# FACULTY OF ENGINEERING & TECHNOLOGY

# SYLLABUS

# FOR

M.TECH. COMPUTER SCIENCE & ENGINEERING (Credit Based Evaluation and Grading System) (SEMESTER: I to IV)

Session: 2019-20



# GURU NANAK DEV UNIVERSITY AMRITSAR

Note: (i) Copy rights are reserved. Nobody is allowed to print it in any form. Defaulters will be prosecuted.

> (ii) Subject to change in the syllabi at any time. Please visit the University website time to time.

# **SCHEME:**

Sr.No.	Subject Code	Subject		Cred	its
		Semester-I	L	T	Р
1	CSL553	Programming Concepts	3	0	1
2	CSL554	Distributed Operating System	4	0	0
3	CSL555	Network Security	4	0	0
4	CSL548	Data Mining	4	0	0
		Sub Total:	15	0	1
		Grand Total:		16	

S. No.	Subject Code	Subject		Cred	its	
	Semester-II					
1	CSL560	Computer System Architecture &	4	0	0	
		Organization				
2	CSL569	Software Engineering & Testing		0	1	
3	CSL567	Wireless Networks	4	0	0	
4		Elective-I	4	0	0	
6.	CSP569	TERM PAPER	0	0	1	
		Sub Total:	15	0	2	
		Grand Total:		17		
		List of Electives – I				
1	CSL564	Advanced Software Engineering	4	0	0	
2	CSL565	Network Programming	4	0	0	
3	CSL566	Advanced Cloud Computing	4	0	0	

Note:

PSL-053 ID Course Human Rights & Constitutional Duties (Compulsory Paper) Students can opt. in any semester except Semester 1<sup>st</sup>. This ID Paper is one of the total ID Papers of this course.

		Semester-III			
S. No.	Code	Subject	L	Т	Р
1.	CSL656	Introduction to Data Science	3	0	1
2.		Elective-II	4	0	0
3.		Inter-disciplinary subject-I	4	0	0
4.	CSD651	Dissertation (Part–I)	0	0	4
		Sub Total	11	0	5
		Grand Total		16	
		List of Electives-II:			
1	CSL650	Parallel Computing	4	0	0
2	CSL652	Artificial Neural Networks	4	0	0
3	CSL653	Digital Image Processing	4	0	0
5	CSL655	Software Project Management	4	0	0

CSB2: M.TECH. (COMPUTER SCIENCE & ENGINEERING) SEMESTER SYSTEM (Credit Based Evaluation and Grading System)

		Semester-IV:			
1.		Inter-disciplinary subject-II	4	0	0
2.	CSD660	Dissertation (Part-II)	0	0	12
		Grand Total:		16	1

# **CSL-553: PROGRAMMING CONCEPTS**

Internal Marks:	20	CREDITS
External Marks:	80	LTP
		3 0 1

#### **Instructions for the Paper Setters:**

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

#### **SECTION-A**

Constants, variables, vectors, matrices; arithmetic and vector/matrix operations, operator precedence. Input and display of data; writing programs in MATLAB environment.

Selection structures: IF and SWITCH.

Loops: FOR and WHILE structures. Multiple (nested) loops. Flags. BREAK and CONTINUE statements.

#### **SECTION-B**

Arrays, cell arrays, structures, string manipulation, Debugging programs.

User-defined functions, Scripts, Passing arguments, returning values; NARGIN, NARGOUT, GLOBAL variables, Formatted output.

#### **SECTION-C**

Files: Saving/loading data from files. Opening/closing files, handling file errors. Selecting files through dialog boxes. Repeating operations until end-of-file.

# **SECTION-D**

Applications: Linear Algebra, Curve Fitting and Exploration, Data Analysis and Statistics Plotting: Preparing data to be plotted; formatting plots; 2D chart types. Preparing data for 3D plotting; 3D chart types. Working with chart handles. Simple curve fitting.

- 1. Essential MATLAB for Engineers and Scientists, Fourth Edition; Brian H. Hahn, Daniel T. Valentine.
- 2. Getting started with MATLAB: A quick introduction for scientists and engineers, Rudra Pratap, Oxford University Press.
- 3. Programming exercises based on the subjects covered in first semester

# **CSL-554: DISTRIBUTED OPERATING SYSTEM**

Internal Marks:	20	CI	REDI	ΤS
External Marks:	80	L	Т	Ρ
		4	0	0

# **Instructions for the Paper Setters:**

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

#### **SECTION-A**

Attributes of Distributed Operating System: Performance and Scalability, Connectivity and Security, Reliability and Fault Tolerance, Transparency, Network Operating Systems, Distributed Operating Systems.

