Total no. of Pages:02 Roll No..................

FIFTH SEMESTER-B. TECH
END-SEMESTER EXAMINATION, DECEMBER, 2021

Course Code: CACSC14/ COCSC14

Course Title: Principles of Compiler Construction

Time:3hrs. Max.Marks:40

Note: - Attempt all five questions. Missing data/information (if any), may be suitably assumed and mentioned in the answer.

Q1. Attempt any two parts of the following (4+4) CO1

a) Draw and explain the block diagram of a Compiler. Indicate the output of every stage corresponding to input z = x + y • v. Differentiate pass and phase of a compiler.

b) Explain the language processing system and the role of linker, loader and assembler.

c) Explain compiler construction tools.

Q2. Attempt any two parts of the following (4+4) CO2

a) Construct the minimized DFA from regular expression a(a|b)*ab using Thompson's construction.

b) List out the functions of a Lexical Analyzer? State the reasons for the Separation of Analysis phase into Lexical, Syntax, and Semantic Analysis. Write regular expression for the language containing all strings of 0's and 1's that do not contain 011.

c) Write a lexical analyzer for keywords in 'C' language and show it using LEX.

Q3. Attempt any two parts of the following (4+4) CO3

a) Differentiate between regular expressions and Context free grammar. Consider the following grammar

S->SS+|SS•|a

Show how the string aa+a• can be generated by this grammar. Construct parse tree for this string. Find whether this grammar is ambiguous or not.

b) Construct the SLR parsing table for the given grammar and show the parsing moves for the input string "zxxx"

S -> AxB
S ->Bc
A -> yA
A -> z
B -> xB
B -> ε

c) Construct unambiguous context-free grammar for Arithmetic Expressions in postfix notation show it using YACC.

Q4. Attempt any two parts of the following (4+4) CO4

a) What is basic block? How do you construct basic blocks? Construct basic block and flow graph for the following code

```
i=0;
s=0;
while (i<10)
{ s=s+i;
  i=i+1;
}
```

b) Explain Syntax directed definition and Syntax directed Translation scheme by taking suitable example. Write SDT for converting infix to postfix.

c) Explain the different components of activation record. Translate the expression into quadruples, triples and indirect triples.

-(a+b) • (c+d) + (a+b+c)

Q5. Attempt any two parts of the following (4+4) CO4, CO5

a) What are the issues in generation of target code? Draw directed acyclic graph (DAG) for the following expression. What are the advantages of using DAG.

e = (a+b) • (b-c) + (a+b) • (b-c)

	b) Using the Sethi Ullman algorithm generate target code for the following expression assuming only two registers r1 and r2 are available. Explain the steps. ((a+b) - (c*d)) + ((e-f) * (u-v)) c) What is code optimization? What are the issues associated with code optimization? Explain various code optimization techniques		

FIFTH SEMESTER-BTECH-COE

END SEMESTER EXAMINATION: DECEMBER 2021

Course Code: COCSC15
Course Title: Cloud Computing

Time: 3 Hours

Max Marks: 40

Note: Attempt all questions
Assume suitable missing data, if any

Q.No	Question	Marks	CO
Q1.	**Attempt any two parts of the following**		
~~1a~~	Describe a real-life example to illustrate the concepts behind cloud computing.	4	5
1b	Give a brief note on the merits and demerits of cloud computing.	4	1
1c	Describe several approaches of cloud migration.	4	2
Q2.	**Attempt any two parts of the following**		
~~2a~~	What are SLAs? How SLAs differ for each type of cloud deployment?	4	1
2b	What is outsourced community cloud?	4	2
2c	What are the characteristics of hybrid cloud?	4	1
Q3.	**Attempt any two parts of the following**		
~~3a~~	Write short notes on end user and service provider responsibilities of cloud service models with a suitable diagram.	4	2
3b	Write short notes on cloud service models that emerged after the introduction of cloud computing.	4	3
3c	Explain how cloud computing facilitates individuals and start-up industries.	4	5
Q4.	**Attempt any two parts of the following**		
4a	What are protection rings? Explain how it is used in virtualization.	4	4
4b	Differentiate full virtualization, paravirtualization, and hardware-assisted virtualization techniques.	4	2
~~4c~~	What is the role of hypervisor in virtualization? Briefly explain the different types of hypervisors with a neat diagram.	4	2
Q5.	**Attempt any two parts of the following**		
~~5a~~	Explain how cloud computing is different from virtualization.	4	2
5b	Explain MapReduce workflow with the help of a diagram and suitable example.	4	3
5c	Explain any three components of HDFS architecture. Explain how HDFS deals with over replication and under replication of blocks.	4	4

