

## Curriculum for Data Science (D.Voc.)

NSQF Level	Code	Educational Component	Credit	Marks
3 (1 <sup>st</sup> Year) Semester I	<b>Theory</b>			
	3.GE.01	Language – I	3	50
	3.GE.02	Applied Chemistry	3	50
	3.GE.03	Applied Physics	3	50
	3.GE.04	Applied Mathematics-I	3	50
	<b>Lab / Practical</b>			
	3.GP.01	Applied Chemistry Lab	1.5	50
	3.GP.02	Applied Physics Lab	1.5	50
	<b>On-Job-Training (OJT) / Qualification Packs</b>			
	Domestic Data Entry Operator		(Any one)	15
Domestic IT helpdesk Attendant				
CRM Domestic IT Attendant				
3 (1 <sup>st</sup> Year) Semester II	<b>Theory</b>			
	3.GV.01	General Foundation Course	3	50
	3.GV.02	Basic Electricity and Electronics	3	50
	3.GE.05	Language – II	3	50
	3.GE.06	Applied Mathematics -II	3	50
	<b>Lab / Practical</b>			
	3.VP.01	Basic Electricity and Electronics Lab	1.5	50
	3.GP.03	Language Lab	1.5	50
	<b>On-Job-Training (OJT) / Qualification Packs</b>			
	Any one of the QP's can be opted as offered in Semester I		(Any one)	15
4 (2 <sup>nd</sup> Year) Semester I	<b>Theory</b>			
	4.GV.01	IT Foundations	3	50
	4.GV.02	Programming for Problem Solving (using C / Python)	3	50
	4.GV.03	Introduction to Data Science	3	50
	4.GV.04	Mathematical Foundations	3	50
	<b>Lab / Practical</b>			
	4.VP.01	IT-Tools & Programming Lab (using C / Python)	1.5	50
	4.VP.02	Introduction to Programming Tools for Data Science Lab	1.5	50
	<b>On-Job-Training (OJT) / Qualification Packs</b>			
	IT Attendant		(Any one)	15
Junior Software Developer				
Cyber Awareness Instructor				
4 (2 <sup>nd</sup> Year) Semester II	<b>Theory</b>			
	4.GV.05	Data Structures & Algorithms	3	50
	4.GV.06	Database Management Systems	3	50
	4.GV.07	Computational Data Analytics	3	50
	4.GV.08	Computer Organization & Architecture	3	50
	<b>Lab / Practical</b>			
	4.VP.03	Data Structure Lab	1.5	50
	4.VP.04	Database Management System Lab	1.5	50
<b>On-Job-Training (OJT) / Qualification Packs</b>				

	One more QP to be opted from the QPs mentioned in the Level 4 first semester	(Any one)	15	200
<b>5 (3<sup>rd</sup> Year) Semester I</b>	<b>Theory</b>			
	5.GV.01	Operating System	3	50
	5.GV.02	Web Data Mining	3	50
	5.GV.03	Machine Learning - I	3	50
	5.GV.04	Parallel and Distributed Computing (inc. Cloud computing)	3	50
	<b>Lab / Practical</b>			
	5.VP.01	Web Data Mining Lab	1.5	50
	5.VP.02	Advanced Computer Network & Security Lab	1.5	50
	<b>On-Job-Training (OJT) / Qualification Packs</b>			
	Systems Analyst		(Any one)	15
Cyber Security Auditor				
Cyber Security Trainer				
<b>5 (3<sup>rd</sup> Year) Semester II</b>	<b>Theory</b>			
	5.GV.05	Machine Learning- II	3	50
	5.GV.06	Case Studies of Data Science Application	3	50
	5.GV.07	Penetration Testing	3	50
	5.GV.08	Analysing, Visualizing and Applying data science with python	3	50
	<b>Lab / Practical</b>			
	5.VP.03	Project Work & Seminar-Presentation	3	100
	<b>On-Job-Training (OJT) / Qualification Packs</b>			
One more QP to be opted from the QPs mentioned in the Level 5 first semester		(Any one)	15	200

## Detailed Curriculum

### (Level 3: 1<sup>st</sup> Year, Semester – I)

#### **(3.GE.01): Language – I**

**Reading comprehension (prescribed texts) and functional grammar:** A variety of genres – short stories, expository pieces, biographies, poems, plays, newspaper and magazine excerpts have been included. Teaching of grammar has been integrated with the reading texts. The emphasis is on functional grammar.