**Communication in Distributed Systems:** Remote Procedure Call (RPC), Remote Method Invocation (RMI), CORBA (Common Object Request Broker Architecture), DCOM (Distributed Component Object Model, Process Migration in Distributed Systems.

#### **SECTION-B**

**Synchronization in Distributed Systems:** Mutual Exclusion in Distributed Systems : Mutual Exclusion without Shared Memory, Agrawala and Ricart's Distributed Mutual Exclusion Algorithm.

**Deadlock in Distributed Systems:** Distributed Deadlock, Deadlock Prevention, Deadlock Detection, A Distributed Resource Deadlock Algorithm.

#### SECTION-C

**Distrubuted Systems and Web Services:** Distributed File System Concepts, Network File System (NFS). Multicomputer Systems, Clustering, Distributed Computing, Grid Computing.

#### SECTION-D

**Security in Operating System:** Cryptography, Access Control, Security Attacks and Security Solutions, Key Agreement Protocols, Secure Communication Protocols, Steganography, Open Source Security.

- 1. Gary Nutt : *Operating System, Modern Perspective*, Addison Wesley.
- 2. A.S.Tanenbaum : *Modern Operating Systems*, Prentice Hall.
- 3. M. Milenkovic : Operating Systems, McGraw Hill.
- 4. Deitel and Choffnes : Operating Systems, 3<sup>rd</sup> Edition, Pearson Education.
- 5. Brian L. Stuart : Operating Systems Principles, Design and Applications, Cengage Learning.

# **CSL–555: NETWORK SECURITY**

Internal Marks:	20			CR	REDI	ΤS
External Marks:	80			L	Т	Ρ
				4	0	0

#### **Instructions for the Paper Setters:**

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

#### **SECTION-A**

**Introduction:** Attacks, services, mechanisms, security attacks, security services, Model for network security, Internet standards.

**Conventional encryption and message confidentiality:** Conventional encryption principles, conventional encryption algorithms, cipher block modes of operations, location of encryption devices, key distribution.

#### **SECTION-B**

**Public Key cryptography and authentication:** Approaches to message authentication, Secure Hash Functions and HMAC, Public Key Cryptography, Principles Public Key Cryptography Algorithms, Digital signatures, Key management.

Authentication & E mail Security: Kerberos, X.509 Directory Authentication Services-PGP-S/MIME.

#### **SECTION-C**

**IP** Security: IP security overview, IP Security Architecture, Authentication Header, Encapsulating Security Pay load, Combing Security Associations, Key Management.

#### **SECTION-D**

**Web Security:** Web Security Requirements, SSL and Transport Layer Security, SET Network Management Security.

System Security: Intruders, viruses related threats, Fire Design principles, Trusted Systems.

- 1. William Stallings, *Network Security Essentials Applications and Standards*, Pearson Education Asia, New Delhi, 2006
- 2. Kaufman, *Network Security: Private Communication in a Public World*, Pearson Education Asia, New Delhi, 2005.
- 3. William Stallings, *Cryptography and Network Security*, Pearson Education Asia, New Delhi, 2005.

# CSL-548: DATA MINING

Internal Marks:	20		CREDIT		ΤS
External Marks:	80		L	Т	Ρ
			4	0	0

#### **Instructions for the Paper Setters:**

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

#### **SECTION-A**

Introduction: Basic concepts of Data Mining, Related technologies (Machine Learning, DBMS, OLAP, Statistics), Data Mining Goals, Stages of the Data Mining Process, Data Mining Tasks, Knowledge Representation Methods, Applications of Data Mining, Major Challenges and Issues in Data Mining

Data preprocessing: Data cleaning, Data transformation, Data reduction, Discretization and generating concept hierarchies

#### **SECTION-B**

Association Rule Mining: Introduction and Basic Concepts, Motivation and terminology, Examples of Association rule mining, Basic Algorithms, Parallel and Distributed Algorithms, Comparing Approaches, Incremental Rules, Advanced Association Rule Techniques, Measuring the Quality of Rules