Total No. of Page: 2

Roll No. [redacted]

END SEMESTER EXAMINATION December 2021

Course Code: COCSC16
Course Title: Data Mining

Time: 3 Hours

Max. Marks : 40

Note: - **Attempt all the five questions. Missing data/ information if any, maybe suitably assumed & mentioned in the answer.**

<table>
<tr><th>Q. No.</th><th>Question</th><th>Marks</th><th>CO</th></tr>
<tr><td><b>Q1</b></td><td><b>Attempt any 2 parts of the following.</b></td><td></td><td></td></tr>
<tr><td>1a</td><td>Elaborate various stages of Data Mining Process.</td><td>4</td><td>CO1</td></tr>
<tr><td>1b</td><td>Differentiate classification and Regression for predictive analysis. Explain clustering.</td><td>4</td><td>CO1</td></tr>
<tr><td>1c</td><td>Suppose the fraction of undergraduate students who play football is 15% and the fraction of graduate students who play football is 23%. If one-fifth of the college students are graduate students and the rest are undergraduates, what is the probability that a student who plays football is a graduate student? Also, Suppose 30% of the graduate students live in hostel but only 10% of the undergraduate students live in hostel. If a student plays football and lives in hostel, is he or she more likely to be a graduate or undergraduate student? You can assume independence between students who live in hostel and those who play football.</td><td>4</td><td>CO1</td></tr>
<tr><td><b>Q2</b></td><td><b>Attempt any 2 parts of the following.</b></td><td></td><td></td></tr>
<tr><td>2a</td><td>Discuss four techniques to deal with missing data in dataset along with suitable examples.</td><td>4</td><td>CO1</td></tr>
<tr><td>2b</td><td>Calculate Dissimilarity matrix for given dataset?
Object: 1, 2, 3, 4, 5
Values: 40, 50, 42, 21, 30</td><td>4</td><td>CO2</td></tr>
<tr><td>2c</td><td>If two data objects are given as x ={3,2,0,5,0,0,0,2,0,0} and y= {1,0,0,0,0,0,0,1,0,2}. Calculate its Cosine Similarity.</td><td>4</td><td>CO1</td></tr>
<tr><td><b>Q3</b></td><td><b>Attempt any 2 parts of the following.</b></td><td></td><td></td></tr>
<tr><td>3a</td><td>Consider market basket dataset shown in the following table.
<table>
<tr><th>T. ID</th><th>Items Purchased</th><th>T. ID</th><th>Items Purchased</th><th>T. ID</th><th>Items Purchased</th></tr>
<tr><td>1</td><td>{a,d,e}</td><td>3</td><td>{a,d,b}</td><td>5</td><td>{b,c}</td></tr>
<tr><td>2</td><td>{a,d,b,c}</td><td>4</td><td>{a,e}</td><td>6</td><td>{a,d,b,e,c}</td></tr>
</table>
Compute frequent pattern generated by FP growth algorithm if Minimum Support is 2.</td><td>4</td><td>CO4</td></tr>
<tr><td>3b</td><td>For following given dataset, generate association rules using Apriori Algorithm. Consider Min Support as 50% and Confidence as 75%.
<table>
<tr><th>Transaction ID</th><th>Items Purchased</th><th>Transaction ID</th><th>Items Purchased</th></tr>
<tr><td>T1</td><td>{bread,egg,cheese}</td><td>T3</td><td>{bread}</td></tr>
<tr><td>T2</td><td>{egg, juice}</td><td>T4</td><td>{bread,egg}</td></tr>
</table></td><td>4</td><td>CO4</td></tr>
</table>

~~3c~~ The following contingency table summarizes supermarket transaction data, where *sandwiches* refers to the transactions containing sandwiches, *-sandwiches* refers to the transactions that do not contain sandwiches, *burgers* refers to the transactions containing burgers, and *-burgers* refers to the transactions that do not contain burgers. — 4 — CO2

	sandwiches	*-sandwiches*	**Sum row**
burgers	2000	500	2500
-burgers	1000	1500	2500
Sum col	3000	2000	5000

a. Suppose that the association rule "*sandwiches* → *burgers*" is mined. Given a minimum support threshold of 25% and a minimum confidence threshold of 50%, is this association rule strong?

b. Compare the results using lift and χ^2 correlation measures.