**Non prescribed:** In this section learners will be exposed to newspaper, articles, tables, diagrams, advertisements etc. which they have to read carefully and interpret. In the examination similar pieces will be used.

**Grammar and usage. Functional writing and study skills:** Paragraph writing, Letter writing, Note making, Ending (punctuation, spelling, appropriate vocabulary, structures).

**Reference Books:**

1. Effective Communication Skills, Kulbhushan Kumar, Khanna Publishing House
2. Business Communications, Varinder Bhatia, Khanna Publishing House

#### **(3.GE.02): Applied Chemistry**

Structure of Atom, Periodic Properties of Elements, Chemical Bonds, Fuel and their Classification, Water, Corrosion, Plastic and Polymers.

**Reference Books:**

1. Chemistry, Satyaprakash, Khanna Publishing House
2. Engineering Chemistry, Saiful Islam, Khanna Publishing House

#### **(3.GE.03): Applied Physics**

Units & Dimensions, Surface Tension and Viscosity, Vibrations, Heat, Ultrasonics, Optics.

**Reference Books:**

1. Engineering Physics, Malik and Singh, Tata Mc Graw Hill
2. Engineering Physics, Naidu, Pearson

3. Modern Physics for Engineers, S.P. Taneja, R. Chand

### **(3.GE.04): Applied Mathematics-I**

Sets, Relations and Functions, Sequences and Series, Algebra-I, Co-ordinate Geometry, Statistics and Probability.

#### **Reference Books:**

1. Applied Mathematics-I, J.K. Tyagi, Khanna Publishing House
2. Engineering Mathematics, Reena Garg, Khanna Publishing House

### **(3.GP.01): Applied Chemistry Lab**

1. Proximate analysis of solid fuel.
2. Experiments based on Bomb Calorimeter.
3. Determination of turbidity in a given sample.
4. To determine the flash and fire point of a given lubricating oil.
5. To determine the viscosity of a given lubricating oil by Redwood viscometer.
6. To determine cloud and pour point of a given oil.

### **(3.GP.02): Applied Physics Lab**

1. To determine the surface tension of a liquid by rise in capillary.
2. To determine the viscosity of a given liquid.
3. To determine the frequency of tuning fork using a sonometer.
4. To determine the frequency of AC main using sonometer.
5. Time period of a cantilever.

## **(Level 3: 1<sup>st</sup> Year, Semester – II)**

### **(3.GV.01): General Foundation Course**

Business Management and Entrepreneurship, Computational Skills, Environmental Education, Rural Development.

#### **Reference Books:**

1. Environmental Studies, M.P. Poonia & S.C. Sharma, Khanna Publishing House
2. A Textbook of Environmental Sciences, Rimpi Mehani Ne' Chopra, Khanna Publishing House

### **(3.GV.02): Basic Electricity and Electronics**

Current Electricity, D.C. Circuits, Electric Cells, Lighting Effects of Current, Capacitors, Electromagnetic Effects, A.C. Circuits.

Overview of Atom, Sub-Atomic Particles and CRO, Voltage and Current, Basics of Semiconductor, Bipolar Junction Transistor, Transistor Amplifier and Applications.

#### **Reference Books:**

1. Basic Electrical Engineering, Ritu Sahdev, Khanna Publishing House
2. Basic Electrical Engineering, Pradeep Kumar, Khanna Publishing
3. Basic Electronics, S. Biswas, Khanna Publishing House
4. All in One Electronics Simplifies, A.K. Maini, Khanna Publishing House

### **(3.GE.05): Language – II**

Listening and speaking skills, English for specific purposes (opt any one):

**English for Science:** 1. Health and hygiene, 2. Conservation of (nearly extinct) animals, 3. Plant life, 4. Bio gas / solar energy.