Classifications and Prediction: Basic Concepts, Decision Tree induction, Bayes Classification Methods, Rule Based Classification, Model Evaluation and Selection, Techniques to Improve Classification Accuracy

Classification Advanced concepts: Bayesian Belief Networks, Classification by Back Propagation, Support Vector Machine and Classification using frequent patterns

#### **SECTION-C**

Cluster Analysis: Basic concepts and Methods, Cluster Analysis, Partitioning methods, Hierarchical methods, Density based Methods, Grid Based Methods, Evaluation of Clustering

Advanced Cluster Analysis: Probabilistic model based clustering, Clustering High, Dimensional Data, Clustering Graph and Network Data, Clustering with Constraints

Outlier Analysis: Basic concepts of Outlier analysis, Types of Outliers, Challenges of Outlier Detection, Outlier Detection Methods, Statistical approaches, Proximity-Based Approaches, Clustering-Based Approaches, Classification-Based Approaches, Mining Contextual and Collective Outliers

#### **SECTION-D**

Text mining: Basic Concepts, Extracting attributes (Keywords), structural approaches (parsing, soft parsing)

Web Mining: Introduction, Classifying web pages, extracting knowledge from the web

Overview of Data Mining Software and Applications: Case Study: WEKA

#### **REFERENCES:**

- Jiawei Han, Micheline Kamber, Jain Pei, "Data Mining: Concepts and Techniques", Third Edition (The Morgan Kaufmann Series in Data Management System), 2012.
- ii. David J. Hand, Heikki Mannila and Padhraic Smyth "Principles of Data Mining" (Adaptive Computation and Machine learing), 2005
- iii. Margaret H Dunham, "Data Mining: Introductory and Advanced Topics", 2003
- iv. Soman, K.P., Diwakar Shyam and Ajay V. "Insight into Data Mining: Theory and Practices", PHI, 2009.

# CSL-560: COMPUTER SYSTEM ARCHITECTURE & ORGANIZATION

Internal Marks:	20	CREDITS		ΤS
External Marks:	80	L	Т	Ρ
		4	0	0

#### **Instructions for the Paper Setters:**

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

#### **SECTION-A**

**Introduction to Computers:** Basic Computer Organization: System Buses, Instruction Cycles **CPU Organization:** Design specifications for a Simple CPU, Fetching Instructions from Memory, Decoding Instructions, Executing Instructions, Designing the Control Unit Using Hardwired Control and Microprogrammed control approach.

**Parallelism in Uniprocessor Systems:** Trends in parallel processing, Basic Uniprocessor Architecture, Parallel Processing Mechanism.

#### **SECTION-B**

**Parallel Computer Structures**: Pipeline Computers, Array Computers, Multiprocessor Systems.

**Architectural Classification Schemes:** Multiplicity of Instruction-Data Streams, Serial versus Parallel Processing, Parallelism versus Pipelining.

**Pipelining:** An overlapped Parallelism, Principles of Linear Pipelining, Classification of Pipeline Processors.

**Principles of Designing Pipelined Processors:** Instruction Prefetch and Branch Handling, Data Buffering and Busing Structures, Internal Forwarding and Register Tagging, Hazard Detection and Resolution.

Superscalar and Superpipeline Design: Superscalar Pipeline Design, Superpipelined Design.

**SECTION-C** 

**Structures and Algorithms for Array Processors**: SIMD Array Processors, SIMD Computer Organizations, Inter-PE Communications.

**SIMD Interconnection Networks:** Static versus Dynamic Networks, Mesh-Connected Illiac Network, Cube Interconnection Networks.

# **SECTION-D**

**System Interconnect Architectures:** Network Properties and Routing, Static Connection Networks, Dynamic Connection Networks.

**Multiprocessor Architecture:** Functional Structures: Loosely Coupled Multiprocessors, Tightly Coupled Multiprocessors.

**Interconnection Networks**: Time Shared for Common Buses, Crossbar Switch and Multiport Memories.

- 1. *Computer Architecture and Parallel Processing*, Faye A. Briggs, McGraw-Hill International Editions.
- 2. *Computer Systems Organization & Architecture*, John d. Carpinelli, Addison Wesley.