Q4 **Attempt any 2 parts of the following.**

4a Consider the following 1-D data. — 4 — CO2

X	0.5	3.0	4.5	4.6	4.9	5.2	5.3	5.5	7.0	9.5
Label	N	N	P	P	P	N	N	P	N	N

a. Classify the data point X = 5.0 according to its 3-, and 5- nearest neighbors.

b. What would be the class label if distance-weighted voting approach is used?

4b Consider the dataset shown in the following table having three attributes A1, A2, and A3. Predict the class label for a test sample (A1 = 0, A2 = 1, A3 = 0) using the Naïve Bayes Algorithm. — 4 — CO3

Instance	A1	A2	A3	Class Label
1.	0	0	0	Class 1
2.	0	0	1	Class 2
3.	0	1	1	Class 2
4.	0	1	1	Class2
5.	0	0	1	Class 1
6.	1	0	1	Class 1
7.	1	0	1	Class 2
8.	1	0	1	Class 2
9.	1	1	1	Class 1
10.	1	0	1	Class 1

4c Explain following: GINI Index, Entropy, Average Information Entropy and Information gain. — 4 — CO3

Q5 **Attempt any 2 parts of the following.**

5a Expectation Maximization clustering works on which parameter and how it is done. What is the role of Parsing and soft parsing. — 4 — CO4

5b Given points five points C1, C2, C3, C4 and C5 where d(C_1,C_2): 7, d(C_1,C_3): 4, d(C_1,C_4): 1, d(C_1,C_5): 2, d(C_2,C_3): 8, d(C_2,C_4): 5, d(C_2,C_5): 11, d(C_3,C_4): 9, d(C_3,C_5): 10 and d(C_4,C_5): 3. Perform Agglomerative clustering (Single Linkage) on these points and find Dendrogram. — 4 — CO4

5c Divide the given sample data in three (3) clusters using k-means Algorithm. — 4 — CO4

Height: 183, 171, 167, 176, 180, 177, 180, 180, 182, 183, 185, 185

Weight: 70, 56, 60, 72, 84, 76, 71, 68, 69, 77,72, 74

Total No. of Page: ~~2~~ 4

Roll No. ____________

B.Tech., V Semester, Computer Engineering, CSAI
END SEMESTER EXAMINATION December 2021

Course Code: COCSC17/CACSC17
Course Title: Machine Learning

Time: 3 Hours

Max. Marks : 40

Note: - Attempt all the five questions. Missing data/ information if any, maybe suitably assumed and mentioned in the answer.

Q. No.	Question	Marks	CO
Q1	**Attempt any 2 parts of the following.**		
1a	Elaborate on the following ML tasks and their type, describe suitable performance metrics, and explain how the system can gain experience for learning: (i) a handwriting recognition system (ii) forecasting the network traffic (throughput in bps) during different times of a day.	4	CO1
1b	For the following training dataset, find the linear regression model parameters for Y=a.X+b. Round the parameters to two decimal places. For the derived model, calculate R^2 and adjusted R^2.	4	CO1

X:	1	3	5	7	9
Y:	4	5	7	10	13

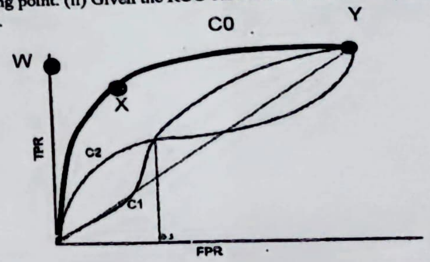
Q. No.	Question	Marks	CO
1c	Distinguish between logistic regression as a discriminative classifier and Naïve Bayes as a generative classifier in terms of their probabilistic assumptions. Suppose you train a logistic regression classifier, and the learned hypothesis function is: $h_\theta(X) = \sigma(\theta_0 + \theta_1 X_1)$, where, $\theta_0 = 6$, $\theta_1 = -1$. Draw a graph to show the output probability. What is the log-odds when x=7?	4	CO2
Q2	**Attempt any 2 parts of the following.**		
2a	Consider the following set of training examples: (Age range in the dataset is 16 to 30 years)	4	CO3

Instance	Have Laptop	Age	Buy new laptop
1	T	16	Y
2	T	19	Y
3	T	18	N
4	F	27	N
5	F	24	Y
6	F	28	Y

Derive the Decision Tree (DT) for the above dataset showing information gain at each node

split.

