SREE VIDYANIKETHAN ENGINEERING COLLEGE
(AUTONOMOUS)
Sree Sainath Nagar, A. Rangampet-517 102
MASTER OF COMPUTER APPLICATION
MCA III – Semester
14MC3HS01: ORGANIZATIONAL BEHAVIOUR AND HUMAN RESOURCE MANAGEMENT

<table>
<thead>
<tr>
<th>Int. Marks</th>
<th>Ext. Marks</th>
<th>Total Marks</th>
<th>L</th>
<th>T</th>
<th>P</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>60</td>
<td>100</td>
<td>3</td>
<td>1</td>
<td>-</td>
<td>3</td>
</tr>
</tbody>
</table>

PRE-REQUISITES:

COURSE OBJECTIVES:

I. To impart basic concepts of organizational behavior, Human Resource Management.
II. To develop skills related to planning and management of Human Resources.
III. To apply the concepts & theory of personality for effective Human Resource Management.

COURSE OUTCOMES:

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

1. Gain Knowledge in
2. Develop skills of effective Human Resource Management.
4. Inculcates ethical principles and commitment to Professionalism.
5. Develops effective communication among the work group of an Organization.
6. Enriches empathy and enables lifelong human relations.
7. Provides a platform to emerge as a responsible individual in an Organization.
8. Provides life skills for effective operation of an organization.

DETAILED SYLLABUS

UNIT–I: INTRODUCTION TO ORGANIZATIONAL BEHAVIOR:
Introduction to organization –Concepts of organization – meaning of Organizational Behavior(Ob), characteristics of Ob, role of Ob, approaches to the study of Ob.
UNIT-II: PERSONALITY, PERCEPTION AND LEARNING
PERSONALITY: concept, determinants of personality, theories of personality, Organizational applications of personality.
PERCEPTION: Process of perception, inter personal perception, managerial applications of perception.
LEARNING: Components of learning process, Importance of learning, Factors affecting learning

UNIT-III: INTRODUCTION TO HRM - HRM: Functions and objectives of HRM
HR PLANNING: Nature and importance of HRP, factors affecting HRP, HRP Process
RECRUITMENT & SELECTION: Nature and importance of recruitment, recruitment process, selection process, barriers to effective selection.

UNIT-IV: INDUCTION, TRAINING AND PERFORMANCE APPRAISAL
INDUCTION: Induction, Orientation Programme, Problems of Orientation
TRAINING: Nature of training and development, gaps in training, the training process, training methods, training for career development
PERFORMANCE APPRAISAL: Managing Performance, Appraisal Process, Job Evaluation Process, Methods of Evaluation

UNIT-V: REMUNERATION, EMPLOYEE PARTICIPATION AND CONTEMPORARY ISSUES
REMUNERATION: Components of Remuneration, Factors influencing employee remuneration-Incentives, Types of incentives-Incentive Schemes-Employee Benefits and Services
EMPLOYEE PARTICIPATION: Empowering employees through participation-Importance and Limitations

TEXT BOOKS:

REFERENCES:
MCA III-SEMESTER
14MC3BS01: OPERATIONS RESEARCH

<table>
<thead>
<tr>
<th>Int. Marks</th>
<th>Ext. Marks</th>
<th>Total Marks</th>
<th>L</th>
<th>T</th>
<th>P</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>60</td>
<td>100</td>
<td>3</td>
<td>1</td>
<td>-</td>
<td>3</td>
</tr>
</tbody>
</table>

PRE-REQUISITES:

COURSE OBJECTIVES:

I. To create awareness, about optimization in utilization of resources.
II. To understand nuance of project management through operational models.
III. To apply Operations research techniques like Linear Programming Problems, Transportation problem, Assignment problem, Replacement problem and PERT/CPM in Research and Industrial operations.

COURSE OUTCOMES:

On successful completion of this course, the students will be able to
1. Apply the knowledge of Linear Programming Problem, Transportation Problem, Assignment Problem, Inventory models PERT/CPM etc., in the fields of business research and industry.
2. Formulate research literature and solve complex computational problems using Linear Programming Problem.
3. Design and development a component or a process that meet a specific need of application software.
4. Analyze and design the data, to synthesize transformation by using operational models like Transportation Problem, Assignment Problem, Sequencing Problem etc.,
5. Demonstrate the knowledge and Management principles by applying PERT/ CPM methods.