# CSL-569: SOFTWARE ENGINEERING & TESTING

Internal Marks:	20		CREDITS		ΤS
External Marks:	80		L	Т	Ρ
			3	0	1

#### **Instructions for the Paper Setters:**

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

#### **SECTION-A**

**Introduction to S/W Engineering** - A layered technology – Software Process models - waterfall, RAD, Evolutionary, Concurrent, specialized - System Engineering – Computer Based System Overview, Business Process Engineering Overview – Product Engineering Overview.

**Introduction to S/W Testing** – Fundamentals of testing process, broad categories of testing – General principles of testing – Major Software Testing Techniques- White-box testing, basis path testing: flow graph notation, cyclomatic complexity; Control structure testing: condition testing, data-flow testing, loop testing ; Black Box testing : Graph based testing methods-BVA

#### **SECTION-B**

**Software Testing Strategies** – Approach-verification and validation; Strategic issues; testing conventional software – Unit testing, Integration testing, Validation testing, System testing; Debugging process, strategies, correcting error - The Testing Phases - Test strategy and Test plan – Test strategy template - Test plan template – Requirement traceability –Test scenario – Test Case.

#### **SECTION-C**

**Software Product metrics**: measures, metrics, indicators of quality challenge- measurement principles, attributes of software metrics- McCall and ISO 9126 factors- metrics for analysis model- metrics for design model-architectural design, object-oriented design metrics, CK, MOOD, LK metric suites, component-level design metrics, operation-oriented metrics and interface design metrics-metrics for source coding, metrics for testing, metrics for maintenance.

#### **SECTION-D**

**Test Estimation techniques**: Approaches of Test effort estimation, Delphi Technique, Analogy based estimation, Software size based estimation, Test case enumeration based estimation, Task (Activity) based Test estimation, Testing size based estimation, Sizing a Testing project, Merits and demerits of various Test Estimation techniques.

**Test Automation Tool**: Introduction to Selenium Tool, Selenium IDE, Selenium Remote control, Selenium Grid.

- 1. Pressman, R., Software Engineering, McGraw Hill, 2005(6e).
- 2. Humphrey, W., Managing Software Process, Pearson Education Asia, 1998.
- 3. Crosby, P.B., Quality is Free: The Art of Making Quality Certain, Mass Market, 1992.
- 4. Senn, J.A., Software Analysis and Design, McGraw Hill, 1989
- 5. Software Testing Foundations Andreas Spillner, Tilo Linz, Hans Schäfer
- 6. Software Estimation Best practices, Tools & Techniques Murali Chemuturi

# **CSL–567: WIRELESS NETWORKS**

Internal Marks:	20	CREDITS
External Marks:	80	LTP
		4 0 0

# **Instructions for the Paper Setters:**

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

# **SECTION-A**

Introduction to Wireless Systems: History, Overview of Wireless Networks, Types & Range of

Wireless Communication, Wireless Network Standards, Security Issues of Wireless Technology.

Wireless Communication Technology: Antennas & Propagation, Signal Encoding Techniques,

Spread Spectrum Coding and Error Control.

# **SECTION-B**

**Wireless Networking:** Satellite Communications, Cellular Wireless Networks, Cordless Systems and Wireless Local Loop, Mobile Communications, Mobile IP and Wireless Access Protocol.

**Wireless LANs:** Introduction, Benefits, WLANs Configurations and Standards, Security, IEEE 802.11, Wireless LAN Standard, Blue Tooth.

WAP: History of WAP, Architecture, and Components.

# SECTION-C

Cellular Technology: Design and Principles of Cellular Operation, Cellular Telephony Operations, GSM

# **SECTION-D**

Adhoc & Sensor Networks: Introduction, Protocols, and Applications.

- 1. Wireless Communications & Networks by Stallings (2005) Pearson Education.
- 2. Wireless Communication, Principles of Practice by Rappaport, Prentice Hall, 2<sup>nd</sup> Edition.
- 3. Introduction to Wireless Technology by Rogers (2003), Pearson Education.
- Data Over Wireless Networks: Blue tooth, WAP and Wireless LANs by Held (2001), Tata McGraw Hill (Osborne reprint)

# **CSP–569: TERM PAPER**

 Credits

 L
 T
 P

 0
 0
 1

# **CSL-564: ADVANCED SOFTWARE ENGINEERING**

Internal Marks:	20		CREDITS		ΤS
External Marks:	80		L	Т	Ρ
			4	0	0

#### **Instructions for the Paper Setters:**

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

#### SECTION-A

**Software Project Management:** Fundamentals of Software project planning, Conventional Software Management, Evolution of Software Economics, Improvement of Software Economics, Comparison of old and modern ways of Software Management.