Q. No.	Question	Marks	CO
2b	Describe the working of the Random Forest Algorithm, explaining why aforest of stubs (collection of very small Trees) is preferred rather than a single Tree? How will the performance of the ensemble be affected if the attributes and the training examples are not distributed randomly?	4	CO3
2c	Explain how boosting can reduce both bias error and variance error. In an Adaboost ensemble, there are three DTs (1,2,3). They give the odds (probability of no-error / probability of error) as: $odds_1 = 1.6$, $odds_2 = 2.7$, $odds_3 = 0.5$. Their predictions for a given datapoint are: $\widehat{y_1} = -1, \widehat{y_2} = 1, \widehat{y_3} = -1$. Calculate the final prediction of the ensemble.	4	CO3
Q3	**Attempt any 2 parts of the following.**		
3a	(i) Given the ROC curve C0 and the operating points W, X and Y, justify which of the three is the best operating point. (ii) Given the ROC curves for two classifiers C1 and C2, compare their performance.	4	CO2



Q. No.	Question	Marks	CO
3b	Given below are the test results of two classifiers developed to detect COVID patients. Derive their confusion matrices given that the threshold for output is at 0.6. Which of the two would a doctor use if she/he does not want to miss people with COVID, even if it means some normal people are diagnosed as COVID positive?	4	CO1

Prediction 1

y	y pred
0	0.5
1	0.9
0	0.7
1	0.7
1	0.3
0	0.4
1	0.5

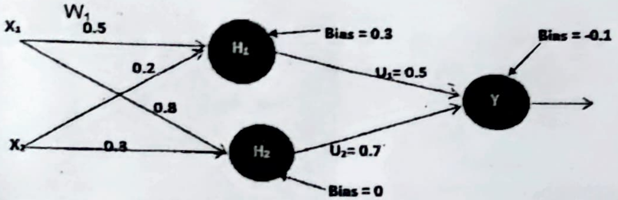
Prediction 2

y	y pred
0	0.6
1	0.9
0	0.7
1	0.7
1	0.3
0	0.8
1	0.5

Q. No.	Question	Marks	CO
3c	What is the purpose of performing cross validation. Given the dataset {(x,y)}: { (0.1,1)(0.4,1)(0.5,0) (0.4,0) (0.9,1) (0.3,0) (0.8,1) (0.8,0) (0.2,1) (0.6,1) (0.9,1) (0.2,0) (0.3,0) (0.6,1) (0.5,1) (0.7,0) }, construct the decomposition of runs using 4-fold cross validation. Calculate the overall accuracy if the predictions for the four cycles are: (1010), (1110), (0111), (0101).	4	CO1
Q4	**Attempt any 2 parts of the following.**		

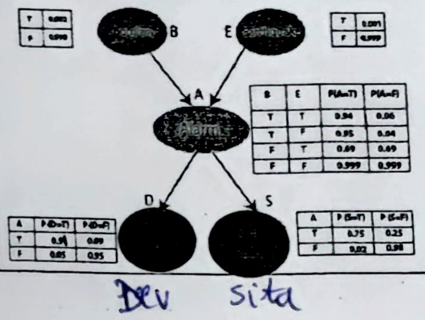
<table>
<tr><td>4a</td><td>Explain any two of the following:
- Any one Recurrent Neural Network architecture and its application
- Any one Long Short Term Memory network architecture, and its activation equations
- Dropout Regularization in Convolutional Neural Networks</td><td>4</td><td>CO1</td></tr>
<tr><td>4b</td><td>Given the MLP FNN below with sigmoid activations and training data, calculate the signal values at the input and output of each neuron, the error at the output and weight adjustment of U1 and W1 after a cycle of back-propagation.