DETAILED SYLLABUS

UNIT-I: INTRODUCTION TO OPERATIONS RESEARCH AND LPP (10 Periods)
Introduction to OR: Introduction, modeling in OR- Phases of OR study. Linear Programming: Formulation of LPP, Graphical solution of LPP, Simplex method, Artificial variable technique-Big M-method.

UNIT-II: TRANSPORTATION AND ASSIGNMENT PROBLEM (12 Periods)
Transportation Problem: Finding an initial basic feasible solution using North-West corner rule, Least cost Entry method, Vogel’s Approximation Method. Degeneracy in Transportation Problem, Optimality test- MODI method Unbalanced Transportation Problem.
Assignment Problem: Hungarian method of Assignment Problem, Traveling salesman Problem and its restrictions.
UNIT-III: SEQUENCING PROBLEM AND REPLACEMENT PROBLEM
(10 Periods)

Sequencing Problem: Optimal solution for processing n-jobs through two machines, n-jobs through three machines.

Replacement Problem: Introduction, Replacement of items that deteriorate when money value is constant and variable- Individual Replacement policy and group Replacement policy.


UNIT-V: INVENTORY MODELS AND PROJECT MANAGEMENT BY PERT/CPM (13 Periods)

Total Periods: 55

TEXT BOOKS:

REFERENCE BOOKS:
MCA III – Semester
14MC30101: OBJECT ORIENTED ANALYSIS AND DESIGN

<table>
<thead>
<tr>
<th>Int. Marks</th>
<th>Ext. Marks</th>
<th>Total Marks</th>
<th>L</th>
<th>T</th>
<th>P</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>60</td>
<td>100</td>
<td>3</td>
<td>1</td>
<td>-</td>
<td>3</td>
</tr>
</tbody>
</table>

PRE-REQUISITES:
Courses on “Object Oriented Programming Through JAVA” and “Software Engineering”

COURSE OBJECTIVES:
I. To understand the principles of object orientation.
II. To understand the behavior of an Application Software before deployment by using State Chart Diagrams.
III. To analyze logical, behavioral and architectural models and apply for real world applications.

COURSE OUTCOMES:
After successful completion this course, the student will be able to:
1. Acquire the principles of Object Oriented Development.
2. Recognize the elements of Class, Use case, Activity, Sequence and State, diagrams and develop models using them.
3. Design the domain and application artifacts to construct the Logical, Behavioral and Architectural model of an Application Software.
4. Make use of UML Tool, Rational Rose to design Class, Use Case, Sequence, Collaboration, Activity, State Chart, Component and Deployment Diagrams for the an Application Software.
5. Communicate effectively with all the team members about various logical and behavioral objects of an Application Software.
6. Function effectively as an individual team member or leader.

DETAILED SYLLABUS
UNIT -I: INTRODUCTION TO UML
The meaning of Object Orientation, object identity, Importance of modeling, principles of modeling, object oriented modeling, An overview of UML, conceptual model of the UML, Architecture.

Classes - Terms and concepts, Common Modeling Techniques.
Relationships - Modeling simple dependencies, single Inheritance and structural relationships, Common Mechanisms and UML Diagrams.

UNIT-III: STRUCTURAL MODELING
Advanced classes, advanced relationships, Interfaces, Types and Roles, Packages, Instances.

Class Diagrams - Terms, concepts, Modeling techniques for Class Diagram, Modeling Simple collaboration, Logical database Schema, Forward and Reverse Engineering.

Object Diagrams – Modeling object structures, Forward and Reverse engineering.


Total Periods: 52

TEXT BOOK:

REFERENCE BOOKS:
MCA III – Semester
14MC30102: DATA WAREHOUSING AND DATA MINING

<table>
<thead>
<tr>
<th>Int. Marks</th>
<th>Ext. Marks</th>
<th>Total Marks</th>
<th>L</th>
<th>T</th>
<th>P</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>60</td>
<td>100</td>
<td>3</td>
<td>1</td>
<td>-</td>
<td>3</td>
</tr>
</tbody>
</table>

PRE-REQUISITES:
A course on "Data Base Management Systems"

COURSE OBJECTIVES:
I. To understand the concepts of Data Warehouse system, Data Warehouse system architecture, data mining principles and techniques.
II. To analyze and demonstrate data mining algorithms like Apriori, FP-growth, k-mediod and gaining better understanding of business trends.
III. To be relevant in applying critical thinking, problem-solving, and decision-making skills effectively.