**English for Receptionist:** 1. Receiving messages, making request etc., 2. Supplying information, 3. Giving advice and making suggestions, 4. Dealing with complaints, 5. Making entries in an appointment book, register etc.

**English for Office Use:** 1. Using the telephone taking and passing messages, 2. Receiving messages, 3. Marking noting on files and circular, 4. Writing office notes, memos, notices, agendas for meetings, 5. Telegrams and fax messages, 6. Writing business letters, application enquires, complaints, 7. Filling in forms, cheques, pay in slips etc.

#### **Reference Books:**

1. Effective Communication Skills, Kulbhushan Kumar, Khanna Publishing House
2. Business Communications, Varinder Bhatia, Khanna Publishing House

### **(3.GE.06): Applied Mathematics – II**

Algebra-II, Relations and Functions, Calculus, Vectors and Three Dimensional Geometry, Linear Programming and Mathematical Reasoning.

#### **Reference Books:**

1. Applied Mathematics-II, J.K. Tyagi, Khanna Publishing House
2. Elements of Mathematical Analysis, R. Agor, Khanna Publishing House

### **(3.VP.01): Basic Electricity and Electronics Lab**

1. Verify that resistance of conductor is directly proportional to resistivity and length and inversely proportional to cross-sectional area of the conductor.
  2. Verification of Ohm's Law.
  3. Verification of temperature co-efficient of resistance:
    - (i) Positive for Tungsten and Nichrome and
    - (ii) Negative for carbon.
  4. Study of series resistive circuits.
  5. Study of parallel resistive circuits.
  6. Study of series and parallel connection of cells in circuits.
  7. Preparation of Electrolyte for lead acid battery and its charging and measurement of Specific gravity with the help of hydrometer.
  8. To find heat efficiency of an electric kettle.
  9. Charging and Discharging of a capacitor.
  10. Verification of magnetic field of a Solenoid with:
    - (i) Iron core and
    - (ii) Air core.
  11. Verification of Faraday's Laws of electromagnetic induction.
  12. Verification of Torque development in a current carrying coil in magnetic field.
  13. Study of R.L. series circuit and measurement of power and power factor.
  14. Study of R.C. series circuit and measurement of power and power factor.
  15. Study of R.L.C. series circuit and measurement of power and power factor.
  16. Study of R.L.C. series circuit for calculation of inductive reactance, capacitive reactance, impedance and Q- Factor.
- 
1. Study of current and voltage measurement using Ammeter and Voltmeter.
  2. Study of current and voltage measurement using Galvanometer.
  3. Study of current, voltage and resistance measurement using of Multi-meter
  4. Study of Power and Energy measurement using Wattmeter and Energy meter.
  5. Study of working principle of Signal Generator and measurement of amplitude, time period and frequency of signal using Oscilloscope.
  6. Study of V-I Characteristic of Diode.
  7. Study of V-I Characteristic of Zener Diode. And use of Zener Diode as voltage regulator.
  8. Study of Half wave rectifier with and without filter circuit.
  9. Study of Full wave rectifier with and without filter circuit.
  10. Study CE configuration for NPN and PNP transistors and measurement of voltage and current gain.
  11. Study CB configuration for NPN and PNP transistors and measurement of voltage and current gain.
  12. Study CC configuration for NPN and PNP transistors and measurement of voltage and current gain.
  13. Study of working of single layer PCB manufacturing
  14. Study of working of double layer PCB manufacturing.
  15. Design of 7 segment display using LED and bread board

### **(3.GP.03): Language Lab**

Technical Report Writing, Language Laboratory Practice, Conversation Practice Sessions, Group Discussion Sessions, Presentation, Competitive Examination.

