RAJIV GANDHI PROUDYOGIKI VISHWAVIDYALAYA, BHOPAL

New Scheme Based On AICTE Flexible Curricula

CSE- Data Science/Data Science, VIII-Semester

CD-801 Compiler Design

UNIT I - Introduction of Compiler, Major data Structure in compiler, BOOT Strapping & Porting, Compiler structure: analysis-synthesis model of compilation, various phases of a compiler, Lexical analysis: Input buffering , Specification & Recognition of Tokens, LEX.

UNIT II - Syntax Analysis &Syntax Directed TranslationSyntax analysis: CFGs, Top down parsing, Brute force approach, recursive descent parsing, transformation on the grammars, predictive parsing, bottom up parsing, operator precedence parsing, LR parsers (SLR,LALR, LR),Parser generation.Syntax directed definitions: Construction of Syntax trees, Bottom up evaluation of S-attributed definition, L-attribute definition, Top down translation, Bottom Up evaluation of inherited attributes Recursive Evaluation, Analysis of Syntax directed definition.

UNIT III - Type Checking & Run Time EnvironmentType checking: type system, specification of simple type checker, equivalence of expression, types, type conversion, overloading of functions and operations, polymorphic functions. Run time Environment: storage organization, Storage allocation strategies, parameter passing, dynamic storage allocation , Symbol table

UNIT IV - Code GenerationIntermediate code generation: Declarations, Assignment statements, Boolean expressions, Case statements, Back patching, Procedure calls Code Generation: Issues in the design of code generator, Basic block and flow graphs, Register allocation and assignment, DAG representation of basic blocks, peephole optimization, generating code from DAG.

UNIT IV -Code OptimizationIntroduction to Code optimization: sources of optimization of basic blocks, loops in flow graphs, dead code elimination, loop optimization, Introduction to global data flow analysis, Code Improving transformations ,Data flow analysis of structure flow graph Symbolic debugging of optimized code.

Lab –

Assignments using Lex and Yaac

Books:

1. Compilers: Principles, Techniques and Tools, V. Aho, R. Sethi and J. Ullman. PearsonEducation

2. Lex &Yacc, Levine R. John, Tony Mason and Doug Brown, O'Reilly

Reference Books:

1. The Design and Evolution of C++, Bjarne Stroustrup.
2. Compiler Design, Raghavan, TMH Pub.
3. Compiler Construction: Principles and Practice, Louden, Cengage Learning
4. Compiler Design in C, A. C. Holub. Prentice-Hall Inc., 1993.
5. Writing compiler & Interpreters, Mak, Willey Pub.

RAJIV GANDHI PROUDYOGIKI VISHWAVIDYALAYA, BHOPAL

New Scheme Based On AICTE Flexible Curricula

CSE- Data Science/Data Science, VIII-Semester

802(A) Reinforcement Learning

Unit I - Introduction: State of the art applications in Atari, Alpha Go, relation to other problems in artificial intelligence. Markov Decision Processes (model based): Formulation, Value Iteration (VI), Policy Iteration (PI), Linear Programming (LP).

Unit II -Approximate Dynamic Programming (approximate model based): curse-of-dimensionality, representations, Approximate value iteration, approximate policy iteration, approximate linear program, approximation and convergence guarantees.

Unit III -Stochastic Approximation: Single and multi-timescale stochastic approximation, introduction to ordinary differential equation based convergence results. Value function learning (approximate model-free): Temporal difference (TD learning, TD(0), TD(lambda), Q-learning, State-Action-Reward-State Algorithm (SARSA) , TD with function approximation, on/off-policy learning, gradient temporal difference learning.

Unit IV - Actor-Critic: Policy gradient, Natural Actor-Critic, Deep RL.

Unit V -Exploration vs Exploitation: Upper Confidence Bound (UCB), Upper Confidence Reinforcement Learning (UCRL).

Text Books :

1. Richard S. Sutton and Andrew G. Barto, Introduction to Reinforcement Learning, 2nd Edition, MIT Press. 2017. ISBN-13 978-0262039246.
2. Dimitri Bertsekas and John G. Tsitsiklis, Neuro Dynamic Programming, Athena Scientific. 1996. ISBN-13: 978-1886529106.

