SUBJECT CODE: 305 **SUBJECT : BIGDATA**

Unit 1 : Introduction Bigdata

- Answer the following :

1) What is Bigdata?
2) Enlist tools used for Bigdata?
3) What are the 5 V's of Bigdata?
4) What is Digital Data?
5) What are the applications of Bigdata?
6) What is Bigdata Analytics?
7) Differentiate between unstructured and structured Digital data?
8) Explain advantages and disadvantages of Bigdata?
9) What is meant by petabyte and zettabyte?
10) Which are the bigdata Data sources with example?

Unit 2 : Introduction Data Science

- Answer the following :

1) What is Data science?
2) Enlist phases of data analytics life cycle?
3) Which are types of data analytics?
4) What is the statistical inference?
5) What is regression and correlation?
6) Explain needs of data analytics?
7) What is probability distribution modeling?
8) Explain phases of Data science projects?
9) Differentiate between population and sample?
10) What are the example of statistical inference?

Unit 3 : Introduction to Machine Learning

- Answer the following :

1) What is Machine Learning?
2) What is supervised machine learning?
3) Write short note on: Naive Bayes'.
4) What is KNN?
5) What is regression analysis? Define it.
6) With the help of example describe K-means algorithm.
7) Short note on Apriori algorithm.
8) Short note on Linear regression.
9) Define: decision node
10) Define analysis

Unit 4 : Data Analytics with R/Weka machine learning

- Answer the following :

1) What is Data manipulation?
2) What is data visualization?
3) What is WEKA?
4) What is pipe operator? Explain?
5) What is summaries ? Explain?
6) What is Box Plot?
7) What is Histogram?
8) Define structured Data.
9) How to install it in windows.
10) What is Base R Graphics?