## **(Level 4: 2<sup>nd</sup> Year, Semester – I)**

### **(4.GV.01): IT Foundations**

Computer Organization & OS: User perspective, Networking and Internet, Office automation tools, Multi Media Design: (Open Source Design Tools), Troubleshooting: Hardware, Software and Networking, Work Integrated Learning IT.

#### **Reference Books:**

1. IT Tools, R.K. Jain, Khanna Publishing House
2. Information Security & Cyber Laws, Sarika Gupta, Khanna Publishing House
3. Mastering PC Hardware & Networking, Ajit Mittal, Khanna Publishing House

### **(4.GV.02): Programming for Problem Solving (using C / Python)**

Introduction, Conditional Statements, Control Statements, String Manipulation, Lists, Tuple, Dictionaries, Functions, Modules, Exception Handling.

**Reference Books:**

1. E. Balaguruswamy, Programming in ANSI C, Tata McGraw-Hill
2. Brian W. Kernighan and Dennis M. Ritchie, The C Programming Language, Prentice Hall of India
3. Byron Gottfried, Schaum's Outline of Programming with C, McGraw-Hill
4. Python Crash Course, 2Nd Edition: A Hands-On, Project-Based Introduction to Programming, Eric Matthes
5. Python Programming: Using Problem Solving Approach, Reema Thareja
6. Core Python Programming, R. Nageswara Rao
7. Introduction to Computing and Problem Solving with Python, J. Jose, Khanna Publications

**(4.GV.03): Introduction to Data Science**

Toolkits using Python: Matplotlib, NumPy, Scikit-learn, NLTK, Visualizing Data: Bar Charts, Line Charts, Scatterplots, Working with data: Reading Files, Scraping the Web, Using APIs (Example: Using the Twitter APIs), Cleaning and Munging, Manipulating Data, Rescaling, Dimensionality Reduction

**Reference Books:**

1. Jain V.K., "Data Sciences", Khanna Publishing House, Delhi.
2. Aurélien Géron, "Hands-On Machine Learning with Scikit-Learn and Tensor Flow: Concepts, Tools, and Techniques to Build Intelligent Systems", 1st Edition, O'Reilly Media

**(4.GV.04): Mathematical Foundations**

1. Linear Algebra: Vectors, Matrices,
2. Statistics: Describing a Single Set of Data, Correlation, Simpson's Paradox, Correlation and Causation
3. Probability: Dependence and Independence, Conditional Probability, Bayes's Theorem, Random Variables, Continuous Distributions, The Normal Distribution, The Central Limit Theorem
4. Hypothesis and Inference: Statistical Hypothesis Testing, Confidence Intervals, Phacking, Bayesian Inference

**Reference Books:**

1. Advanced Level Pure Mathematics, S. L. Green, UTP
2. A Course of Pure Mathematics (Centenary Edition), G. H. Hardy, Cambridge University Press

**(4.VP.01): IT-Tools & Programming Lab (using C / Python)**

Spreadsheets, Word, Presentation, Multimedia Design, Troubleshooting, Study of computer components, Booting of Computer and its shutdown, Practicing some fundamental DOS Commands, Simple Programs in BASIC to compute Mean, Variance, Correlation and Regression, Creating database in MS-Access, structuring with different types of fields and use of query facility for accessing the information, Project / Practical File, Viva Voce.

Interactive interpreter and Python Script, Indentation Error, decision making and looping, one dimensional and two dimensional arrays, explore string functions, mean, median, mode, find all duplicates in the list, find all unique elements of a list, compute gcd, lcm of two numbers, use of Lists, Dictionaries, implement Turtle, Linear and Binary Search, Exceptions in Python.

**(4.VP.02): Introduction to Programming Tools for Data Science Lab**

1. Write a programme in Python to predict the class of the flower based on available attributes.
2. Write a programme in Python to predict if a loan will get approved or not.
3. Write a programme in Python to predict the traffic on a new mode of transport.
4. Write a programme in Python to predict the class of user.
5. Write a programme in Python to identify the tweets which are hate tweets and which are not.
6. Write a programme in Python to predict the age of the actors.
7. Mini project to predict the time taken to solve a problem given the current status of the user. & Viva Voce.