Training data = {X1= 0.2, X2=0.4, Y=0.6}



X₁
W₁
0.5
0.2
0.8
X₂
0.3
H₁
H₂
Bias = 0.3
U₁= 0.5
U₂= 0.7
Bias = 0
Y
Bias = -0.1

</td><td>4</td><td>CO3/5</td></tr>
<tr><td>4c</td><td>For a CNN architecture, explain the need for (i) stride >1, (ii) Maxpool layer, (iii) higher dropout rate at output classification layers, (iv) use of ReLU in inner feature learning layers.</td><td>4</td><td>CO5</td></tr>
<tr><td><b>Q5</b></td><td><b>Attempt any 2 parts of the following.</b></td><td></td><td></td></tr>
<tr><td>5a</td><td>Classify a Red.Domestic SUV using Naïve Bayes classification using the dataset given below:
<table>
<tr><th>Example No.</th><th>Color</th><th>Type</th><th>Origin</th><th>Stolen?</th></tr>
<tr><td>1</td><td>Red</td><td>Sports</td><td>Domestic</td><td>Yes</td></tr>
<tr><td>2</td><td>Red</td><td>Sports</td><td>Domestic</td><td>No</td></tr>
<tr><td>3</td><td>Red</td><td>Sports</td><td>Domestic</td><td>Yes</td></tr>
<tr><td>4</td><td>Yellow</td><td>Sports</td><td>Domestic</td><td>No</td></tr>
<tr><td>5</td><td>Yellow</td><td>Sports</td><td>Imported</td><td>Yes</td></tr>
<tr><td>6</td><td>Yellow</td><td>SUV</td><td>Imported</td><td>No</td></tr>
<tr><td>7</td><td>Yellow</td><td>SUV</td><td>Imported</td><td>Yes</td></tr>
<tr><td>8</td><td>Yellow</td><td>SUV</td><td>Domestic</td><td>No</td></tr>
<tr><td>9</td><td>Red</td><td>SUV</td><td>Imported</td><td>No</td></tr>
<tr><td>10</td><td>Red</td><td>Sports</td><td>Imported</td><td>Yes</td></tr>
</table></td><td>4</td><td>CO4</td></tr>
<tr><td>5b</td><td>Derive the dual form of the optimized SVM, explaining its advantage over the original form. Explain how the "Kernel trick" handles non-linearly separable data.</td><td>4</td><td>CO3</td></tr>
<tr><td>5c</td><td>Given the Bayesian network shown in the Fig. below, calculate the probability that the alarm has sounded, but there is neither a burglary, nor an earthquake occurred, and Hari received a call only from Dev, but not from Sita.</td><td>4</td><td>CO4</td></tr>
</table>

<table>
<tr><td></td><td>

T
0.002
F
0.998
B
E
T
0.001
F
0.999
A
Alarm
B
E
P(A=T)
P(A=F)
T
T
0.94
0.06
T
F
0.95
0.04
F
T
0.69
0.69
F
F
0.999
0.999
D
S
A
P(D=T)
P(D=F)
T
0.91
0.09
F
0.05
0.95
A
P(S=T)
P(S=F)
T
0.75
0.25
F
0.02
0.98
Dev
Sita

</td><td></td><td></td></tr>
</table>

END SEMESTER EXAMINATION December 2021

Course Code: COCSE06/CACSE03
Course Title: Cryptography Techniques

Time: 3 Hours

Max. Marks : 50

Note: - **Attempt all the five questions. Missing data/ information if any, maybe suitably assumed & mentioned in the answer.**

Q. No.	Question	Marks	CO
Q1	**Attempt any 2 parts of the following.**		
1a	Define message integrity, non-repudiation, message authentication and message confidentiality.	5	CO2
~~1b~~	Explain Euclidean algorithm. What is the primitive root of a number?	5	CO2
1c	List and define five security services.	5	CO5
Q2	**Attempt any 2 parts of the following.**		
~~2a~~	Briefly discuss Diffie-hellman key exchange. Consider two parties Alice and Bob trying to establish a secret key between them using Diffie-hellman key exchange. They select prime number p = 23, g = 5, and secret integers x = 6, y = 15, respectively. Find out the messages sent by Alice and Bob, and the secret key.	5	CO2
2b	Explain meet in the middle attack. Also, provide the ways by which it can be secure.	5	CO5
2c	Define encryption and decryption in RSA algorithm. Consider p = 3, q = 11, and e = 7. Encrypt and decrypt plaintext M = "2". Also, explain how to determine the strength of the RSA algorithm.	5	CO2
Q3	**Attempt any 2 parts of the following.**		
3a	Explain the digital Signature scheme. How it ensures authentication, data integrity, and non-repudiation. Show how digital signatures can also ensure confidentiality.	5	CO3
3b	Define a hash function. Write the properties of hash function in cryptography. Explain the secure hash algorithm.	5	CO3
~~3c~~	Explain a scenario of secret key distribution protocol where man-in-the-middle attacks are ineffective.	5	CO4
Q4	**Attempt any 2 parts of the following.**		
4a	Discuss the steps in user authentication through Kerberos with a suitable diagram.	5	CO5
4b	Discuss the details of X.509 authentication service. How is an X 509 certificate revoked?	5	CO3
~~4c~~	Discuss the need for email security Explain the sequence of steps involved in the message generation and reception in PGP with block diagrams.	5	CO4
Q5	**Attempt any 2 parts of the following.**		
5a	Write in detail the definition, characteristics, types, and limitations of firewalls.	5	CO4
5b	Explain Intrusion Detection System and methods to counter it.	5	CO1
~~5c~~	What do you mean by IP security protocol? Explain the basic issue with IPSEC clients.	5	CO5