References :

1. V. S. Borkar, Stochastic Approximation: A Dynamical Systems Viewpoint, Hindustan Book Agency, 2009. ISBN-13: 978-0521515924
2. Deep Learning. Ian Goodfellow and Yoshua Bengio and Aaron Courville. MIT Press. 2016.ISBN-13: 978-0262035613.

Course outcomes :

The student should be able to -

a) model a control task in the framework of MDPs.

b) Identify the model based from the model free methods.

c) Identify stability/convergence and approximation properties of RL algorithms.

d) Use deep learning methods to RL problems in practice.

RAJIV GANDHI PROUDYOGIKI VISHWAVIDYALAYA, BHOPAL

New Scheme Based On AICTE Flexible Curricula

CSE- Data Science/Data Science, VIII-Semester

802(B) Project Management

Course Outcome: Students will be able to

1. Learn activities involved in IT projects management.
2. Apply agile process to project management.
3. Plan application development using Scrum.
4. Develop abilities to use DevOps in projects.
5. Develop understanding of Containers use in projects.

UNIT I - Project Overview and Feasibility Studies- Identification, Market and Demand Analysis, Project Cost Estimate, Financial Appraisal.

UNIT II - Project Scheduling: Project Scheduling, Introduction to PERT and CPM, Critical Path Calculation, Precedence Relationship, Difference between PERT and CPM, Float Calculation and its importance, Cost reduction by Crashing of activity. Cost Control and Scheduling: Project Cost Control (PERT/Cost), Resource Scheduling & Resource Leveling

UNIT III -Project Management Features: Risk Analysis, Project Control, Project Audit and Project Termination Agile Project Management: Introduction, Agile Principles, Agile methodologies, Relationship between Agile Scrum, Lean, DevOps and IT Service Management (ITIL).

UNITIV- Scrum: Various terminologies used in Scrum (Sprint, product backlog, sprint backlog, sprint review, retro perspective), various roles (Roles in Scrum), Best practices of Scrum.

UNIT V -DevOps: Overview and its Components, Containerization Using Docker, Managing Source Code and Automating Builds, Automated Testing and Test Driven Development, Continuous Integration, Configuration Management, Continuous Deployment, Automated Monitoring. Other Agile Methodologies: Introduction to XP, FDD, DSDM, Crystal.

Text Book(s):

1. Mike Cohn, “Succeeding with Agile: Software Development Using Scrum”, Addison Wesley, 2009
2. Pearson, Robert C. Martin, Juli, James Shore, “The Art Of Agile Development”, O'Reilly, 2013
3. John Hunt, “Agile Software Construction”, 1st Edition, Springer,2005
4. Somerville, “Software Engineering”, 10th edition (Chapter 3, Chapters 22 to 26), Pearson, 2017
5. Deepak Gaikwad, Viral Thakkar, “DevOps Tools from Practitioner's Viewpoint”, Wiley, 2019
6. James Turnbill, “The Docker Book”, 2019

Reference Book(s):

1. Roman Pichler, “Agile Product Management with Scrum”.
2. Ken Schwaber, “Agile Project Management with Scrum” (Microsoft Professional)
3. Andrew Stellman, Jenifer Greene, “Head First Agile”, Oreilly, 2017
4. Peggy Gregory, Casper Lassenius, Xiaofeng Wang Philippe Kruchten (Eds.), “Agile Processes in Software Engineering and Extreme Programming”, 22nd International Conference on Agile Software Development, XP 2021 Virtual Event, June 14–18, 2021, Proceedings, Springer
5. Joseph Phillips, IT Project Management: On Track from Start to Finish, 3rd Edition, McGraw-Hill, 2010
6. Clinton Keith, “Agile Game Development”, Addison Wesley, 2010
7. Scott M Graffius, “Agile Scrum: Your Quick Start Guide with Step-by-Step Instructions”, CreateSpace, 2016

RAJIV GANDHI PROUDYOGIKI VISHWAVIDYALAYA, BHOPAL

New Scheme Based On AICTE Flexible Curricula

CSE- Data Science/Data Science, VIII-Semester

802(C) Computational Statistics

UNIT I -Computational Statistics- Probability concepts, Sampling Concepts, Generating Random Variables, Exploratory Data Analysis, Monte Carlo Methods for Inferential Statistics.