**(Level 4: 2<sup>nd</sup> Year, Semester – II)**

**(4.GV.05): Data Structures & Algorithms**

Program structures, Variables, Data Types, Declarations, Operators (Arithmetic, Relational, Logical), increment and decrement operators, Assignment operators and expressions, Arithmetic expressions, Functions, external variables, scope rules, header files, Pointers and addresses, pointers and function arguments, pointer and arrays, Data Structures: Arrays, Linked list, Stacks and queues, Trees, Graphs, Hashing, Searching & Sorting.

**Reference Books:**

1. Fundamentals of Data Structures, Sartaj Sahni, University Press
2. Data Structures through C, Yashwant Kanetkar, BPB Publications

3. Data Structures Through C In Depth, S.K.Srivastava/Deepali Srivastava
4. Data Structures using C & C++, Rajesh K. Shukla
5. Introduction to Algorithms, Thomas H. Cormen

#### **(4.GV.06): Database Management Systems**

Database system architecture, Data models, Relational query languages, Relational database design, Query processing and optimization, Storage strategies, Transaction processing, Database Security, Advanced topics.

##### **Reference Books:**

1. “Database System Concepts”, 6th Edition by Abraham Silberschatz, Henry F. Korth, S. Sudarshan, McGraw-Hill.
2. “Principles of Database and Knowledge – Base Systems”, Vol 1 by J. D. Ullman, Computer Science Press.
3. Database Management Systems, R.P. Mahapatra, Khanna Publishing House, New Delhi.
4. “Fundamentals of Database Systems”, 5th Edition by R. Elmasri and S. Navathe.
5. Pearson Education “Foundations of Databases”, Reprint by Serge Abiteboul, Richard Hull, Victor Vianu, Addison-Wesley

#### **(4.GV.07): Computational Data Analytics**

Introduction to R Computing language. Best practices in executing Reproducible Research in data science, Sampling and Simulation. Descriptive statistics, and the creation of good observational sampling designs, Data visualization, Data import and visualization, Introduction to various plots, Frequentist Hypothesis Testing, Z-Tests, Power Analysis, Linear regression, diagnostics, visualization, Likelihoodist Inference, Fitting a line with Likelihood, Model Selection with one predictor, Bayesian Inference, Fitting a line with Bayesian techniques, Multiple Regression and Interaction Effects, Information Theoretic Approaches

##### **Reference Books:**

1. Practical Data Science with R, Nina Zumel, John Wiley & Sons.
2. N. C. Das, Experimental Designs in Data Science with Least Resources, Shroff Publisher Publisher.
3. Hadley Wickham, Garret Golemund, R for Data Science, Shroff Publisher/O’Reilly Publisher Publisher
4. Benjamin M. Bolker. Ecological Models and Data in R. Princeton University Press, 2008.

#### **(4.GV.08): Computer Organization & Architecture**

Basic organization of the stored program computer and operation sequence for execution of a program, Role of operating systems and compiler/assembler, Fetch, decode and execute cycle, Concept of operator, operand, registers and storage, Instruction format. Instruction sets and addressing, modes, Commonly used number systems. Fixed and floating point representation of numbers, Overflow and underflow. Design of adders – ripple carry and carry look ahead principles, Design of ALU, Memory unit design with special emphasis on implementation of CPU-memory interfacing, Memory organization, static and dynamic memory, memory hierarchy, associative memory, Cache memory, Virtual memory, Design of control unit - hardwired and microprogrammed control, Introduction to instruction pipelining, Introduction to RISC architectures. RISC vs CISC architectures, I/O operations - Concept of handshaking, interrupt and DMA, Pipelining, Exception handling, Pipeline optimization techniques, Hierarchical memory technology, Instruction-level parallelism, Multiprocessor architecture, Non von Neumann architectures.