COURSE OUTCOMES:
On successful completion of this course the students will be able to:
1. Understand the role of data warehouse, architecture and its support for quality data.
2. Ability to identify the association rules, classifying by using decision tree algorithms and clustering large data sets using k-mediods, k-means, etc.
3. Design and implement Hierarchical Methods-Agglomerative clustering, divisive hierarchical clustering and Classification by Decision Tree to perform proactive analysis, predictive modeling, and identifying new trends and behaviors.
4. Investigate preprocessing of data and apply mining techniques on it.
5. Apply data mining algorithms like J48,ID3,Random forest and evaluate results by using tool WEKA.

DETAILED SYLLABUS
UNIT-I: INTRODUCTION AND DATA WAREHOUSE COMPONENTS (10 periods)
Introduction: The need for Data Warehousing, Paradigm Shift, Business Problem Definition, operational and informational Data Stores, Data Warehouse Definition and Characteristics, Data Warehouse Architecture.
Data Warehouse Components: Overall Architecture, Data Warehouse Database, Sourcing, Acquisition , Cleanup and Transformation tools, meta data, data marts, Data Warehouse Administration and Management.

UNIT-II: BUILDING A DATA WAREHOUSE AND INTRODUCTION TO DATA MINING (12 periods) Building A Data Warehouse:
Business Consideration, Design considerations, Technical considerations, Implementation considerations,
Integrated solutions, Benefits of Data Warehousing, Multidimensional Data Model-From tables and spreadsheets to Data Cubes & Star, Snowflake and fact constellation Schemas.

**Introduction To Data Mining:** Motivated Data Mining, Definition of Data Mining, Data Mining-On What Kind of Data?, Data mining Functionalities, classification of Data mining systems, Data mining primitives, Integration of Data mining Systems with a Database or Data Warehouse System, Major issues in Data Mining.

**UNIT-III: DATA PREPROCESSING AND ASSOCIATION RULE MINING (11 periods)**

**Data Preprocessing:** Need for Preprocessing the Data, Data Cleaning, Data Integration and Transformation, Data Reduction, Discretization and Concept Hierarchy Generation.

**Association Rule Mining:** Attribute Oriented Introduction, Mining Frequent Patterns, Associations and Correlations- Basic Concepts, The Apriori algorithm for finding frequent itemsets using candidate generation, Generating association rules from frequent itemsets, Mining frequent itemsets without candidate generation.

**UNIT-IV: CLASSIFICATION AND CLUSTERING (11 periods)**

**Classification:** Definition of classification, Definition of prediction, issues in classification and prediction, Classification by Decision Tree Induction, Bayesian Classification, Rule-Based Classification, Classification by Backpropagation.

**Clustering:** Introduction to cluster Analysis, Types of Data in Cluster Analysis, A Categorization of Major Clustering Methods, k-means and k-mediods methods, CLARANS

**UNIT-V: ADVANCES IN CLUSTERING, MULTIMEDIA, TEXT & WEB DATA MINING AND APPLICATIONS (11 periods)**

**Advances in Clustering:** Hierarchical Methods-Agglomerative and divisive hierarchical clustering, Constraint-Based Cluster Analysis, Outlier Analysis.

**Multimedia, Text and Web Data Mining:** Multimedia Data Mining, Text Mining, Mining the World Wide Web.

**Data Mining Applications:** Financial data Analysis, Retail Industry, Telecommunication Industry.

**Total Periods:** 55

**TEXT BOOKS:**
1. Jiawei Han, Micheline Kamber and Jian Pei, "Data Mining–Concepts and Techniques," 2nd Edition, Morgan Kaufmann Publishers, 2006

**REFERENCE BOOKS:**

SVEC14 - Master of Computer Applications
MCA III-Semester
14MC30103: COMPUTER NETWORKS

<table>
<thead>
<tr>
<th>Int. Marks</th>
<th>Ext. Marks</th>
<th>Total Marks</th>
<th>L</th>
<th>T</th>
<th>P</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>60</td>
<td>100</td>
<td>3</td>
<td>1</td>
<td>-</td>
<td>3</td>
</tr>
</tbody>
</table>

PRE-REQUISITES

COURSE OBJECTIVES:

I. To acquire basic knowledge of network components, topologies, network models, protocols and algorithms.

II. To comprehend in-depth knowledge about the communication technologies like Domain Name Service, E-mail.

III. To analyze techniques in designing network applications.

COURSE OUTCOMES:

After successful completion of the course, the student able to

1. Understand the concept of networks, different topologies, network devices and OSI reference model.

2. Analyze Error detection and correction mechanisms to verify correct data.

3. Investigate congestion handling mechanisms and network security.

4. Design and develop a route to reach to the correct destination by using shortest path routing and Distance Vector Routing Techniques.