#### **SECTION-B**

**Software Re-engineering:** Introduction Re-engineering, Restructuring and Reverse Engineering, Reengineering existing systems, Data Re-engineering and migration, Software Reuse and Re-engineering.

**Object-Oriented (OO) Measurements:** Introduction, Why metrics?, Classification of OO metrics, Study of Design Metrics- method size, method internals, class size, class inheritance, Method inheritance, class intervals and class externals.

#### SECTION-C

**Object-Oriented Analysis and Design:** What is Object-Oriented Design? Objects, Abstraction, Collaboration among Objects, Polymorphisms, Classes, specifying State, Specifying Behavior, Class Relationships, Grouping, Hiding.

#### **SECTION-D**

**Software Agents:** Definition, Applications, Types and Classes, Multi-Agent systems, Characteristics & Properties Agents.

- 1. Walker Royce, Software Project Management, Pearson Education, ISBN: 9780201309584, 2004.
- 2. Robert S. Arnold, Software Re-engineering, IEEE Comp. Society, ISBN: 9780818632723, 2003.
- 3. Lorenz and Kidd, Object Oriented Software Metrics, Prentice Hall, ISBN 978-0131792920, 2001.
- Booch, Object-Oriented Analysis and Design with Applications, Addison-Wesley Professional, 3<sup>rd</sup> Edition, ISBN: 978-0201895513, 2007.

# **CSL-565: NETWORK PROGRAMMING**

Internal Marks:	20		CR	EDI	ΤS
External Marks:	80		L	Т	Ρ
			4	0	0

#### **Instructions for the Paper Setters:**

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

#### **SECTION-A**

**Introduction:** TCP/IP Architecture, TCP/IP addressing, services, FTP, SMTP, TFTP, SNMP, Network file system, domain name system, transport layer protocols, user datagram protocol, transmission control protocol.

#### **SECTION-B**

**Interprocess Communications:** File and record locking, pipes, FIFO's, stream and messages, message quues, samphorers.

#### **SECTION-C**

**Sockets:** Sockets system cells, reserved parts, stream pipes, socket option, asynchronous I/O, Sockets and signals Transport Lay Interface : Elementary TLI functions, stream and stream pipes, asynchronous I/O multiplexing.

#### **SECTION-D**

**Remote Procedure Calls:** Remote login, remote command execution, external data representation.

- 1. A. Stevens, "TCP/IP Illustrated", Vol. 1-3, Pearson Education, 2004.
- 2. R. Stevens, "Unix Network Programming", PHI 2002.
- 3. D.E. Comer, "Internetworking with TCP/IP, Vol. 1, Vol. 2, Vol. 3 Principles, Protocols, and Architecture, PHI, 2000.

# **CSL566: ADVANCED CLOUD COMPUTING**

Internal Marks:	20			CR	EDI	ΤS
External Marks:	80			L	Т	Ρ
				4	0	0

#### **Instructions for the Paper Setters:**

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

#### **SECTION-A**

Introduction: Definition, Vision, Reference Model, Benefits, Limitations, Open Challenges, Grid and Utility Computing.

Virtualization: Definition, Type of Virtualization, Benefits, Limitations, Virtualization and Cloud, Virtual Appliance.

#### **SECTION-B**

Cloud Computing Architecture: Service Models, Deployment Models, Cloud Entities, Cloud Clients, Cloud Programming Models.

Cloud Terminology: Resource Provisioning, Bill Management, Multitenancy and Isolation, Service Level Agreement (SLA) and Quality of Service (QoS), Mobile Cloud Computing.

#### **SECTION-C**

Cloud Security: Infrastructure Security, Data Security, Identity and Access Management, Privacy Management, Security as a Service on Cloud.

#### **SECTION-D**

Big-Data and Internet of Things (IoT): Definition of Big-Data, Structured and Unstructured Data, V's of Big-Data, Hadoop, Definition of IoT, Characteristics of IoT, Combining Big-Data, IoT and Cloud Computing.