UNIT II -Data Partitioning, Probability Density Estimation, Statistical Pattern Recognition, Nonparametric Regression.

UNITIII -Data Mining- data mining algorithms-Instance and Features, Types of Features (data), Concept Learning and Concept Description, Output of data mining Knowledge Representation; Decision Trees- Classification and Regression trees constructing.

UNIT IV -Classification trees, Algorithm for Normal Attributes, Information Theory and Information. Entropy, Building tree, Highly-Branching Attributes, ID3 to c4.5, CHAID, CART, Regression Trees, Model Trees, Pruning.

UNIT V -Preprocessing and Post processing in data mining – Steps in Preprocessing, Discretization, Manual Approach, Binning, Entropy- based Discretization, Gaussian Approximation, K-tile method, Chi Merge, Feature extraction, selection and construction, Feature extraction, Algorithms, Feature selection, Feature construction, Missing Data, Post processing. Association Rule Mining- The Apriori Algorithm. Multiple Regression Analysis, Logistic Regression, k- Nearest Neighbor Classification, Constructing new attributes for algorithms of decision trees. Induction, Quick, Unbiased and Efficient Statistical tree.

Text Books / References

1. Wendy L. Martinez and Angel R, “Martinez Computational Statistics,” Chapman & Hall/CRC, 2002.
2. Ian H. Witten, “Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations”, Morgan Kaufmann, 2000.
3. Jiawei Han and Micheline Kamber, “Data Mining: Concepts and Techniques,” Morgan Kaufmann Publishers, 2001.
4. K. P. Soman, V. Ajay and Diwakar Shyam, “Insight into Data Mining: Theory and Practice”, Prentice Hall India, 2005.

RAJIV GANDHI PROUDYOGIKI VISHWAVIDYALAYA, BHOPAL

New Scheme Based On AICTE Flexible Curricula

CSE- Data Science/Data Science, VIII-Semester

802(D) Machine Learning for Data Science

UNIT I INTRODUCTION TO MACHINE LEARNING

Review of Linear Algebra for machine learning; Introduction and motivation for machine learning; Examples of machine learning applications, Vapnik-Chervonenkis (VC) dimension, Probably Approximately Correct (PAC) learning, Hypothesis spaces, Inductive bias, Generalization, Bias variance trade-off.

UNIT II SUPERVISED LEARNING

Linear Regression Models: Least squares, single & multiple variables, Bayesian linear regression, gradient descent, Linear Classification Models: Discriminant function – Perceptron algorithm, Probabilistic discriminative model - Logistic regression, Probabilistic generative model – Naive Bayes, Maximum margin classifier – Support vector machine, Decision Tree, Random Forests

UNIT III ENSEMBLE TECHNIQUES AND UNSUPERVISED LEARNING

Combining multiple learners: Model combination schemes, Voting, Ensemble Learning - bagging, boosting, stacking, Unsupervised learning: K-means, Instance Based Learning: KNN, Gaussian mixture models and Expectation maximization.

UNIT IV NEURAL NETWORKS

Multilayer perceptron, activation functions, network training – gradient descent optimization – stochastic gradient descent, error backpropagation, from shallow networks to deep networks –Unit saturation (aka the vanishing gradient problem) – ReLU, hyperparameter tuning, batch normalization, regularization, dropout.

UNIT V DESIGN AND ANALYSIS OF MACHINE LEARNING EXPERIMENTS

Guidelines for machine learning experiments, Cross Validation (CV) and resampling – K-fold CV, bootstrapping, measuring classifier performance, assessing a single classification algorithm and comparing two classification algorithms – t test, McNemar's test, K-fold CV paired t test

COURSE OUTCOMES:

At the end of this course, the students will be able to:
CO1: Explain the basic concepts of machine learning.
CO2 : Construct supervised learning models.
CO3 : Construct unsupervised learning algorithms.
CO4: Evaluate and compare different models

TEXTBOOKS:

