Roll No.											

PRESIDENCY UNIVERSITY

BENGALURU

End - Term Examinations – MAY 2025

Date: 31-05-2025 **Time:** 09:30 am – 12:30 pm

School: SOE	**Program:** B. Tech (EEE)	
Course Code: EEE3032	**Course Name:** Big Data Analytics in Power Systems	
Semester: VI	**Max Marks**: 100	**Weightage**: 50%

CO - Levels	**CO1**	**CO2**	**CO3**	**CO4**	**CO5**
Marks	**10**	**10**	**30**	**26**	**24**

Instructions:

(i) Read all questions carefully and answer accordingly.

(ii) Do not write anything on the question paper other than roll number.

Part A

Answer ALL the Questions. Each question carries 2marks. **10Q x 2M=20M**

1.	What is the primary goal of optimization in big data for electric power systems?	**2 Marks**	**L1**	**CO3**
2.	Define Big Data in the context of power systems?	**2 Marks**	**L1**	**CO3**
3.	Mention any two challenges faced in Big Data optimization for electric power systems.	**2 Marks**	**L1**	**CO3**
4.	Name any two commonly used databases for scientometric analysis?	**2 Marks**	**L1**	**CO3**
5.	Mention one application of Big Data in power system operation.	**2 Marks**	**L1**	**CO3**
6.	State the primary purpose of data mining in a power distribution system?	**2 Marks**	**L1**	**CO4**
7.	List the technical losses in power distribution systems?	**2 Marks**	**L1**	**CO4**
8.	List the types of data mining commonly used in electricity theft detection?	**2 Marks**	**L1**	**CO4**
9.	Define contingency analysis in power systems?	**2 Marks**	**L1**	**CO5**
10.	List the main advantages of using Big Data in contingency analysis?	**2 Marks**	**L1**	**CO5**

Part B

Answer the Questions. **Total Marks 80M**

11.	**a.**	Explain the scope and opportunities of big data in power system applications.	**10 Marks**	**L2**	**CO1**
		Or			
12.	**a.**	Explain an overview of traditional power systems and the need for modernization, and differentiate the Traditional Power Grid and Smart Grid.	**10 Marks**	**L2**	**C01**

13.	**a.**	Describe the working principle of Phasor Measurement Units (PMUs) with block diagram.	**10 Marks**	**L2**	**CO2**
		Or			
14.	**a.**	Describe various data traffic patterns and flows within smart grid environments.	**10 Marks**	**L2**	**CO2**

15.	**a.**	Explain the concept of optimization and its relevance to Big Data in electric power systems.	**10 Marks**	**L2**	**CO3**
	b.	A generator (G) needs to supply power to a load (L). There are two possible paths from G to L: Path 1: G → A → L → Loss = 6 units Path 2: G → B → L → Loss = 3 units Use ACO to determine the better path (i.e., the one with less power loss).	**10 Marks**	**L3**	**CO3**
		Or			
16.	**a.**	Classify optimization techniques into deterministic and stochastic methods based on problem characteristics.	**10 Marks**	**L2**	**CO3**
	b.	A power system has three generators—Gen1, Gen2, and Gen3—that must collectively supply 100 MW of electrical load. Each generator has a linear fuel cost function defined as follows: Gen1 Cost = 2 × P_1 (₹/MW), Gen2 Cost = 3 × P_2 (₹/MW), Gen3 Cost = 5 × P_3 (₹/MW) Where: P_1, P_2, and P_3 are the power outputs of Gen1, Gen2, and Gen3 respectively. The total power output must satisfy the load demand: $P_1 + P_2 + P_3 = 100$ MW. Determine the optimal power distribution (P_1, P_2, P_3) among the generators such that the total fuel cost is minimized while meeting the load demand constraint by using the Genetic Algorithm (GA) technique.	**10 Marks**	**L3**	**CO3**

17.	a.	Apply the k-means clustering technique to classify a set of substations into two groups based on their annual failure rates and average load. Explain how the clustering helps in identifying: Group 1: Substations with lower failure rates and lower load Group 2: Substations with higher failure rates and higher load Provide the steps you would follow in performing this clustering and interpret how the results can be used for reliability planning. Table 1: Data for 5 substations (Failure rate: failures/year, Load: MW)	20 Marks	L3	CO4

Substation	Failure Rate (failures/year)	Average Load (MW)
A	1	50
B	2	45
C	6	90
D	7	85
E	3	60

Or

18.	a.	Apply the Logistic Regression classification technique to categorize transformers in a power system as either "Healthy" or "Needs Maintenance" based on the following features: Age (in years) Number of faults in the last year Given a new transformer with Age = 8 years and Faults = 2, use the trained model to predict its maintenance status. Clearly outline the steps involved in model training, feature scaling (if needed), and prediction. Table 2: Data for 6 Transformers	20Marks	L3	CO4

Transformer	Age (years)	Faults (count)	Status (Label)
T1	5	1	Healthy (0)
T2	10	3	Needs Maintenance (1)
T3	7	0	Healthy (0)
T4	15	4	Needs Maintenance (1)
T5	3	0	Healthy (0)
T6	12	5	Needs Maintenance (1)

19.	**a.**	Explain the process of contingency analysis in power systems and how big data improves the accuracy of predictions.	10 **Marks**	**L**2	**CO5**
	b	Apply data processing and preprocessing techniques for contingency analysis.	**10 Marks**	**L3**	**CO5**
Or					
20.	**a.**	Describe how big data technologies support contingency analysis in modern power systems.	**10 Marks**	**L2**	**CO**5
	b	Given a power distribution network with critical components such as transformers and circuit breakers, along with external challenges like weather events, how would you apply predictive modeling techniques to assess the total outage severity for the system over the course of a year? Discuss the steps involved in creating a model to predict outage durations, considering both internal and external factors.	**10 Marks**	**L3**	**CO5**