#### **Textbooks/Journals:**

- Rajkumar Buyya, Christian Vecchiola, and Thamarai Selvi, Mastering Cloud Computing: Foundation and Application Programming, Tata McGraw Hill, ISBN-13: 978-1-25-902995-0, New Delhi, India, Feb 2013.
- 2. Tim Mather, Subra Kumaraswamy and Shahed Latif, Cloud Security and Privacy, O'Reilly, ISBN-13: 978-8-18-404815-5.
- 3. Alan Nugent, Fern Halper, Judith Hurwitz and Marcia Kaufman, Big Data for Dummies, Wiley India, ISBN-13: 978-8-12-654328-1, April, 2013.
- Jayavardhana Gubbi, Rajkumar Buyya, Slaven Marusic and Marimuthu Palaniswami, Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions, Future Generation Computer Systems, vol. 29, no. 7, pp. 1645-1660, September, 2013.

#### **Reference Books:**

- Barrie Sosinsky, Cloud Computing Bible, Wiley India Pvt. Ltd, ISBN-13: 978-8-12-6529803, New Delhi, India, 2011.
- Dr. Saurabh Kumar, Cloud Computing: Insights Into New-Era Infrastructure, Wiley India Pvt. Ltd, ISBN-13: 978-8-12-6528837, New Delhi, India, 2011.
- Fern Halper, Hurwitz, Robin Bloor, Marcia Kaufman, Cloud Computing For Dummies, Wiley India Pvt. Ltd, ISBN-13: 978-0-47-0597422, New Delhi, India, 2011.

#### **CSL–656: Introduction to Data Science**

Internal Marks:	20			CR	REDI	ΤS	
External Marks:	80			L	Т	Ρ	
				3	0	1	

#### **Instructions for the Paper Setters:**

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

#### **SECTION-A**

- 1. Introduction: Data Science
  - Big Data and Data Science hype and getting past the hype
  - Why now? Datafication
  - Current landscape of perspectives
  - Skill sets needed
- 2. Statistical Inference
  - Populations and samples
  - Statistical modeling, probability distributions, fitting a model
- 3. Exploratory Data Analysis and the Data Science Process
  - Basic tools (plots, graphs and summary statistics) of EDA
  - Philosophy of EDA
  - The Data Science Process
  - Case Study: RealDirect (online real estate firm)

#### **SECTION-B**

- 4. Three Basic Machine Learning Algorithms
  - Linear Regression
  - k-Nearest Neighbors (k-NN)
  - k-means
- 5. More on Machine Learning Algorithm and Usage in Applications
  - Motivating application: Filtering Spam
  - Why Linear Regression and k-NN are poor choices for Filtering Spam
  - Naive Bayes and why it works for Filtering Spam
  - Data Wrangling: APIs and other tools for scrapping the Web

#### **SECTION-C**

6. Feature Generation and Feature Selection (Extracting Meaning from Data)

- Motivating application: user (customer) retention

- Feature Generation (brainstorming, role of domain expertise, and place for imagination)
- Feature Selection algorithms
- -Filters; Wrappers; Decision Trees; Random Forests
- 7. Mining Social-Network Graphs
  - Social networks as graphs
  - Clustering of graphs
  - Direct discovery of communities in graphs
  - Partitioning of graphs
  - Neighborhood properties in graphs

#### SECTION-D

- 8. Data Visualization
  - Basic principles, ideas and tools for data visualization
  - Examples of inspiring (industry) projects
  - Exercise: create your own visualization of a complex dataset
- 9. Data Engineering
  - -Map Reduce: word count problem, other examples of Map-reduce,
  - -Introduction to Pregel
  - -Hadoop Architecture
- 10. Data Science and Ethical Issues
  - Discussions on privacy, security, ethics
  - Next-generation data scientists

- 1. Cathy O'Neil and Rachel Schutt. Doing Data Science, Straight Talk From The Frontline. O'Reilly, 2014.
- 2. Jure Leskovek, Anand Rajaraman and Jerey Ullman. Mining of Massive Datasets. v2.1,Cambridge University Press, 2014.
- Kevin P. Murphy. Machine Learning: A Probabilistic Perspective. MIT Press-ISBN 0262018020, 2013.
- 4. Foster Provost and Tom Fawcett. Data Science for Business: What You Need to Know about Data Mining and Data-analytic Thinking. O'Reilly, ISBN 1449361323, 2013.
- 5. Trevor Hastie, Robert Tibshirani and Jerome Friedman. Elements of Statistical Learning, Springer- Verlag, Second Edition, ISBN 0387952845, 2009.
- 6. Mohammed J. Zakiand, Wagner, Miera Jr. Data Mining and Analysis: Fundamental Concepts and Algorithms, Cambridge University Press, 2014.
- 7. Jiawei Han, Micheline Kamber and Jian Pei. Data Mining: Concepts and Techniques, Morgan Kaufmann Publishers, Third Edition, ISBN 0123814790, 2011.