1. Ethem Alpaydin, "Introduction to Machine Learning", MIT Press, Fourth Edition, 2020.
2. Stephen Marsland, "Machine Learning: An Algorithmic Perspective, "Second Edition", CRC Press,

RAJIV GANDHI PROUDYOGIKI VISHWAVIDYALAYA, BHOPAL

New Scheme Based On AICTE Flexible Curricula

CSE- Data Science / Data Science, VIII-Semester

803 (A) Block chain Technologies

UNIT I - Introduction: Overview of Block chain, Public Ledgers, Bit coin, Smart Contracts, Block in a Block chain, Transactions, Distributed Consensus, Public vs Private Block chain, Understanding Crypto currency to Block chain, Permissioned Model of Block chain, Overview of Security aspects of Block chain; Basic Crypto Primitives: Cryptographic Hash Function, Properties of a hash function, Hash pointer and Merkle tree, Digital Signature, Public Key Cryptography, A basic crypto currency.

UNIT II - Understanding Block chain with Crypto currency: Bit coin and Block chain: Creation of coins, Payments and double spending, Bit coin Scripts, Bit coin P2P Network, Transaction in Bit coin Network, Block Mining, Block propagation and block relay. Working with Consensus in Bit coin: Distributed consensus in open environments, Consensus in a Bitcoin network, Proof of Work (PoW) – basic introduction, Hash Cash PoW, Bit coin PoW, Attacks on PoW and the monopoly problem, Proof of Stake, Proof of Burn and Proof of Elapsed Time, The life of a Bitcoin Miner, Mining Difficulty, Mining Pool.

UNIT III - Understanding Block chain for Enterprises: Permissioned Block chain: Permissioned model and use cases, Design issues for Permissioned block chains, Execute contracts, State machine replication, Overview of Consensus models for permissioned block chain- Distributed consensus in closed environment, Paxos, RAFT Consensus, Byzantine general problem, Byzantine fault tolerant system, Lamport-Shostak-Pease BFT Algorithm, BFT over Asynchronous systems.

UNIT IV - Enterprise application of Block chain: Cross border payments, Know Your Customer (KYC), Food Security, Mortgage over Block chain, Block chain enabled Trade, We Trade – Trade Finance Network, Supply Chain Financing, and Identity on Block chain.

UNIT V - Block chain application development: Hyperledger Fabric- Architecture, Identities and Policies, Membership and Access Control, Channels, Transaction Validation, Writing smart contract using Hyperledger Fabric, Writing smart contract using Ethereum, Overview of Ripple and Corda.

References:

1. Melanie Swan, “Block Chain: Blueprint for a New Economy”, O’Reilly, 2015
2. Josh Thompsons, “Block Chain: The Block Chain for Beginners- Guide to Block chain Technology and Leveraging Block Chain Programming”
3. Daniel Drescher, “Block Chain Basics”, Apress; 1stedition, 2017
4. Anshul Kaushik, “Block Chain and Crypto Currencies”, Khanna Publishing House, Delhi.
5.Imran Bashir, “Mastering Block Chain: Distributed Ledger Technology, Decentralization and Smart Contracts Explained”, Packt Publishing
6. Ritesh Modi, “Solidity Programming Essentials: A Beginner’s Guide to Build Smart Contracts for Ethereum and Block Chain”, Packt Publishing
7. Salman Baset, Luc Desrosiers, Nitin Gaur, Petr Novotny, Anthony O’Dowd, Venkatraman Ramakrishna, “Hands-On Block Chain with Hyper ledger: Building Decentralized Applications with Hyperledger Fabric and Composer”, Import, 2018

RAJIV GANDHI PROUDYOGIKI VISHWAVIDYALAYA, BHOPAL

New Scheme Based On AICTE Flexible Curricula

CSE- Data Science/Data Science, VIII-Semester

803 (B) Time-Series Analysis

Unit I - INTRODUCTION OF TIMESERIES ANALYSIS: Introduction to TimeSeries and Forecasting, Different types of data, Internal structures of timeseries. Models for time series analysis, Autocorrelation and Partialautocorrelation. Examples of Time series Nature and uses of forecasting,
Forecasting Process, Data for forecasting, Resources for forecasting.