DETAILED SYLLABUS

UNIT- I: INTRODUCTION & PHYSICAL LAYER (12 Periods)


UNIT-III: THE NETWORK LAYER  (11 Periods)

UNIT-IV: THE TRANSPORT LAYER  (10 Periods)

UNIT-V: THE APPLICATION LAYER AND NETWORK SECURITY (10 Periods)
The Application Layer: Introduction to Application Layer, DNS-The Domain name space, Resource records and Name servers. Electronic Mail-Architecture and services, the user agent, message formats, message transfer and Final Delivery.

Total Periods: 55

Text Book:

Reference Books:
MCA – III Semester
14MC30121: OBJECT ORIENTED ANALYSIS AND DESIGN & NETWORKS LAB

<table>
<thead>
<tr>
<th>Marks</th>
<th>Marks</th>
<th>Total Marks</th>
<th>L</th>
<th>T</th>
<th>P</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>50</td>
<td>75</td>
<td></td>
<td></td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

PRE-REQUISITES:
Courses on “Object Oriented Programming Through JAVA” and “Software Engineering”

COURSE OBJECTIVES:
I. To understand the principles of object orientation.
II. To understand the behavior of an Application Software before deployment by using State Chart Diagrams.
III. To analyze logical, behavioral and architectural models and apply for real world applications.
IV. To comprehend in-depth knowledge about the communication technologies like Domain Name Service, E-mail.
V. To analyze techniques to design simple network.

COURSE OUTCOMES:
After successful completion this course, the student will be able to:
1. Acquire the principles of Object Oriented Development.
2. Recognize the elements of Class, Use case, Activity, Sequence and State, diagrams and develop models using them.
3. Design the domain and application artifacts to construct the Logical, behavioral and Architectural model of an Application Software.
4. Make use of UML Tool, Rational Rose to design Class, Use Case, Sequence, Collaboration, Activity, State Chart, Component and Deployment Diagrams for the an Application Software.
5. Communicate effectively with all the team members about various logical and behavioral objects of an Application Software.
6. Function effectively as an individual team member or leader.
7. Understand the concept of networks, different topologies, network devices and OSI reference model.
8. Analyze Error detection and correction mechanisms to verify correct data.
9. Design and develop a route to reach to the correct destination by using shortest path routing and Distance Vector Routing Techniques.

LIST OF LAB PROGRAMMS:
1. a) Identification of Classes, Objects, Attributes, Operations and Subclasses
   From your everyday experience identify 2 original classes of objects. Choose your second one carefully, with some computerized system in mind that it could fit into.
For each class draw a graphical representation encapsulating seven major attributes and seven operations that apply to it. For each of your classes state at least four examples or instantiations of it.

b) Introduction of Inheritance

Choose a third class of object of any type suitable as a base class for a hierarchy of subclasses which will inherit all the attributes and operations of their super classes.

In a tree diagram draw this class and (some of) its subclasses, at least one of which must be three levels of hierarchy below the base class. Include significant attributes and operations.

1. Write a program for error detecting code using CRC-CCITT (16-bits).

2. Write a program for frame sorting technique used in buffers.

3. Noun-Verb Parsing

Apply a noun-verb parse to the following passage to identify classes and operations. Make a table listing the classes with their operations alongside, and use it to identify any subclasses which are present.

The Blueberry Muffin Factory buys in raw ingredients, including flour, milk, eggs and blueberries. All the ingredients except the blueberries are made into a batter. The blueberries are washed and sorted (to remove any bad berries or foreign objects) and then added to the batter. The batter is then poured into baking tins and baked. The baked muffins are cooled and packaged. The packages are put in boxes and finally distributed.

4. Write a program for distance vector algorithm to find suitable path for transmission.

5. a) Understanding the Logical View of the Application: Library Management System

   a) Identification of Objects
   b) Identification of Attributes
   c) Identification of Behaviors
   d) Identification of Relationships

The LMS would be used by the Librarian to keep track of books, Library Members and Borrowing activities. All members cannot borrow all books; few books may be available to review in the Library itself. These books are called as Reference Books.

