**FACULTY OF COMPUTER APPLICATIONS**
**Bachelor of Computer Applications**

---

- **Sem.** : 6
- **Subject Code** : 05BC3603
- **Subject** : Machine Learning
- **Course Objectives** :

1. To get the idea about concepts of machine learning and its applications
2. Understanding various algorithms of classification for supervised machine learning
3. Understanding various algorithms of regression for supervised machine learning
4. Understanding various algorithms of clustering and association rule mining for unsupervised machine learning

- **Prerequisites** : Concepts of statistics, basics of Python programming

| Unit No | Topics Covered | No of lectures required |
|---|---|---|
| **1** | Introduction<br>Introduction to machine learning and preparing a model Introduction, What is human learning? Types of machine learning, Applications of machine learning, Languages and tools in machine learning<br>Machine learning activities, Basic types of data, Exploring data structure, Data quality and remediation, Data pre-processing | **8** |
| **2** | Supervised learning – Classification<br>Example of supervised learning, Classification model, Classification learning steps<br>Importance of Bayesian method, Bayes theorem and concept learning<br>Common algorithms – K nearest neighbour (kNN), Decision tree, Random forest model, Support Vector | **10** |

| | | |
|---|---|---|
| | Machine (SVM) Evaluating performance of classification models | |
| **3** | Supervised learning – Regression<br>Introduction, Example of regression, Common regression algorithms – Simple linear, Multiple linear, Main problems in regression analysis, Improving accuracy of linear regression model, Polynomial regression model, Logistic regression, Maximum likelihood estimation<br>Evaluating performance of regression models | **6** |
| **4** | Unsupervised learning<br>Introduction, Unsupervised vs supervised learning, Applications, Clustering, Evaluating performance of clustering<br>Finding pattern using association rule mining | **6** |

**Course Outcomes :** Students will able to

1. Understand applications of ML in real life.
2. Apply classification algorithms for supervised learning
3. Apply regression algorithms for predictive analysis
4. Apply clustering algorithms and association rule mining algorithms for real life problems

Course Outcomes – Program Outcomes Mapping Table : (Change as per the program)

| | PO1 | PO2 | PO3 | PO4 | PO5 | PO6 | PO7 | PO8 | PSO1 | PSO2 | PSO3 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| CO1 | H | | L | | | M | | L | L | | L |
| CO2 | H | L | M | L | M | M | L | | M | L | M |
| CO3 | H | M | H | L | L | H | M | H | M | L | M |
| CO4 | H | M | H | L | L | H | M | H | M | L | H |

**Text Book :**

1. Machine Learning, Saikat Dutt et al. , Pearson Education, Third impression, 2019

**Reference Books :**

1. "Machine Learning", Tom M Mitchell, McGraw Hill, First Edition
2. "Machine Learning" , Anuradha Srinivarasa raghavan, Vincy Joseph,Wiley India, First Edition
3. "Machine Learning in Action",Peter Harrington, DreamTech, First Edition

**FACULTY OF COMPUTER APPLICATIONS**
**Bachelor of Computer Applications**

**Web References :**

1. www.edx.org
2. www. coursera.org
3. www.kaggle.com

**App References :**

1. Learn Machine Learning - ML Tutorials & Programs
2. Data Science 101 - Machine Learning Tutorials

**Syllabus Coverage from text /reference book & web/app reference:**

| Unit # | Chapter Numbers |
|---|---|
| 1 | 1,2 |
| 2 | 6.1,6.2,6.3,6.4, 7, 3.5.1 |
| 3 | 8, 3.5.2 |
| 4 | 9,3.5.3 |

**FACULTY OF COMPUTER APPLICATIONS**
**Bachelor of Computer Applications**

# PRACTICAL

<table>
<tr><th>Unit No.</th><th>List of Practicals</th></tr>
<tr><td>1 & 2</td><td>
Data Identification and Feature Engineering :<br><br>
Practical – 1 : Download Auto-MPG data set (https://www.kaggle.com/uciml/autompg-dataset) and perform following operations<br>
i. Write program to read dataset ( Text,CSV,JSON,XML)<br>
ii. Which of the attributes are numeric and which are categorical?<br>
iii. Performing Data Cleaning<br>
☐ Handling Missing Data<br>
☐ Removing Null<br>
data iv. Rescaling Data<br>
v. Encoding Data<br>
vi. Feature Selection and Dimensionality Reduction<br>
☐ Implement Principle Component Analysis,<br><br>
Pracitcal – 2 : Download Airline data set (https://www.kaggle.com/open- flights/airline-database) and perform all the above operations.
</td></tr>
<tr><td>3</td><td>
Supervised Learning – Classification :<br><br>
Practical – 1 : Write a python code to apply Naive Bayesian algorithm to classify that whether a person can buy computer or not based on given test data :
<table>
<tr><th>Age</th><th>Income</th><th>Student</th><th>Creditrating</th><th>Buyscomputer</th></tr>
<tr><td>Youth</td><td>High</td><td>No</td><td>Fair</td><td>No</td></tr>
<tr><td>Youth</td><td>High</td><td>No</td><td>Excellent</td><td>No</td></tr>
<tr><td>Middle</td><td>High</td><td>No</td><td>Fair</td><td>Yes</td></tr>
<tr><td>Senior</td><td>Medium</td><td>No</td><td>Fair</td><td>Yes</td></tr>
<tr><td>Senior</td><td>Low</td><td>Yes</td><td>Fair</td><td>Yes</td></tr>
<tr><td>Middle</td><td>Low</td><td>Yes</td><td>Excellent</td><td>No</td></tr>
<tr><td>Senior</td><td>Low</td><td>Yes</td><td>Excellent</td><td>Yes</td></tr>
</table>
</td></tr>
</table>