#### CSD-651: DISSERTATION (PART-I)

CR	EDI	ΤS
L	Т	Ρ
0	0	4

#### The dissertation will normally contain:

- 1. A clear indication, at appropriate stages, of original and creative elements. The level of originality expected is likely to include the application of existing techniques to new environments, the use of original materials, the re–working of existing materials, and the use of comparative approaches to the provision of information technology;
- 2. A discussion of its scope and aims, and its theoretical and professional significance, including a discussion of the context in which the problem is seen as important;
- 3. An analysis of the topic within a critical review of the relevant literature;
- 4. An evaluation of methods used in the dissertation, their reliability, validity, and a comparison with alternative methods;
- 5. An account of the process of obtaining the data required for the dissertation and the results obtained;
- 6. An analysis of the results of the dissertation to include a discussion of their significance, their relationship to other research, and any methodological or theoretical implications; and
- 7. The relationship of the findings to existing professional understanding and, where appropriate, potential implementation difficulties.

It is not intended to restrict students to a precisely defined format for the dissertation but it should follow the standard practices of dissertation writing. Although a written report will normally be expected, it should be accompanied by soft copy on CD.

The evaluation shall be done as per the common ordinances for courses under Credit Based Continuous Evaluation Grading System

#### **CSL650: PARALLEL COMPUTING**

Internal Marks:	20			CR	REDI	ΤS
External Marks:	80			L	Т	Ρ
				4	0	0

#### **Instructions for the Paper Setters:**

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

#### **SECTION-A**

**Paradigms of Parallel Computing:** Synchronous – Vector/Array, SIMD, Systolic **Asynchronous** – MIMD, reduction Paradigm, Hardware taxonomy: Flynn's classification, Handler's classification, Software taxonomy: Kung's taxonomy, SPMD.

#### **SECTION-B**

Abstract Parallel Computational Models: Combinational circuits, Sorting Network, PRAM Models, Interconnection RAMs.

**Parallelism approaches** – Data parallelism, control parallelism.

#### **SECTION-C**

**Performance Metrics**: Laws governing performance measurements: Amdahl's law, Gustafson's law, Sun–Ni law, Metrics such as Speedup, efficiency, iso-efficiency, utilization, sizeup, communication overheads etc., Benchmarks.

**Scheduling and Parallelization:** Load Scheduling, Types of scheduling algorithm, Load Balancing, Loop Scheduling, Parallelization of sequential programs.

#### **SECTION-D**

**Overview of Parallel Programming Development & Support Environments:** Shared memory programming, distributed memory programming, object oriented programming, data–parallel programming, functional and data flow programming.

- 1. Michael J. Quinn, *Parallel Computing: Theory and Practice*, Tata McGraw–Hill, 4th Edition, ISBN: 9780070512948, 2004.
- 2. C. Xavier and S. S. Iyenger, *Introduction to Parallel Algorithms*, Wiley–Interscience Publication, ISBN: 9780471251828, 1998.
- 3. Wilkinson: *Parallel Programming*, Prentice Hall, ISBN: 9780131405639, 2004.

#### CSL-652: ARTIFICIAL NEURAL NETWORKS

Internal Marks:	20	CREDITS
External Marks:	80	LTP
		4 0 0

#### **Instructions for the Paper Setters:**

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

#### **SECTION-A**

**Neural Network Technology:** Evolution of ANN, Architecture of ANN, Knowledge representation. **Neural Network Learning:** Basic learning rules, supervised & unsupervised learning, LMS Algorithm.

#### **SECTION-B**

**Single Layer Perceptrons–I:** Preceptron Model, Preceptron learning algorithms: Simple learning algorithm, pocket algorithm without and with Ratches, Linear Machines, Kessler's construction, Linear Machines Learning algorithm, Representing Boolean functions.