Unit II - STATISTICS BACKGROUND FOR FORECASTING: Graphical Displays,Time Series Plots, Plotting Smoothed Data, Numerical Description of TimeSeries Data, Use of Data Transformations and Adjustments, GeneralApproach to Time Series Modeling and Forecasting, Evaluating andMonitoring Forecasting Model Performance.

Unit III - TIME SERIES REGRESSION MODEL: Introduction Least SquaresEstimation in Linear Regression Models, Statistical Inference in LinearRegression, Prediction of New Observations, Model Adequacy Checking,Variable Selection Methods in Regression, Generalized and Weighted LeastSquares, Regression Models for General Time Series Data, ExponentialSmoothing, First order and Second order.

Unit IV - AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA)MODELS: Autoregressive Moving Average (ARMA) Models – Stationarityand Invertibility of ARMA Models - Checking for Stationarity usingVariogram- Detecting Nonstationarity - Autoregressive Integrated MovingAverage (ARIMA) Models - Forecasting using ARIMA - Seasonal Data -Seasonal ARIMA Models Forecasting using Seasonal ARIMA ModelsIntroduction - Finding the "BEST" Model -Example: Internet Users DataModel Selection Criteria - Impulse Response Function to Study theDifferences in Models Comparing Impulse Response Functions forCompeting Models .

Unit V - MULTIVARIATE TIME SERIES MODELS AND FORECASTING: Multivariate Time Series Models and Forecasting, Multivariate StationaryProcess, Vector ARIMA Models, Vector AR (VAR) Models, NeuralNetworks and Forecasting Spectral Analysis, Bayesian Methods inForecasting.

TEXTBOOKS:

1. **Introduction To Time Series Analysis AndForecasting**, 2nd Edition, Wiley Series In Probability And Statistics, By Douglas C. Montgomery, Cheryl L. Jen.
2. **Master Time Series Data Processing, Visualization, And Modeling Using Python** Dr. Avishek Pal Dr. Pks Prakash.

Time Series Analysis Lab

1. Time Series Data Cleaning
2. Loading and Handling Times series data
3. Preprocessing Techniques
4. How to Check Stationarity of a Time Series.
5. How to make a Time Series Stationary?
6. Estimating & Eliminating Trend.
   - Aggregation
   - Smoothing
   - Polynomial Fitting
7. Eliminating Trend and Seasonality
   - Differencing
   - Decomposition

a) Moving Average time analysis data.
b) Smoothing the Time analysis Data.
c) Check out the Time series Linear and non-linear trends.
d) Create a modelling.

8. Modelling time series
   - Moving average
   - Exponential smoothing
   - ARIMA
9. Seasonal autoregressive integrated moving average model (SARIMA)

10. Dependence Techniques
    - Multivariate Analysis of Variance and Covariance
    - Canonical Correlation Analysis
    - Structural Equation Modeling
11. Inter-Dependence Techniques
12. Factor Analysis
13. Cluster Analysis

RAJIV GANDHI PROUDYOGIKI VISHWAVIDYALAYA, BHOPAL

New Scheme Based On AICTE Flexible Curricula

CSE- Data Science/Data Science, VIII-Semester

CD 803 (C) Internet of Things

Unit I -IoT definition, Characteristics, IoT conceptual and architectural framework, Physical and logical design of IoT, IoT enablers, Modern day IoT applications, M2M communications, IoT vs M2M, IoT vs WoT, IoT reference architecture, IoT Network configurations, IoT LAN, IoT WAN, IoT Node, IoT Gateway, IoT Proxy, IPv4 vs IPV6

Unit II - Sensor, Basic components and challenges of a sensor node, Sensor features, Sensor resolution; Sensor classes: Analog, Digital, Scalar, Vector Sensors; Sensor Types, bias, drift, Hysteresis error, quantization error; Actuator; Actuator types: Hydraulic, Pneumatic, electrical, thermal/magnetic, mechanical actuators, soft actuators

Unit III - Basics of IoT Networking, IoT Components, Functional components of IoT, IoT service oriented architecture, IoT challenges, 6LowPAN, IEEE 802.15.4, ZigBee and its types, RFID Features, RFID working principle and applications, NFC (Near Field communication), Bluetooth, Wireless Sensor Networks and its Applications