| | | | | | |
|---|---|---|---|---|---|
| | Youth | Medium | No | Fair | No |
| | Youth | Low | Yes | Fair | Yes |
| | Senior | Medium | Yes | Fair | Yes |
| | Youth | Medium | Yes | Excellent | Yes |

| | | | | |
|---|---|---|---|---|
| Middle | Medium | No | Excellent | Yes |
| Middle | High | Yes | Fair | Yes |
| Senior | Medium | No | Excellent | No |

Practical – 2 : Write a python code to implement decision tree for below given dataset. Identify Job offered or not.

Practical – 3 : Write a python code to implement K-nearest neighbourhood
program for the given dataset. (for above both the data sets)

Practical – 4 : Implement supervised machine learning algorithm (Classification
– K Nearest Neighbourhood) in python to classify breast tumour data into malignant breast tumour or benign breast tumour (use breast tumour dataset) and obtain its accuracy level.

Practical – 5 : Implement supervised machine learning algorithm (Classification

| | |
|---|---|
| | – K Nearest Neighbourhood) in python to classify iris data into setosa, virginica, versicolor using iris dataset and obtain its accuracy level. |

**FACULTY OF COMPUTER APPLICATIONS**
**Bachelor of Computer Applications**

| | |
|---|---|
| | Practical – 6 : Build a classification model in python that classifies if a student gets admission in a course or not given his last two examination scores for the dataset available at https://docs.google.com/spreadsheets/d/1g0mjTUZ9Ado5prXA1UnAvNjmdzT rV0TzkFkIoU-Lpbk/edit?usp=sharing<br><br>Practical – 7 : Implement supervised machine learning algorithm (Classification<br>– Support Vector Machine) in python to classify breast tumour data into malignant breast tumour or benign breast tumour (use breast tumour dataset) and obtain its accuracy level.<br><br>Practical – 8 : Write a python program to build an email spam classifier using support vector machines for the Spam base dataset from UCI machine learning repository.<br><br>Practical – 9 : Implement supervised machine learning algorithm (Classification - Naïve Bayes algorithm) in python on Pima Indians Diabetes dataset and obtain its accuracy level. |

**FACULTY OF COMPUTER APPLICATIONS**
**Bachelor of Computer Applications**

| | |
|---|---|
| 4 | Supervised Learning – Regression :<br><br>Pracitcal – 1 : Predict the CPU time if the DiskIO = 40 on the basis of following data :<br><br>Practical – 2 : Write a python code to predict profit of hotel chain given the population of the area (city) using the data at https://docs.google.com/spreadsheets/d/1Ks20skBgEefHFU36sFqVzozoFtz2E ZE2rxBIgXOrUg/edit?usp=sharing.<br><br>Practical – 3 : Write a python code to predict salary on the basis of experience in years using the data at https://github.com/tarunlnmiit/machine_learning/blob/master/SimpleLinearReg ression.csv |

| | |
|---|---|
| | Practical – 4 : Write a python code to predict the price of house given square feet and number of bed rooms in the house for the dataset available at https://docs.google.com/spreadsheets/d/1DHVK7gKo4TSyj7mFLwofHamj1Sl4SOZma2q51w1ZvyE/edit?usp=sharing<br><br>Practical – 5 : Build a logistic regression model to classify flower type based on the dataset of iris flower.<br><br>Practical – 6 : Build a multivariate logistic regression model to classify glass type of glass given different glass mixture features using the Glass Identification Dataset from UCI Machine Learning Repository. |
| 5 | Unsupervised Learning – Clustering and Association Rule Mining :<br><br>Practical – 1 : Implement unsupervised machine learning algorithm (Clustering – K Means) in python on Titanic dataset to cluster data (use Titanic dataset) by removing the class label.<br><br>Practical – 2 : Implement unsupervised machine learning algorithm (Clustering – K Means) in python on Breast Tumour dataset to cluster data (use Breast Tumour dataset) by removing the class label.<br><br>Practical – 3 : Implement unsupervised machine learning algorithm (Clustering – Hierarchical) in python on Titanic dataset to cluster data (use Titanic dataset).<br><br>Practical – 4 : Implement unsupervised machine learning algorithm (Clustering – Hierarchical) in python on Breast Tumour dataset to cluster data (use Breast Tumour dataset) by removing the class label.<br><br>Practical – 5 : Implement Apriori algorithm in python to find rules which explain association between different products for given transactions at a retail store. (The data is available at |