#### **SECTION-C**

**Single Layer Perceptrons–II:** Anderson's BSB Model, Hopfield's Model, K–Means Clustering, Topology–Preserving Maps, ART1 and ART2.

#### **SECTION-D**

**Multilayer Perceptrons:** Back–Propagation, Applications of Back–propagation: NETtalk, Handwritten Character Recognition, Pattern Recognition.

- 1. [SG] Gallant S.L., *Neural Networks Learning & Expert Systems*, MIT Press, ISBN: 9780262071451, 1993.
- [SH] Haykin S., Neural Networks: A comprehensive Foundation, Prentice Hall, 3<sup>rd</sup> Edition, ISBN: 9780131471399, 2007.
- 3. [FS] Freeman J.A., Skapura D.M., *Neural Network Algorithms, Applications and Programming Techniques*, Pearson Education, ISBN: 9780201513769, 2003.

#### **CSL-653: DIGITAL IMAGE PROCESSING**

Internal Marks:	20			CR	REDI	ΤS
External Marks:	80			L	Т	Ρ
				4	0	0

#### **Instructions for the Paper Setters:**

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

#### **SECTION-A**

**Introduction and Digital Image Fundamentals :** The origins of Digital Image Processing, Fundamentals Steps in Image Processing, Elements of Digital Image Processing Systems, Image Sampling and Quantization, Some basic relationships like Neighbours, Connectivity, Distance Measures between pixels, Linear and Non Linear Operations.

**Image Enhancement:** Point Operations, Histograms, Spatial Domain methods, Frequency domain methods, Enhancement by point processing, Spatial filtering, low pass filtering, High pass filtering, Homomorphic filtering, Colour Image Processing.

#### **SECTION-B**

Image Restoration Degradation Model, Algebraic approach to Restoration, Inverse Filtering, Wiener Filter, Constrained least square restoration, Interactive restoration, Restoration in spatial domain. **Image Compression:** Coding, Interpixel and Psychovisual Redundancy, Image Compression Models, Elements of Information Theory, Error free comparison Lossy Compression, Image Compression Standards.

#### **SECTION-C**

**Image Segmentation:** Detection of Discontinuities, Edge Linking and Boundary Detection, Thresholding, Region Oriented Segmentation, Motion based segmentation.

**Representation and Description:** Representation, Boundary Descriptors, Regional Descriptors, Use of Principal Components for Description, Introduction to Morphology, Some basic Morphological Algorithms.

#### **SECTION-D**

**Object Recognition:** Patterns and Pattern Classes, Decision – Theoretic Methods, Structural Methods.

#### **References:**

1. Rafael C. Gonzalez & Richard E. Woods, "Digital Image Processing", 2<sup>nd</sup> Edition.

2. A.K. Jain, "Fundamental of Digital Image Processing", PHI

#### **CSL-655: SOFTWARE PROJECT MANAGEMENT**

Internal Marks:	20			CR	EDI	ΤS	
External Marks:	80			L	Т	Ρ	
				4	0	0	

#### **Instructions for the Paper Setters:**

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

#### **SECTION-A**

Planning Fundamentals: Major issues in software project planning, Planning activities Project,

master schedule, Software risk management, Risk monitoring, Risk analysis

#### **SECTION-B**

**Software cost:** Major issues in estimating software cost, Cost estimation methods, Experience based model, Parameter based model, COCOMO, Versions of COCOMO, Software size estimation, Function points, Software project schedule, Rayleigh model.

# SECTION-C

**Functional organization:** Project organization, Matrix organization, Staffing, Quality replacements, Turnover management.

Directing a software engineering project: Issues, activities, Conflict management.

#### **SECTION-D**

**Issues in controlling a software project:** Controlling activities, Threads of control, Work breakdown structures, Earned value tracking

#### **References:**

1. Richard Thayer, Software Engineering Project Management, Tata Mc Graw Hill, 2004.

2. Donald J. Reifer, Software Management, Pearson Education, 2003.

# CSD-660: DISSERTATION (PART-II)

 Credits

 L
 T
 P

 0
 0
 12

# In continuation from 3<sup>rd</sup> Semester

The evaluation shall be done as per the common ordinances for courses under Credit Based Continuous Evaluation Grading System