

--	--	--	--	--	--	--	--	--	--

**GANDHI INSTITUTE OF ENGINEERING AND TECHNOLOGY UNIVERSITY, ODISHA, GUNUPUR
(GIET UNIVERSITY)**

Ph.D. (Second Semester) Examinations, May – 2025

**WPPECA2014/23SPPECS2012 – Data Preparation and Analysis
(CSE)**



Time: 3 hrs

Maximum: 70 Marks

The figures in the right hand margin indicate marks.

Answer ANY FIVE Questions.

(14 x 5 = 70 Marks) Marks

1. a. What is data preprocessing ? Explain the different methods for data cleansing phase. 7
 b. List out various scalability and real time issues of data gathering and preparation methods during data pre-processing techniques? 7
2. a. Explain briefly about consistency checking after the data cleaning technique? 7
 b. What do you mean by data transformation? Explain the different techniques of data transformation during data preprocessing? 7
3. a. What are the data cleaning methods? Discuss. 7
 b. How to clean the Heterogeneous data? Explain. 7
4. a. Define support vector machine(SVM) and further explain the maximum margin linear separators concept. 6
 b. Apply k-means algorithm on given data for K=2. 8
 Use C1(2,4) & C2C2(6,3) as initial cluster centres.
 Data : a(2,4),b(3,3),c(5,5),d(6,3),e(4,3)
5. a. Explain what are data correlations and why is it important to identify correlations? 7
 Differentiate the difference between correlation and regression analysis?
 b. Explain what are the important aspects of data analysis? Explain what are the challenges of interactive visual data analysis? 7
6. a. Discuss about different types of data visualization 7
 b. Explain how does visualization of big data help in interpreting information? 7
7. a. Draw and explain the architecture of convolutional network. 7
 b. How do you differentiate the OLSR and classification algorithms ? State and explain with the suitable examples through Python or R language 7
8. a. What is neural network ? Explain BPN algorithm with suitable example 6
 b. Write a short notes of the followings : 8
 - i) LDA
 - ii) PCA

---End of Paper---