VIT
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)
CHENNAI

Reg. Number: 21BLC1122

Continuous Assessment Test (CAT) – II – APR 2024

Programme	:	B.Tech (ECE/ECM)	Semester	:	WS 2023-24
Course Code & Course Title	:	BCSE352E & ESSENTIALS OF DATA ANALYTICS	Class Number	:	CH2023240500995 CH2023240501020 CH2023240501038 CH2023240501079
Faculty	:	Dr. JEETASHREE APARAJEETA Dr. LOGESH R Dr. ROHITH G Dr. VIJAYARAJAN R	Slot	:	TC2
Duration	:	90 Minutes	Max. Mark		50

General Instructions:

- Write only your registration number on the question paper in the box provided and do not write other information.
- Use statistical tables supplied from the exam cell as necessary
- Use graph sheets supplied from the exam cell as necessary
- Only non-programmable calculator without storage is permitted

Answer all questions

Q. No	Sub Sec.	Description	Marks
1.		**A study is performed to analyze the relationship between the concentration of a certain pollutant in a lake (measured in parts per million, ppm) and the distance from a factory (measured in kilometers) over a period of six months. The data collected is as follows:** **Distance from Factory (x)**: 1, 2, 3, 4, 5, 6 **Pollutant Concentration (y)**: 10, 8, 6, 5, 4, 3 **a) Calculate the least squares regression line y = a x + b.** **b) Find the Standard Error of the Estimate (SEE).** **c) Predict the pollutant concentration in the lake at a distance of 7 kilometers from the factory using the regression model.**	10
2.		**Mr. Joe working as a Data Analyst at a Krish Tech Inc. He has been assigned a new task to build a simple and efficient clinical expert system based on set of rules designed by a Human Medical Practitioner. Solve the problem of designing a simple and effective rule-based classifier to support Mr. Joe's current task. Illustrate the structure and working of the classifier you propose to Mr. Joe and discuss in detail about its advantages and disadvantages.**	10
3.		**A university is investigating the effectiveness of three different teaching methods (Traditional Lecture, Active Learning, and Online Modules) for improving student performance in a statistics course.**	15

After completing the course, the final exam scores of students taught using each method are recorded. Test at a 5% level of significance (α=0.05) whether there is a significant difference in mean final exam scores among the three teaching methods.

Traditional Lecture	75	80	70	78	82
Active Learning	82	85	88	80	85
Online Modules	78	75	80	82	77

4.

Yash is working as a data scientist for a music streaming service. His task is to develop a simple model to predict whether a user will like a new song recommendation based on their listening history. He decided to use the Naive Bayes classifier for this task.

Yash was provided with a dataset containing a sample of user listening history and their corresponding preferences for new song recommendations (like or dislike). Each user's listening history is represented by a list of genres they've listened to.

User ID	Listening History	Preference
1	Rock, Pop, Hip-Hop	Like
2	Classical, Jazz	Like
3	Pop, Hip-Hop, R&B	Dislike
4	Country, Rock	Like
5	Hip-Hop, Rap, Pop	Dislike
6	Electronic, Dance, Pop	Like
7	Rock, Metal, Alternative	Like
8	Jazz, Blues, Classical	Dislike
9	Pop, Indie, Alternative	Like
10	Hip-Hop, Rap, R&B	Dislike

Using the provided dataset, formulate the frequency table and likelihood table for the Naive Bayes classifier. Implement Bayes' theorem to determine while given a user's listening history of Rock, Pop, and Hip-Hop genres, should the user be recommended a new song or not?

15

**********All the best **********

1·4

Critical Values of the F-Distribution: $\alpha = 0.05$

Denom. d.f.	Numerator Degrees of Freedom									
	1	2	3	4	5	6	7	8	9	10
1	161.448	199.500	215.707	224.583	230.162	233.986	236.768	238.883	240.543	241.882
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385	19.396
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.786
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438	3.388	3.347
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420	2.366	2.321
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397	2.342	2.297
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375	2.320	2.275
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355	2.300	2.255
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337	2.282	2.236
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321	2.265	2.220
27	4.210	3.354	2.960	2.728	2.572	2.459	2.373	2.305	2.250	2.204
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291	2.236	2.190
29	4.183	3.328	2.934	2.701	2.545	2.432	2.346	2.278	2.223	2.177
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266	2.211	2.165
31	4.160	3.305	2.911	2.679	2.523	2.409	2.323	2.255	2.199	2.153
32	4.149	3.295	2.901	2.668	2.512	2.399	2.313	2.244	2.189	2.142
33	4.139	3.285	2.892	2.659	2.503	2.389	2.303	2.235	2.179	2.133
34	4.130	3.276	2.883	2.650	2.494	2.380	2.294	2.225	2.170	2.123
35	4.121	3.267	2.874	2.641	2.485	2.372	2.285	2.217	2.161	2.114
36	4.113	3.259	2.866	2.634	2.477	2.364	2.277	2.209	2.153	2.106
37	4.105	3.252	2.859	2.626	2.470	2.356	2.270	2.201	2.145	2.098
38	4.098	3.245	2.852	2.619	2.463	2.349	2.262	2.194	2.138	2.091
39	4.091	3.238	2.845	2.612	2.456	2.342	2.255	2.187	2.131	2.084
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180	2.124	2.077
41	4.079	3.226	2.833	2.600	2.443	2.330	2.243	2.174	2.118	2.071
42	4.073	3.220	2.827	2.594	2.438	2.324	2.237	2.168	2.112	2.065
43	4.067	3.214	2.822	2.589	2.432	2.318	2.232	2.163	2.106	2.059
44	4.062	3.209	2.816	2.584	2.427	2.313	2.226	2.157	2.101	2.054
45	4.057	3.204	2.812	2.579	2.422	2.308	2.221	2.152	2.096	2.049
46	4.052	3.200	2.807	2.574	2.417	2.304	2.216	2.147	2.091	2.044
47	4.047	3.195	2.802	2.570	2.413	2.299	2.212	2.143	2.086	2.039
48	4.043	3.191	2.798	2.565	2.409	2.295	2.207	2.138	2.082	2.035
49	4.038	3.187	2.794	2.561	2.404	2.290	2.203	2.134	2.077	2.030
50	4.034	3.183	2.790	2.557	2.400	2.286	2.199	2.130	2.073	2.026
60	4.001	3.150	2.758	2.525	2.368	2.254	2.167	2.097	2.040	1.993
70	3.978	3.128	2.736	2.503	2.346	2.231	2.143	2.074	2.017	1.969
80	3.960	3.111	2.719	2.486	2.329	2.214	2.126	2.056	1.999	1.951
90	3.947	3.098	2.706	2.473	2.316	2.201	2.113	2.043	1.986	1.938
100	3.936	3.087	2.696	2.463	2.305	2.191	2.103	2.032	1.975	1.927
120	3.920	3.072	2.680	2.447	2.290	2.175	2.087	2.016	1.959	1.910
140	3.909	3.061	2.669	2.436	2.279	2.164	2.076	2.005	1.947	1.899
180	3.894	3.046	2.655	2.422	2.264	2.149	2.061	1.990	1.932	1.884
200	3.888	3.041	2.650	2.417	2.259	2.144	2.056	1.985	1.927	1.878
∞	3.841	2.996	2.605	2.372	2.214	2.099	2.010	1.938	1.880	1.831

3