##### **Reference Books:**

1. Mano, M.M., “Computer System Architecture”, PHI.
2. Hayes J. P., “Computer Architecture & Organisation”, McGraw Hill.
3. Hamacher, “Computer Organisation”, McGraw Hill.
4. Chaudhuri P. Pal, “Computer Organisation & Design”, PHI.
5. P N Basu- “Computer Organization & Architecture”, Vikas Pub.
6. Rajaraman – “Computer Organization & Architecture”, PHI.
7. B.Ram – “Computer Organization & Architecture”, Newage Publications.
8. J. L. Hennessy and D. A. Patterson, “Computer Architecture A Quantitative Approach”, Morgan Kauffman.

#### **(4.VP.03): Data Structure Lab**

Implement stack, Write functions like push, pop, Initialize, Empty or Full, Implement concept of queues, Implement queues in a circular array, Implement queues as a circular linked list, Implementing doubly linked list, Binary search tree to sort an array.

#### **(4.VP.04): Database Management System Lab**

Programs / Practical Questions, SQL Queries, Creating Database, Table and Record Handling, Retrieving Data from a Database, Clause, Database Management, Cursors in Oracle PL / SQL, Writing Oracle PL / SQL Stored Procedures, Programs, Operating Web Based Application, Project / Practical File, Viva Voce.

## (Level 5: 3<sup>rd</sup> Year, Semester – I)

### **(5.GV.01): Operating System**

System Software: Operating System, Compiler, Interpreter and Assembler, Operating System: Need for Operating System, Functions of Operating System (Processor Management, Memory Management, File Management and Device Management), Types of Operating System-Interactive (GUI based), Time Sharing, Real Time and Distributed, commonly used Operating System: UNIX, LINUX, Windows, Solaris, BOSS (Bharat Operating System Solutions); Mobile OS – Android, Symbian, IOS. Utility Software: Anti-Virus, File Management tools, Compression tools and Disk Management tools (Disk Cleanup, Disk Defragmenter, Backup). Processes, Thread, Process Scheduling, Inter-process Communication, Deadlocks, Memory Management, Virtual Memory, I/O Hardware, File Management, Disk Management.

#### **Reference Books:**

1. Operating Systems, Ekta Walia, Khanna Publishing House.
2. Operating System Concepts Essentials, Avi Silberschatz, Peter Galvin, Greg Gagne, Wiley Asia Student Edition.
3. Operating Systems: Internals and Design Principles, William Stallings, Prentice Hall of India.
4. Operating Systems: A Modern Perspective, Gary J. Nutt, Addison-Wesley.

### **(5.GV.02): Web Data Mining**

Introduction to internet and WWW, Data Mining Foundations, Association Rules and Sequential Patterns, Basic Concepts of Association Rules, Apriori Algorithm, Frequent Itemset Generation, Association Rule Generation, Data Formats for Association Rule Mining, Mining with multiple minimum supports, Extended Model, Mining Algorithm, Rule Generation, Mining Class Association Rules, Basic Concepts of Sequential Patterns, Mining Sequential Patterns on GSP, Mining Sequential Patterns on Prefix Span, Generating Rules from Sequential Patterns, Concepts of Information Retrieval, IR Methods, Boolean Model, Vector Space Model and Statistical Language Model, Relevance Feedback, Evaluation Measures, Text and Web Page Pre-processing, Stopword Removal, Stemming, Web Page Preprocessing, Duplicate Detection, Inverted Index and Its Compression, Inverted Index, Search using Inverted Index, Index Construction, Index Compression, Latent Semantic Indexing, Singular Value Decomposition, Query and Retrieval, Web Search, Meta Search, Web Spamming.