https://drive.google.com/file/d/1NUXoptUlHY8z4KcFKpFA6sQN5KnWzk3p/ view?usp=sharing )

Practical – 6 : Generating Association rule mining for following data set.

| | |
|---|---|
| 1 | Red,White,Green |
| 2 | White,Orange |
| 3 | White,Blue |
| 4 | Red,White,Orange |

| CGPA | Communication | Aptitude | Programming_skill | Job_offered |
|---|---|---|---|---|
| High | Good | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | Low | Good | No |
| Low | Good | Low | Bad | No |
| High | Good | High | Bad | Yes |
| High | Good | High | Good | Yes |
| Medium | Bad | Low | Bad | No |
| Medium | Bad | Low | Good | No |
| High | Bad | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | High | Bad | No |
| Low | Bad | High | Bad | No |
| Medium | Good | High | Bad | Yes |
| Low | Good | Low | Good | No |
| High | Bad | Low | Bad | No |

| CGPA | Communication | Aptitude | Programming_skill | Job_offered |
|---|---|---|---|---|
| High | Good | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | Low | Good | No |
| Low | Good | Low | Bad | No |
| High | Good | High | Bad | Yes |
| High | Good | High | Good | Yes |
| Medium | Bad | Low | Bad | No |
| Medium | Bad | Low | Good | No |
| High | Bad | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | High | Bad | No |
| Low | Bad | High | Bad | No |
| Medium | Good | High | Bad | Yes |
| Low | Good | Low | Good | No |
| High | Bad | Low | Bad | No |

| | | |
|---|---|---|
| | 5 | Red,Blue |

| CGPA | Communication | Aptitude | Programming_skill | Job_offered |
|---|---|---|---|---|
| High | Good | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | Low | Good | No |
| Low | Good | Low | Bad | No |
| High | Good | High | Bad | Yes |
| High | Good | High | Good | Yes |
| Medium | Bad | Low | Bad | No |
| Medium | Bad | Low | Good | No |
| High | Bad | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | High | Bad | No |
| Low | Bad | High | Bad | No |
| Medium | Good | High | Bad | Yes |
| Low | Good | Low | Good | No |
| High | Bad | Low | Bad | No |

| CGPA | Communication | Aptitude | Programming_skill | Job_offered |
|---|---|---|---|---|
| High | Good | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | Low | Good | No |
| Low | Good | Low | Bad | No |
| High | Good | High | Bad | Yes |
| High | Good | High | Good | Yes |
| Medium | Bad | Low | Bad | No |
| Medium | Bad | Low | Good | No |
| High | Bad | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | High | Bad | No |
| Low | Bad | High | Bad | No |
| Medium | Good | High | Bad | Yes |
| Low | Good | Low | Good | No |
| High | Bad | Low | Bad | No |

| | | |
|---|---|---|
| | 7 | White,Orange |

| CGPA | Communication | Aptitude | Programming_skill | Job_offered |
|---|---|---|---|---|
| High | Good | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | Low | Good | No |
| Low | Good | Low | Bad | No |
| High | Good | High | Bad | Yes |
| High | Good | High | Good | Yes |
| Medium | Bad | Low | Bad | No |
| Medium | Bad | Low | Good | No |
| High | Bad | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | High | Bad | No |
| Low | Bad | High | Bad | No |
| Medium | Good | High | Bad | Yes |
| Low | Good | Low | Good | No |
| High | Bad | Low | Bad | No |

| CGPA | Communication | Aptitude | Programming_skill | Job_offered |
|---|---|---|---|---|
| High | Good | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | Low | Good | No |
| Low | Good | Low | Bad | No |
| High | Good | High | Bad | Yes |
| High | Good | High | Good | Yes |
| Medium | Bad | Low | Bad | No |
| Medium | Bad | Low | Good | No |
| High | Bad | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | High | Bad | No |
| Low | Bad | High | Bad | No |
| Medium | Good | High | Bad | Yes |
| Low | Good | Low | Good | No |
| High | Bad | Low | Bad | No |