There are two kinds of Library Members namely, Student Members and Faculty Members. Faculty Members can check out Research papers and Magazines, where as a Student Member can check out only Books. The system need to send an Alert, whenever a book has not returned within a specific due date to both the Librarian and the Borrower through e-mail.

SVEC14 - Master of Computer Applications
The system maintains a Catalogue of having a description of each book available in the Library.

b) Developing a Class Diagram: Library Management System
   a) Development of Class Diagram
   b) Applying Forward Engineering and Reverse Engineering of a Class Diagram.

1. Using TCP/IP sockets, write a client-server program to make client sending the file name and the server to send back the contents of the requested file if present.

2. Understanding the Interaction between Objects: Library Management System / E-Banking System
   a) Identification of Objects
   b) Identification of Messages
   c) Development of Sequence Diagrams
   d) Development of Collaboration Diagrams

3. Write a program for Hamming Code generation for error detection and correction.

4. Write a program for congestion control using Leaky bucket algorithm

5. Understanding the Advanced Behavioral Model of the Application: Library Management System
   a) Identification of Actions
   b) Identification of Activities
   c) Development of Activity Diagram
   d) Development of State Chart Diagram

6. Design TCP concurrent Client and Server application to reverse the given input sentence

7. Understanding the Architectural Model of the Application: Library Management System
   a) Identification of Components
   b) Development of Component Diagram
   c) Modeling a Library Management System as Distributed System

REFERENCE BOOKS:
2. www.uml.org
# MCA III – Semester
## 14MC30122: DATA WAREHOUSING AND DATA MINING LAB

<table>
<thead>
<tr>
<th></th>
<th>Int. Marks</th>
<th>Ext. Marks</th>
<th>Total Marks</th>
<th>L</th>
<th>T</th>
<th>P</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25</td>
<td>50</td>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**PRE-REQUISITES:**
A course on “Data Base Management Systems”

**COURSE OBJECTIVES:**
1. To understand the concepts of Data Warehouse system, Data Warehouse system architecture, data mining principles and techniques.
2. To analyze and demonstrate data mining algorithms like Apriori, FP-growth, k-medioids and gaining better understanding of business trends.
3. To develop and apply critical thinking, problem-solving, and decision-making skills.

**COURSE OUTCOMES:**
On successful completion of this course the students will be able to:
1. Understand the role of data warehouse, architecture and its support for quality data.
2. Ability to identify the association rules, classifying by using decision tree algorithms and clustering large data sets using k-medioids, k-means, etc.
3. Design and implement Hierarchical Methods-Agglomerative clustering, divisive hierarchical clustering and Classification by Decision Tree to perform proactive analysis, predictive modeling, and identifying new trends and behaviors.
4. Investigate preprocessing of data and apply mining techniques on it.
5. Apply data mining algorithms like J48, ID3, Random forest and evaluate results by using tool WEKA.

**LIST OF LAB PROGRAMS**
Using Open Source for Data Warehousing (Business Intelligence)
Integrating data from different databases to Extracting, Transformation and Loading

1. To perform various commands given in PL/SQL in Oracle 8.0(For brushing up.)
2. Knowing IDE of Open Source for Data Warehousing (Business Intelligence) for Integrating data from different databases to Extracting, Transformation and Loading
3. a. Construct a data acquisition process to perform the Source to target.
   b. Construct a data acquisition process to perform the Sorter Transformation and Aggregator Transformation form one data source to target.

4. a. Construct a data acquisition process to perform the Filter Transformation form one data source to target.
   b. Construct a data acquisition process to perform the Joiner Transformation from oracle data source to target.
   c. Construct a data acquisition process to perform the Aggregator Transformation form one data source to target.

5. a. Construct a data acquisition process to perform the Filter Transformation and Expression Transformation in source Qualifier form one data source to target.
   b. Construct a data acquisition process to perform the Joiner Transformation in source qualifier form one data source to target.

6. a. Construct a data acquisition process to perform the Router Transformation form one data source to a minimum of three target tables.
   b. Construct a data acquisition process to perform the Ranker Transformation form one data source to target.
   c. Construct a data acquisition process to perform the Expression Transformation form one data source to target.

