

END SEM EXAMINATION[CSE 4th YEAR]

DATA-MINING & DATA WAREHOUSING [CSE-S-521]

[DURATION-3 HOURS]

NOTE--- ALL QUESTIONS ARE COMPULSORY

[MAX MARKS-50]

[SECTION A]

Q1(i). Explain Hunt's algorithm to build a decision tree. [2]

(ii) Explain Apriori algo drawbacks and its improvements [3]

Q2. (i) Explain Bayesian classification.

(ii) By using Bayesian classification Predict playing tennis in the day with the condition $P(v)$ outlook= Rain, temperature=Mild , humidity=high wind=strong) using the following training data: [3+5]

Day	Outlook	Temperature	Humidity	Wind	Play Tennis
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

Q3(i). Explain Genetic Algorithm and all its steps. [3]

(ii).Maximize $f(x) = x^3$.where x is should vary between 0 and 20. Represent X as binary unsigned integer of length 4 .Use population size as 4.Repeat all the steps two times . [5]

Q4.(i) Explain ECLAT Algorithm by an example. [3]

(ii)In FP-Growth Algorithm how can we represent the structure of node to generate prefix path . [3]