| | | |
|---|---|---|
| | 9 | Red,White,Blue |

| CGPA | Communication | Aptitude | Programming_skill | Job_offered |
|---|---|---|---|---|
| High | Good | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | Low | Good | No |
| Low | Good | Low | Bad | No |
| High | Good | High | Bad | Yes |
| High | Good | High | Good | Yes |
| Medium | Bad | Low | Bad | No |
| Medium | Bad | Low | Good | No |
| High | Bad | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | High | Bad | No |
| Low | Bad | High | Bad | No |
| Medium | Good | High | Bad | Yes |
| Low | Good | Low | Good | No |
| High | Bad | Low | Bad | No |

| | | |
|---|---|---|
| | 10 | Yellow |

| CGPA | Communication | Aptitude | Programming_skill | Job_offered |
|---|---|---|---|---|
| High | Good | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | Low | Good | No |
| Low | Good | Low | Bad | No |
| High | Good | High | Bad | Yes |
| High | Good | High | Good | Yes |
| Medium | Bad | Low | Bad | No |
| Medium | Bad | Low | Good | No |
| High | Bad | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | High | Bad | No |
| Low | Bad | High | Bad | No |
| Medium | Good | High | Bad | Yes |
| Low | Good | Low | Good | No |
| High | Bad | Low | Bad | No |

| CGPA | Communication | Aptitude | Programming_skill | Job_offered |
|---|---|---|---|---|
| High | Good | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | Low | Good | No |
| Low | Good | Low | Bad | No |
| High | Good | High | Bad | Yes |
| High | Good | High | Good | Yes |
| Medium | Bad | Low | Bad | No |
| Medium | Bad | Low | Good | No |
| High | Bad | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | High | Bad | No |
| Low | Bad | High | Bad | No |
| Medium | Good | High | Bad | Yes |
| Low | Good | Low | Good | No |
| High | Bad | Low | Bad | No |

| CGPA | Communication | Aptitude | Programming_skill | Job_offered |
|---|---|---|---|---|
| High | Good | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | Low | Good | No |
| Low | Good | Low | Bad | No |
| High | Good | High | Bad | Yes |
| High | Good | High | Good | Yes |
| Medium | Bad | Low | Bad | No |
| Medium | Bad | Low | Good | No |
| High | Bad | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | High | Bad | No |
| Low | Bad | High | Bad | No |
| Medium | Good | High | Bad | Yes |
| Low | Good | Low | Good | No |
| High | Bad | Low | Bad | No |

| CGPA | Communication | Aptitude | Programming_skill | Job_offered |
|---|---|---|---|---|
| High | Good | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | Low | Good | No |
| Low | Good | Low | Bad | No |
| High | Good | High | Bad | Yes |
| High | Good | High | Good | Yes |
| Medium | Bad | Low | Bad | No |
| Medium | Bad | Low | Good | No |
| High | Bad | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | High | Bad | No |
| Low | Bad | High | Bad | No |
| Medium | Good | High | Bad | Yes |
| Low | Good | Low | Good | No |
| High | Bad | Low | Bad | No |

| CGPA | Communication | Aptitude | Programming_skill | Job_offered |
|---|---|---|---|---|
| High | Good | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | Low | Good | No |
| Low | Good | Low | Bad | No |
| High | Good | High | Bad | Yes |
| High | Good | High | Good | Yes |
| Medium | Bad | Low | Bad | No |
| Medium | Bad | Low | Good | No |
| High | Bad | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | High | Bad | No |
| Low | Bad | High | Bad | No |
| Medium | Good | High | Bad | Yes |
| Low | Good | Low | Good | No |
| High | Bad | Low | Bad | No |

data set.

| CGPA | Communication | Aptitude | Programming_skill | Job_offered |
|---|---|---|---|---|
| High | Good | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | Low | Good | No |
| Low | Good | Low | Bad | No |
| High | Good | High | Bad | Yes |
| High | Good | High | Good | Yes |
| Medium | Bad | Low | Bad | No |
| Medium | Bad | Low | Good | No |
| High | Bad | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | High | Bad | No |
| Low | Bad | High | Bad | No |
| Medium | Good | High | Bad | Yes |
| Low | Good | Low | Good | No |
| High | Bad | Low | Bad | No |

| CGPA | Communication | Aptitude | Programming_skill | Job_offered |
|---|---|---|---|---|
| High | Good | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | Low | Good | No |
| Low | Good | Low | Bad | No |
| High | Good | High | Bad | Yes |
| High | Good | High | Good | Yes |
| Medium | Bad | Low | Bad | No |
| Medium | Bad | Low | Good | No |
| High | Bad | High | Good | Yes |
| Medium | Good | High | Good | Yes |
| Low | Bad | High | Bad | No |
| Low | Bad | High | Bad | No |
| Medium | Good | High | Bad | Yes |
| Low | Good | Low | Good | No |
| High | Bad | Low | Bad | No |