Learn to perform data mining tasks using a data mining toolkit (such as open source WEKA) Understand the data sets and data preprocessing, Demonstrate the working of algorithms for data mining tasks such association rule mining, classification, clustering and regression, Exercise the data mining techniques with varied input values for different parameters. To obtain practical experience using data mining techniques on real world data sets. Emphasize hands-on experience working with all real data sets.

7. Introduction to the Weka machine learning toolkit
   a. Create a Buys_Computer dataset in .arff format. The following are the fields: age, income, student, credit_rating, buys_computer
   b. Create a Student Results Data set in .CSV format (Minimum required fields should be seven) Credit Risk Assessment- Data mining case study
The business of banks is making loans. Assessing the credit worthiness of an applicant is of crucial importance. You have to develop a system to help a loan officer decide whether the credit of a customer is good, or bad. A bank's business rules regarding loans must consider two opposing factors. On the one hand, a bank wants to make as many loans as possible. Interest on these loans is the bank's profit source. On the other hand, a bank cannot afford to make too many bad loans. Too many bad loans could lead to the collapse of the bank. The bank's loan policy must involve a compromise: not too strict, and not too lenient. To do the assignment, you first and foremost need some knowledge about the world of credit. You can acquire such knowledge in a number of ways.

1. Knowledge Engineering. Find a loan officer who is willing to talk. Interview her and try to represent her knowledge in the form of production rules.

2. Books. Find some training manuals for loan officers or perhaps a suitable textbook on finance. Translate this knowledge from text form to production rule form.

3. Common sense. Imagine yourself as a loan officer and make up reasonable rules which can be used to judge the credit worthiness of a loan applicant.

4. Case histories. Find records of actual cases where competent loan officers correctly judged when, and when not to, approve a loan application.

The German Credit Data:

Actual historical credit data is not always easy to come by because of confidentiality rules. Here is one such dataset, consisting of 1000 actual cases collected in Germany. credit dataset (original) Excel spreadsheet version of the German credit data.

In spite of the fact that the data is German, you should probably make use of it for this assignment. (Unless you really can consult a real loan officer!)

A few notes on the German dataset:

- DM stands for Deutsche Mark, the unit of currency, worth about 90 cents Canadian (but looks and acts like a quarter).
- owns_telephone. German phone rates are much higher than in Canada so fewer people own telephones.
- foreign_worker. There are millions of these in Germany (many from Turkey). It is very hard to get German citizenship if you were not born of German parents.
- There are 20 attributes used in judging a loan applicant. The goal is to classify the applicant into one of two categories, good or bad.
(Turn in your answers to the following tasks)

- List all the categorical (or nominal) attributes and the real-valued attributes separately.
- What attributes do you think might be crucial in making the credit assessment? Come up with some simple rules in plain English using your selected attributes.
- One type of model that you can create is a Decision Tree - train a Decision Tree using the complete dataset as the training data. Report the model obtained after training.
- Suppose you use your above model trained on the complete dataset, and classify credit good/bad for each of the examples in the dataset. What % of examples can you classify correctly? (This is also called testing on the training set) Why do you think you cannot get 100 % training accuracy?
- Is testing on the training set as you did above a good idea? Why or Why not One approach for solving the problem encountered in the previous question is using cross validation?
- Describe what cross-validation is briefly. Train a Decision Tree again using cross validation and report your results. Does your accuracy increase/decrease? Why?
- Check to see if the data shows a bias against "foreign workers" (attribute 20), or "personal-status" (attribute 9). One way to do this (perhaps rather simple minded) is to remove these attributes from the dataset and see if the decision tree created in those cases is significantly different from the full dataset case which you have already done. To remove an attribute you can use the preprocess tab in Weka’s GUI Explorer. Did removing these attributes have any significant effect? Discuss.
- Another question might be, do you really need to input so many attributes to get good results?
- Maybe only a few would do. For example, you could try just having attributes 2, 3, 5, 7, 10, 17 (and 21, the class attribute (naturally)). Try out some combinations. (You had removed two attributes in problem Remember to reload the arff data file to get all the attributes initially before you start selecting the ones you want.)

8.

III. Performing data preprocessing for data mining in Weka

JJJ. Sometimes, the cost of rejecting an applicant who actually has a good credit (case 1) might be higher than accepting an applicant who has bad credit (case 2). Instead of counting the mis-classifications equally in both cases, give a higher cost to the first case (say cost 5) and lower cost to the second case. You can do this by using a cost matrix in Weka. Train your Decision Tree again and report the Decision Tree and cross-validation results. Are they significantly different from results obtained in problem 6 (using equal cost)?
Do you think it is a good idea to prefer simple decision trees instead of having long complex decision trees? How does the complexity of a Decision Tree relate to the bias of the model?