Unit IV -MQTT, MQTT methods and components, MQTT communication, topics and applications, SMQTT, CoAP, CoAP message types, CoAP Request-Response model, XMPP, AMQP features and components, AMQP frame types

Unit V -IoT Platforms, Arduino, Raspberry Pi Board, Other IoT Platforms; Data Analytics for IoT, Cloud for IoT, Cloud storage models & communication APIs, IoT case studies

References:

1. Vijay Madisetti, Arshdeep Bahga, “Ïnternet of Things, A Hands on Approach”, University Press
2. Dr. SRN Reddy, Rachit Thukral and Manasi Mishra, “Introduction to Internet of Things: A practical Approach”,
ETI Labs
3. Pethuru Raj and Anupama C. Raman, “The Internet of Things: Enabling Technologies, Platforms, and Use
Cases”, CRC Press
4. Jeeva Jose, “Internet of Things”, Khanna Publishing House, Delhi
5. Adrian McEwen, “Designing the Internet of Things”, Wiley
6. Raj Kamal, “Internet of Things: Architecture and Design”, McGraw Hill
7. Cuno Pfister, “Getting Started with the Internet of Things”, O Reilly Media

Course Outcomes:

After the completion of this course, the students will be able to:

1. Understand Internet of Things and its hardware and software components
2. Interface I/O devices, sensors & communication modules
3. Analyze data from various sources in real-time and take necessary actions in an intelligent fashion
4. Remotely monitor data and control devices
5. Develop real life IoT based projects

RAJIV GANDHI PROUDYOGIKI VISHWAVIDYALAYA, BHOPAL

New Scheme Based On AICTE Flexible Curricula

CSE- Data Science/Data Science, VIII-Semester

803(D) Quantum Computing

COURSE OBJECTIVES:

Course Objectives: The objective of this course is to impart necessary knowledge to the learner so that he/she can develop and implement algorithm and write programs using these algorithm

Unit I Motivation for studying Quantum Computing , Major players in the industry (IBM, Microsoft, Rigetti, D-Wave etc.), Origin of Quantum Computing Overview of major concepts in Quantum Computing: Qubits and multi-qubits states, Braket notation, Bloch Sphere representation, Quantum Superposition, Quantum Entanglement

Unit II Math Foundation for Quantum Computing: Matrix Algebra: basis vectors and orthogonality, inner product and Hilbert spaces, matrices and tensors, unitary operators and projectors, Dirac notation, Eigen values and Eigen vectors

Unit III Building Blocks for Quantum Program: Architecture of a Quantum Computing platform, Details of q-bit system of information representation: Block Sphere, Multi-qubits States, Quantum superposition of qubits (valid and invalid superposition), Quantum Entanglement, Useful states from quantum algorithmic perceptive e.g. Bell State, Operation on qubits: Measuring and transforming using gates. Quantum Logic gates and Circuit: Pauli, Hadamard, phase shift, controlled gates, Ising, Deutsch, swap etc.

Unit IV Programming model for a Quantum Computing Program: Steps performed on classical computer, Steps performed on Quantum Computer, Moving data between bits and qubits. Basic techniques exploited by quantum algorithms, Amplitude amplification, Quantum Fourier Transform, Phase Kick-back, Quantum Phase estimation, Quantum Walks

Unit V Major Algorithms: Shor's Algorithm, Grover's Algorithm, Deutsch's Algorithm, Deutsch - Jozsa Algorithm OSS Toolkits for implementing Quantum program: IBM quantum experience, Microsoft Q, RigettiPyQuil (QPU/QVM)

References:

1. Michael A. Nielsen, "Quantum Computation and Quantum Information", Cambridge University Press.

2. David McMahon, "Quantum Computing Explained", Wiley

Course Outcomes:

After the completion of this course, the students will be able to:

1. Understand major concepts in Quantum Computing
2. Explain the working of a Quantum Computing program, its architecture and program model
3. Develop quantum logic gate circuits
4. Develop quantum algorithm
5. Program quantum algorithm on major toolkits.