#### **Reference Books:**

1. Mining the Web: Discovering Knowledge from Hypertext Data, Soumen Chakrabarti, Morgan Kaufmann Publishers.
2. Bing Liu, Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data, Springer Publications, 2011.
3. Jiawei Han, Micheline Kamber, Data Mining: Concepts and Techniques, Second Edition, Elsevier Publications 2010.
4. Anthony Scime, Web Mining: Applications and Techniques, 2005.
5. Kowalski, Gerald, Mark T Maybury: Information Retrieval Systems: Theory and Implementation, Kluwer Academic Press, 1997.
6. Mathew Russell, Mining the Social Web 2nd Edition, Shroff Publisher/O'Reilly Publisher Publication.
7. Data Mining and Data Warehousing Principles and Practical Techniques, Parteek Bhatia, Cambridge University Press

### **(5.GV.03): Machine Learning -I**

Overview of Machine learning concepts – Over fitting and train/test splits, Types of Machine learning – Supervised, Unsupervised, Reinforced learning, Introduction to Bayes Theorem, Linear Regression- model assumptions, regularization (lasso, ridge, elastic net)

#### **Reference Books:**

1. Jeeva Jose, “Machine Learning”, Khanna Publishing House, Delhi.
2. Chopra Rajiv, “Machine Learning”, Khanna Publishing House, Delhi.

### **(5.GV.04): Parallel and Distributed Computing (inc. Cloud computing)**

Introduction, Parallel Programming Platforms, Principles of Parallel Algorithm Design, CUDA programming model, Analytical Modeling of Parallel Programs, Dense Matrix Algorithms, Graph Algorithms, Cloud Computing.

#### **Reference Books:**

1. A Grama, A Gupta, G Karypis, V Kumar. Introduction to Parallel Computing, Addison Wesley.
2. C Lin, L Snyder. Principles of Parallel Programming. USA: Addison-Wesley Publishing Company.

3. J Jeffers, J Reinders. Intel Xeon Phi Coprocessor High-Performance Programming. Morgan Kaufmann Publishing and Elsevier.
4. T Mattson, B Sanders, B Massingill. Patterns for Parallel Programming. Addison-Wesley Professional.
5. Cloud Computing Bible by Barrie Sosinsky, Wiley India Pvt. Ltd, 2013.
6. Mastering Cloud Computing by Rajkumar Buyya, Christian Vecchiola, S. Thamarai Selvi, McGraw Hill Education (India) Private Limited.
7. Cloud computing: A practical approach, Anthony T. Velte, Tata Mcgraw-Hill.
8. Cloud Computing, Miller, Pearson.
9. Cloud Computing, Kumar Saurabh, Wiley India.

### **(5.VP.01): Web Data Mining Lab**

1. Demonstrate knowledge of Data Science and Machine Learning.
2. Apply Data Science process to a real life scenario.
3. Explore New York City - 311 Complaints and Housing datasets.
4. Analyze and Visualize data using Python.
5. Perform feature engineering exercise using Python.
6. Build and validate predictive machine learning model using Python.
7. Create and share Actionable Insights to real life data problems.

### **(5.VP.02): Advanced Computer Network & Security Lab**

1. Configuration and logging to a CISCO Router and introduction to the basic user Interfaces. Introduction to the basic router configuration and basic commands.
2. Configuration of IP addressing for a given scenario for a given set of topologies.
3. Configure a DHCP Server to serve contiguous IP addresses to a pool of four IP devices with a default gateway and a default DNS address. Integrate the DHCP server with a BOOTP demon to automatically serve Windows and Linux OS Binaries based on client MAC address.
4. Configure, implement and debug the following: Use open source tools for debugging and diagnostics.
  - a. ARP/RARP protocols
  - b. RIP routing protocols
  - c. BGP routing
  - d. OSPF routing protocols
  - e. Static routes (check using netstat)
5. Configure DNS: Make a caching DNS client, and a DNS Proxy; implement reverse DNS and forward DNS, using TCP dump/Wireshark characterise traffic when the DNS server is up and when it is down.
6. Configure FTP Server on a Linux/Windows machine using a FTP client/SFTP client characterise file transfer rate for a cluster of small files 100k each and a video file of 700mb. Use a TFTP client and repeat the experiment.
7. Configure a mail server for IMAP/POP protocols and write a simple SMTP client in C/C++/Java client to send and receive mails.
8. Implement Open NMS+ SNMPD for checking Device status of devices in community MIB of a Linux PC. Using yellow pages and NIS/NFS protocols implement Network Attached Storage Controller (NAS). Extend this to serve a windows client using SMB. Characterise the NAS traffic using wireshark.
9. Configure wireless network with AP, PC, Mobile Device.

