to determine the factor loadings for only First Principal Component for the above matrix. (12)

Q.6[a] A prospective MBA student wanted to predict starting salary upon graduation, based on program per- year tuition. If $RSS=21.8$ and $TSS=64$ for this problem; determine the coefficient of determination R^2 and interpret its meaning. (6)

[b] Explain logistic regression model by giving a suitable example. (6)