ARUNACHAL UNIVERSITY OF STUDIES

Established by Govt. of Arunachal Pradesh vide Act 9 of 2012, the Arunachal University of Studies Act, 2012 Recognized as per u/s 2(f) of University Grants Commission Act, 1956 NH-52, Namsai, Arunachal Pradesh -792103

BACHELOR OF COMPUTER APPLICATION MASTER OF COMPUTER APPLICATION INTEGRATED – SEMESTER EIGHT

Eighth Semester			
S. No.	**Name of Subject**	**Credits**	**Total Marks**
1	Data Mining & Warehousing	5	100
2	Compiler Design	5	100
3	Analysis & Design of Algorithms	5	100
4	Minor Project Work	7	100
Total		**22**	

Subject Name: DATA MINING & WAREHOUSING

Unit I: Introduction

Introduction to data mining and knowledge discovery from databases. Scalability issues of data mining algorithms.

Introduction to Data warehousing: General principles, modeling, design, implementation, and optimization.

Data preparation: Preprocessing, sub-sampling, feature selection.

Unit II: Classification and prediction

Bayes learning, discriminant analysis, decision trees, CART, C4.5 etc, neural learning, support vector machines, active learning. Combination of classifiers/ ensemble learning. Associations, dependence analysis, correlation, rule generation – apriori algorithm, FP Trees etc. and evaluation.

Unit III: Cluster analysis and deviation detection

Partitioning algorithms, density based algorithms, hierarchical algorithms, model based algorithms, grid based algorithms, graph theoretic clustering etc.

Temporal and spatial data mining: Mining complex types of data. Visualization of data mining results.

Unit IV: Advanced topics

High performance computing for data mining, distributed data mining, soft-computing tools for data mining.

Applications of data mining in bioinformatics, information retrieval, web mining, image and text mining.

References

1. J. Han, M. Kamber, J. Pei, Data Mining: Concepts and Techniques 3E, Elsevier, 2011.
2. J. Han, M. Kamber: Data Mining: Concepts and Techniques, Morgan Kaufmann, 2000.
3. D. J. Hand, H. Mannila and P. Smyth: Principles of Data Mining, MIT Press, 2000.

4. M. Berry and G. Linoff: Mastering Data Mining, John Wiley & Sons, 2000.
5. A. K. Pujari: Data Mining Techniques, Sangam Books Ltd., 2001.

Subject Name: COMPILER DESIGN

Unit I: Introduction

Phases of compilation and overview Lexical Analysis (scanner): Regular language, finite automata, regular expression, from regular expression to finite automata, scanner generator (lex, flex).

Unit II: Syntax Analysis (Parser)

Context-free language and grammar, push-down automata, LL(1) grammar and top- down parsing, operator grammar, LR(0), SLR(1), LR(1), LALR(1) grammars and bottom-up parsing, ambiguity and LR parsing, LALR(1) parser generator (yacc, bison). Semantic Analysis: Attribute grammar, syntax directed definition, evaluation and flow of attribute in a syntax tree.

Unit III: Symbol Table

Its structure, symbol attributes and management. Run-time environment: Procedure activation, parameter passing, value return, memory allocation, and scope. Intermediate Code Generation: Translation of different language features, different types of intermediate forms.

Unit IV: Code Improvement (optimization)

Analysis: control-flow, data-flow dependence etc.; Code improvement local optimization, global optimization, loop optimization, peep-hole optimization etc. Architecture dependent code improvement: instruction scheduling (for pipeline), loop optimization (for cache memory) etc. Register allocation and target code generation Advanced topics: Type systems, data abstraction, compilation of object oriented features and non-imperative programming languages.

References

1. A. V. Aho, R. Sethi, J. D. Ullman, Compilers: Principles, Techniques and Tools, Addison-Wesley, 2007.
2. M. L. Scott, Programming Language Pragmatics, Elsevier, 2009.
3. A. W. Appel, Modern Compiler Implementation in C/Java, Cambridge University Press, 2004.
4. K. D. Cooper and Linda Torczon, Engineering a Compiler, Elsevier, 2011.
5. A. I. Holob, Compiler Design in C, Prentice-Hall, 1994.
6. S. S. Muchnik, Advanced Compiler Design and Implementation, Elsevier, 1997.
7. R. Allen, K. Kennedy, Optimizing Compilers for Modern Architectures, Elsevier, 2007.

Subject Name: ANALYSIS & DESIGN OF ALGORITHMS

Unit I: Introduction and basic concepts

Complexity measures, worst-case and average-case complexity functions, problem complexity, quick review of basic data structures and algorithm design principles.

Sorting and selection: Finding maximum and minimum, k largest elements in order; Sorting by selection, tournament and heap sort methods, lower bound for sorting, other sorting algorithms - radix sort, quick sort, merge sort; Selection of k-th largest element.

Unit II: Searching and set manipulation

Searching in static table – binary search, path lengths in binary trees and applications, optimality of binary search in worst cast and average-case, binary search trees, construction of optimal weighted binary search trees; Searching in dynamic table – randomly grown binary search trees, AVL and (a,b) trees.

Hashing: Basic ingredients, analysis of hashing with chaining and with open addressing.

Union-Find problem: Tree representation of a set, weighted union and path compression-analysis and applications.

Unit III: Graph problems

Graph searching – BFS, DFS, shortest first search, topological sort; connected and biconnected components; Minimum spanning trees – Kruskal's and Prim's algorithms – Johnson's implementation of Prim's algorithm using priority queue data structures.

Algebraic problems: Evaluation of polynomials with or without preprocessing. Winograd's and Strassen's matrix multiplication algorithms and applications to related problems, FFT, simple lower bound results.

Unit IV: String processing

String searching and Pattern matching, Knuth-Morris-Pratt algorithm and its analysis.

NP-completeness: Informal concepts of deterministic and nondeterministic algorithms, P and NP, NP-completeness, statement of Cook's theorem, some standard NP-complete problems, approximation algorithms.

References

1. T. H. Cormen, C. L. Leiserson, R. L. Rivest, and C. Stein, Introduction to Algorithms, MIT Press, 2003.
2. J. Kleinberg and E. Tardos, Algorithm Design, Addison-Wesley, 2006.
3. Harry R. Lewis and Larry Denenberg, Data Structures and Their Algorithms, Harper Collins, 1991.
4. A. Gibbons, Algorithmic Graph Theory, Cambridge University Press, 1985.
5. Michael T. Goodrich and Roberto Tamassia, Algorithm Design: Foundations, Analysis, and Internet Examples, John Wiley, 2006.
6. R. Sedgewick, Algorithms in C: Part 5, Addison Wesley, 2001.
7. M. H. Alsuwaiyel, Algorithm Design Techniques and Analysis, World Scientific, 1999.
8. Gilles Brassard and Paul Bratley, Algorithmics: theory and practice, Prentice-Hall, 1988.
9. Udi Manber, Introduction to Algorithms: A Creative Approach, Addison-Wesley, 1989.
10. Sara Baase and Allen Van Gelder, Computer Algorithms: Introduction to Design and Analysis, Addison- Wesley, 2000.

Subject Name: MINOR PROJECT WORK