## **(Level 5: 3<sup>rd</sup> Year, Semester – II)**

### **(5.GV.05): Machine Learning -II**

Classification and Regression algorithms- Naïve Bayes, K-Nearest Neighbors, logistic regression, support vector machines (SVM), decision trees, and random forest, Classification Errors, Analysis of Time Series- Linear Systems Analysis, Nonlinear Dynamics, Rule Induction, Neural Networks Learning And Generalization, Overview of Deep Learning

#### **Reference Books:**

1. Jeeva Jose, "Machine Learning", Khanna Publishing House, Delhi.
2. Chopra Rajiv, "Machine Learning", Khanna Publishing House, Delhi.
3. Ian Goodfellow, Yoshua Bengio and Aaron Courville, "Deep Learning", MIT Press <http://www.deeplearningbook.org>
4. Jiawei Han and Jian Pei, "Data Mining Concepts and Techniques", Third Edition, Morgan Kaufmann Publishers

### **(5.GV.06): Case Studies of Data Science Application**

Weather forecasting, Stock market prediction, Object recognition, Real Time Sentiment Analysis



**Reference Books:**

1. Jain V.K., “Big Data and Hadoop”, Khanna Publishing House, Delhi
2. Jain V.K., “Data Sciences”, Khanna Publishing House, Delhi
3. Ian Goodfellow, Yoshua Bengio and Aaron Courville, "Deep Learning", MIT Press

**(5.GV.07): Penetration Testing**

Comprehensive Pen Test Planning, Scoping, and Recon, In-Depth Scanning, Exploitation, Password Attacks and Merciless Pivoting, Domain Domination and Azure Annihilation, Penetration Test and Capture-the-Flag Workshop.

**Reference Books:**

1. Michael Sikorski, Andrew Honig “Practical Malware Analysis: The Hands-On Guide to Dissecting Malicious Software” publisher Williampollo.
2. Rafeeq Rehman: “Intrusion Detection with SNORT, Apache, MySQL, PHP and ACID,” Prentice Hall.
3. Christopher Kruegel, Fredrik Valeur, Giovanni Vigna: “Intrusion Detection and Correlation Challenges and Solutions”.
4. Carl Endorf, Eugene Schultz and Jim Mellander “Intrusion Detection & Prevention”, Tata McGraw-Hill.
5. Stephen Northcutt, Judy Novak: “Network Intrusion Detection”, New Riders Publishing.
6. T. Fahringer, R. Prodan, “A Text book on Grid Application Development and Computing Environment”, Khanna Publishers.

**(5.GV.08): Analysing, Visualizing and Applying data science with python**

Data Analysis libraries: will learn to use Pandas DataFrames, Numpy multi-dimensional arrays, and SciPy libraries to work with a various dataset, Pandas, an open-source library, and we will use it to load, manipulate, analyze, and visualize various datasets. Scikit-learn, and we will use some of its machine learning algorithms to build smart models and make predictions, various parameters that can be used to compare various parameters,

Descriptive Statistics, Basic of Grouping, ANOVA, Correlation, Polynomial Regression and Pipelines, R-squared and MSE for In-Sample Evaluation, Prediction and Decision Making.

**Reference Books:**

1. Data Visualization with Python and JavaScript, Kyran Dale, Shroff Publisher/O’Reilly Publisher Publication.
2. Data Science Using Python and R by Chantal D. Larose and Daniel T. Larose, Wiley Publication.
3. Python for Data Science and Visualization -Beginners to Pro, Udemy.

**(5.VP.03): Project Work & Seminar-Presentation**

On the basis of learning in the Diploma of Vocational, a project to be taken up by the student strengthening his / her vocational skills.

\*\*\*\*\*