

Computer Science Courses for Undergraduate Programme of study with **Computer Science** discipline as one of the **two** Core Disciplines
(For e.g. courses for B.A. Programmes with Computer Science as Major discipline)

Computer Science Major

DISCIPLINE SPECIFIC CORE COURSE (DSC-5): Database Management System

CREDIT DISTRIBUTION, ELIGIBILITY AND PRE-REQUISITES OF THE COURSE

Course title & Code	Credits	Credit distribution of the course			Eligibility criteria	Pre-requisite of the course
		Lecture	Tutorial	Practical/ Practice		
DSC05: Database Management System	4	3	0	1	Class XII Passed	NIL

Learning Objectives

The course will give an overview of categories of data models used by database management systems and writing queries in relational algebra. The importance of file indexing and controlled execution of transactions will be taught. The course would give students hands-on practice to write complex queries, nested queries using aggregate functions in SQL and to use basic database administration commands.

Learning outcomes

On successful completion of the course, students will be able to:

- Write queries using relational algebra, a procedural language.
- Use the concept of functional dependencies to remove data anomalies and arrive at normalized database design.
- Understand the data storage system, file organization and need of file indexing.
- Learn the importance of transaction processing and concurrency control.
- Write complex and nested SQL queries, and learn basic database administration commands.
- Acquire information about emerging technologies.

SYLLABUS OF DSC 05

Unit 1 (6 hours)

Data models and Relational Algebra: Categories of data models, types of database users, Relational Algebra Operations from SET Theory, SELECT, PROJECT, JOIN, DIVISION Operations.

Unit 2 (8 hours)

Normalization: Functional dependencies, minimal cover, normalizing database with multiple keys till 3NF, Boyce-Codd Normal Form

Unit 3 (8 hours)

Database Storage and index Structures: Storage of data, file structure, file organization and its types: Sequential, Heap and Indexed, Hash file, need for file Indexing, single- and multi-level indexing.

Unit 4 (8 hours)

Transaction Processing: Concurrent execution of transactions and their handling, ACID properties, need of data recovery and log files.

Unit 5 (8 hours)

Advanced SQL: Nested and complex queries using Inner JOIN, Left JOIN, Right JOIN, Full JOIN, views, Database Administration Commands: COMMIT, ROLLBACK, drop database, control permissions etc.

Unit 6 (7 hours)

Emerging Technology: Distributed Database Concepts, Introduction to emerging technologies like Data Warehousing and OLAP, Data Mining.

Essential/recommended readings

1. Elmasri R. and Navathe B. S. *Fundamentals of Database Systems*, 7th Edition, Pearson Education, 2016.
2. Murach J. *Murach's MySQL*, 3th Edition, Pearson, 2019.

Additional References

1. Silberschatz, A., Korth, H.F., & Sudarshan, S. Database System Concepts, 8th Edition, McGraw Hill, 2019.
2. Ramakrishnan, R. & Gehrke, J. Database Management Systems, 3rd Edition, Tata McGraw Hill Education, 2014.

Suggested Practical List : (30 Hours)

MySQL might be used for practical purposes.

The following concepts must be introduced to the students:

1. Multiple table queries (join on different and same tables)
2. Nested select statements
3. Set manipulation using (any, in, contains, all, not in, not contains, exists, not exists union, intersect, minus, etc.)
4. Categorization using group by... having
5. Arranging using order by

List of sample queries to be done using the above schema.

Create tables with relevant column and integrity constraints and populate the tables with data.

Perform the following queries on the database:

1. Display all the details of all employees working in the company.
2. Display ssn, lname, fname, address of employees who work in department no 7.
3. Retrieve the birthdate and address of the employee whose name is 'Franklin T. Wong' and who is a manager
4. Retrieve the name and salary of every employee working in department 'Production'
5. Display maximum and minimum salaries of each department
6. Retrieve all employee names whose address is in 'Bellaire' and working on at least one project
7. Modify the datatype of hours column to float in works_on table.
8. Retrieve all employees who age is more than 35 years
9. Retrieve all employees in department 5 whose salary is between 50,000 and 60,000(inclusive)
10. Add another column named "Join_Date" in the employee table and "Address" in the department table
11. Retrieve the names of all employees who do not have supervisors
12. Retrieve SSN and department name for all employees
13. Retrieve the name and address of all employees who work for the 'Research' department
14. For every project located in 'Stafford', list the project number, the controlling department number, and the department manager's last name, address, and birthdate