You can make your Decision Trees simpler by pruning the nodes. One approach is to use

Reduced Error Pruning - Explain this idea briefly. Try reduced error pruning for training your Decision Trees using cross-validation (you can do this in Weka) and report the Decision Tree you obtain? Also, report your accuracy using the pruned model. Does your accuracy increase?

(Extra Credit): How can you convert a Decision Trees into “if-then-else rules”. Make up your own small Decision Tree consisting of 2-3 levels and convert it into a set of rules. There also exist different classifiers that output the model in the form of rules - one such classifier in Weka is rules. PART, train this model and report the set of rules obtained. Sometimes just one attribute can be good enough in making the decision, yes, just one! Can you predict what attribute that might be in this dataset? OneR classifier uses a single attribute to make decisions (it chooses the attribute based on minimum error). Report the rule obtained by training a one R classifier. Rank the performance of j48, PART and oneR.

9. Classification using the Weka toolkit
   ü Classify a weather nominal dataset using J48 algorithm.
   ü Classify a weather nominal dataset using ID3 algorithm.


11. Association rule analysis in Weka.


Reference Books:

Task Resources:
• Mentor lecture on Decision Trees
• Andrew Moore’s Data Mining Tutorials (See tutorials on Decision Trees and Cross Validation)
• Decision Trees (Source: Tan, MSU)
• Tom Mitchell’s book slides (See slides on Concept Learning and Decision Trees)
• Weka resources:
  Ø Introduction to Weka (html version) (download ppt version)
  Ø Download Weka
  Ø Weka Tutorial
  Ø ARFF format
  Ø Using Weka from command line
MCA III - Semester
14MCHS02: ENGLISH LANGUAGE AND
COMMUNICATION SKILLS LAB

<table>
<thead>
<tr>
<th>Int. Marks</th>
<th>Ext. Marks</th>
<th>Total Marks</th>
<th>L</th>
<th>T</th>
<th>P</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>3</td>
<td>--</td>
</tr>
</tbody>
</table>

PRE-REQUISITES:

COURSE OBJECTIVES:

I. To impart practical knowledge in English Speech Sounds.
II. To develop language skills for effective communication with clarity and precision in academic, professional and personal situations.
III. To apply the practical knowledge of Functional English and vocabulary building in effective writing.
IV. To develop interest in English language so that the students use it effectively in various formal, informal and neutral situations.

COURSE OUTCOMES:

After completion of this course, a successful student will be able to:

1. Acquire Knowledge in various web related concepts and technologies like client side scripting, validation of forms and AJAX programming.
2. Ability to analyze user requirements and selection of suitable web components for the development of web applications.
3. Apply prototyping technique and architectures to design dynamic, scalable, platform independent and enterprise web applications.
4. Demonstrate skills by applying security and management principles for effective implementation of web applications.
5. Adapt tools like Apache Tomcat Server and XAMPP Control Panel for implementing Web Applications.
6. Update their knowledge about rapidly evolving programming languages by going through books and Internet.
7. Recognize and understand ways of using electronic commerce technologies to improve intra and inter-organizational processes.
8. Demonstrate effective use of Dreamweaver to build and publish professional websites that employ best practices, adhere to current web standards, and pass Validation.
9. Develop simple e-commerce applications for specific areas of business that had potential commercialization values.

LIST OF LAB PRACTICE SESSIONS

1. Introduction to Phonetics
2. Accent Rhythm and Intonation
3. Vocabulary Building
4. Idioms and Phrases
5. Functional English - Tenses and Voice
6. Conversation Practice / Role Plays
7. Just a Minute / Elocution
8. Public Speaking
9. Presentation Skills
10. Listening skills

Total Periods: 48 hours

REFERENCES
1. Departmental Lab Manual

SUGGESTED SOFTWARE
2. Dorling Kindersley Series of Grammar, Punctuation, Composition etc.
3. Language in use 1, 2 & 3
5. English in mind, Herbert Puchta and Jeff Stranks with Meredith Levy, Cambridge.
7. Speech Solutions
9. Centronix - Phonetics
10. Rosetta Stone
11. Let’s Talk English, Regional Institute of English South India