15. For each employee, retrieve the employee's name, and the name of his or her immediate supervisor
16. Retrieve all combinations of Employee Name and Department Name
17. Make a list of all project numbers for projects that involve an employee whose last name is 'Narayan' either as a worker or as a manager of the department that controls the project
18. Increase the salary of all employees working on the 'ProductX' project by 15%. Retrieve employee name and increased salary of these employees
19. Retrieve a list of employees and the project name each works in, ordered by the employee's department, and within each department ordered alphabetically by employee first name
20. Select the names of employees whose salary does not match with salary of any employee in department 10
21. Drop the column "Join_Date" from the employee table
22. Retrieve the name of each employee who has a dependent with the same first name and same sex as the employee
23. Retrieve the employee numbers of all employees who work on project located in Bellaire, Houston, or Stafford.
24. Change the name of existing table DEPT_LOCATIONS to DLOCATIONS
25. Find the sum of the salaries of all employees, the maximum salary, the minimum salary and the average salary for each department. Display with proper headings
26. Find the sum of the salaries and number of employees of the 'Marketing' department, as well as the maximum salary, the minimum salary, and the average salary in this department
27. Select the names of employees whose salary is greater than the average salary of all employees in department 10
28. For each department, retrieve the department number, the number of employees in the department and their average salary
29. For each project, retrieve the project number, the project name, and the number of employees who work on that project for more than 10 hours
30. Change the location and controlling department number for all projects having more than 5 employees to 'Bellaire' and 6 respectively
31. For each department having more than 10 employees, retrieve the department name, number of employees drawing more than 40,000 as salary
32. Display employee names having no dependent in descending order along with their age
33. For each department, find the number of female and number of male employees along with name of manager of that department
34. Find the name and age of youngest employee in each department
35. Change the name of MINIT attribute of an employee table to MNAME using alter command
36. Create a view to keep names of employees and their department names if they are not working on any project
37. Create a view to keep track of all employee names who are working on same projects as that of employee 'Franklin T. Wong'
38. Execute system administrative commands like commit, rollback, granting control permissions, etc.

DISCIPLINE SPECIFIC CORE COURSE (A5): Machine Learning

0 CREDIT DISTRIBUTION, ELIGIBILITY AND PRE-REQUISITES OF THE COURSE

Course title & Code	Credits	Credit distribution of the course			Eligibility criteria	Pre-requisite of the course
		Lecture	Tutorial	Practical/ Practice		
A5: Machine Learning	4	3	0	1	Pass in Class XII	A1 Programming Fundamentals using Python, A2 Data Visualization using Python

Learning Objectives

The course aims at introducing the basic concepts and techniques of machine learning so that a student can apply machine learning techniques to a problem at hand.

Learning outcomes

On successful completion of the course, students will be able to:

- Differentiate between supervised and unsupervised learning tasks.
- Appreciate the need of preprocessing, feature scaling and feature selection.
- Understand the fundamentals of classification, regression and clustering
- Implement various machine learning algorithms learnt in the course.

SYLLABUS OF A5

Unit 1 (8 hours)

Introduction: Basic definitions and concepts, key elements, supervised and unsupervised learning, introduction to reinforcement learning, applications of ML.

Unit 2 (10 hours)

Preprocessing: Feature scaling, feature selection methods. dimensionality reduction (Principal Component Analysis).

Unit 3 (12 hours)

Regression: Linear regression with one variable, linear regression with multiple variables, gradient descent, over-fitting, regularization. Regression evaluation metrics.

Unit 4 (15 hours)

Classification: Decision trees, Naive Bayes classifier, logistic regression, k-nearest neighbor classifier, perceptron, multilayer perceptron, neural networks, Support Vector Machine (SVM). Classification evaluation metrics.

Essential/recommended readings

1. Mitchell, T.M. *Machine Learning*, McGraw Hill Education, 2017.
2. James, G., Witten. D., Hastie. T., Tibshirani., R. *An Introduction to Statistical Learning with Applications in R*, Springer, 2014.
3. Alpaydin, E. *Introduction to Machine Learning*, MIT press, 2009.

Additional References

1. Flach, P., *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*, Cambridge University Press, 2015.
2. Christopher & Bishop, M., *Pattern Recognition and Machine Learning*, New York: Springer-Verlag, 2016.
3. Sebastian Raschka, *Python Machine Learning*, Packt Publishing Ltd, 2019.

Suggested Practicals List : (30 Hours)

Practical exercise such as

Use Python for practical labs for Machine Learning. Utilize publically available datasets from online repositories like <https://data.gov.in/> and <https://archive.ics.uci.edu/ml/datasets.php>

For evaluation of the regression/classification models, perform experiments as follows:

- i. Scale/Normalize the data
- ii. Reduce dimension of the data with different feature selection techniques
- iv. Perform k-cross-validation on datasets for evaluation

Report the efficacy of the machine learning models as follows:

- i. MSE and R^2 score for regression models
- ii. Accuracy, TP, TN, FP, FN, error, Recall, Specificity, F1-score, AUC for classification models

For relevant datasets make prediction models for the following

1. Naïve Bayes Classifier
2. Simple Linear Regression multiple linear regression
3. Polynomial Regression
4. Lasso and Ridge Regression
5. Logistic regression
6. Artificial Neural Network
7. k -NN classifier
8. Decision tree classification
9. SVM classification
10. K-Means Clustering
11. Hierarchical Clustering