

# VELTECH MULTI TECH

Dr RANGARAJAN Dr. SAKUNTHALA ENGINEERING COLLEGE

(Owned by Vel Trust 1997)

*(An ISO 9001: 2008 Certified Institution)*

Accredited By NAAC with 'A' Grade and NBA Accredited  
Institution

(Approved by AICTE New Delhi and Govt. of Tamil Nadu, Affiliated to  
Anna University Chennai)



## SYLLABUS

### WEEKLY SCHEDULE

M.E CSE

III SEMESTER 2017 - 2018

### DEPARTMENT OF CSE II YEARS DEGREE COURSE

42, Avadi – Alamathi Road,

Chennai – 600062

Telefax – 044-26841061

E-mail: [emailto@veltechmultitech.org](mailto:emailto@veltechmultitech.org)

Website : [www.veltechmultitech.org](http://www.veltechmultitech.org)

## **INSTITUTION VISION AND MISSION**

Accredited by NAAC with 'A' grade with an impressive score of 3.49.

### **INSTITUTION VISION**

- Elevating well being of humanity by augmenting human resource potential through quality technical education and training

### **INSTITUTION MISSION**

- To effectuate supremacy in technical education through articulation of research and industry practices for social relevance.
- To inculcate the habit of lifelong learning
- To exhibit professional ethics, commitment and leadership qualities

### **DEPARTMENT VISION**

- To emerge as centre for academic excellence in the field of Computer Science and Engineering by exposure to research and industry practices

### **DEPARTMENT MISSION**

- To provide good teaching and learning environment with conducive research atmosphere in the field of Computer Science and Engineering
- To propagate lifelong learning
- To impart the right proportion of knowledge, attitudes and ethics in students to enable them take up positions of responsibility in the society and make significant contributions

## **PROGRAMME EDUCATIONAL OBJECTIVES (PEOS)**

After 3 to 5 years of completing the Programme on Master of Computer Science and Engineering, the post graduates will become:

1. Competent Computer/Software Engineer rendering expertise to the industrial and societal needs in an effective manner.
2. Sustained learner to bring out novel ideas by addressing the research issues.
3. Trainer/Philosopher to guide others towards the development of technology.

## **PROGRAMME OUTCOMES (POS)**

During the course of the programme on Master of Computer Science and Engineering the learners will acquire the ability to:

1. Apply knowledge of mathematics, science and information science in computer engineering in advance level.
2. Design a Computer system with components and processes of desired needs within realistic constraints such as economic, environmental, social, political, ethical, health and safety.
3. Apply Software Engineering principles, techniques and tools in software development.
4. Design and conduct experiments, as well as to analyze and interpret data to lay a foundation for solving complex problems.
5. Identify and modify the functions of the internal of computer components such as operating systems and compilers.
6. Create, collect, process, view, organize, store, mine and retrieve data in both local and remote locations in a secure and effective manner.

## WEEK DETAILS

<b>SL.NO.</b>	<b>WEEK</b>	<b>FROM</b>	<b>TO</b>
<b>1</b>	<b>WEEK1</b>	<b>24.06.2017</b>	<b>24.06.2017</b>
<b>2</b>	<b>WEEK2</b>	<b>27.06.2017</b>	<b>01.07.2017</b>
<b>3</b>	<b>WEEK3</b>	<b>03.07.2017</b>	<b>08.07.2017</b>
<b>4</b>	<b>WEEK4</b>	<b>10.07.2017</b>	<b>15.07.2017</b>
<b>5</b>	<b>WEEK5</b>	<b>17.07.2017</b>	<b>22.07.2017</b>
<b>6</b>	<b>WEEK6</b>	<b>24.07.2017</b>	<b>29.07.2017</b>
<b>7</b>	<b>WEEK7</b>	<b>31.07.2017</b>	<b>05.08.2017</b>
<b>8</b>	<b>WEEK8</b>	<b>07.08.2017</b>	<b>12.08.2017</b>
<b>9</b>	<b>WEEK9</b>	<b>16.08.2017</b>	<b>19.08.2017</b>
<b>10</b>	<b>WEEK10</b>	<b>21.08.2017</b>	<b>26.08.2017</b>
<b>11</b>	<b>WEEK11</b>	<b>28.08.2017</b>	<b>01.09.2017</b>
<b>12</b>	<b>WEEK12</b>	<b>04.09.2017</b>	<b>09.09.2017</b>
<b>13</b>	<b>WEEK13</b>	<b>11.09.2017</b>	<b>16.09.2017</b>
<b>14</b>	<b>WEEK14</b>	<b>18.09.2017</b>	<b>23.09.2017</b>
<b>15</b>	<b>WEEK15</b>	<b>25.09.2017</b>	<b>28.09.2017</b>
<b>16</b>	<b>WEEK16</b>	<b>03.10.2017</b>	<b>07.10.2017</b>

## **SUBJECT CONTENTS**

<b>SL.NO</b>	<b>SUBJECT CODE</b>	<b>SUBJECT NAME</b>
<b>THEORY</b>		
<b>1</b>	<b>CP7301</b>	<b>SOFTWARE PROCESS AND PROJECT MANAGEMENT</b>
<b>2</b>	<b>CP7019</b>	<b>MANAGING BIG DATA</b>
<b>3</b>	<b>CP7024</b>	<b>INFORMATION RETRIEVAL TECHNIQUES</b>
<b>4</b>	<b>CP7029</b>	<b>INFORMATION STORAGE MANAGEMENT</b>
<b>PRACTICAL</b>		
<b>5</b>	<b>CP7311</b>	<b>Project Work (Phase I)</b>

### **TEST / EXAM SCHEDULE**

<b>SL.NO</b>	<b>SUBJECT CODE</b>	<b>SUBJECT NAME</b>	<b>UNIT TEST I</b>	<b>UNIT TEST II</b>	<b>Pre Model Exam</b>	<b>UNIT TEST 4</b>	<b>MODEL EXAM</b>
1	CP7301	SOFTWARE PROCESS AND PROJECT MANAGEMENT	13.07.2017	28.07.2017	16.08.2017	07.09.2017	28.09.2017
2	CP7019	MANAGING BIG DATA	14.07.2017	29.07.2017	17.08.2017	08.09.2017	04.10.2017
3	CP7024	INFORMATION RETRIEVAL TECHNIQUES	14.07.2017	29.07.2017	18.08.2017	09.09.2017	06.10.2017
4	CP7029	INFORMATION STORAGE MANAGEMENT	15.07.2017	30.07.2017	19.08.2017	10.09.2017	09.10.2017

# **CP7301 : SOFTWARE PROCESS AND PROJECT MANAGEMENT**

## **UNIT I DEVELOPMENT LIFE CYCLE PROCESSES**

**WEEK 1:** Overview of software development life cycle – introduction to processes – Personal Software Process (PSP)

**WEEK 2:** Team software process (TSP) – Unified processes – agile processes

**WEEK 3 :**choosing the right process Tutorial: Software development using PSP

**WEEK 4 :CYCLE TEST 1**

## **UNIT II REQUIREMENTS MANAGEMENT**

Functional requirements and quality attributes – elicitation techniques – Quality Attribute Workshops (QAW)

**WEEK 5:** analysis, prioritization, and trade-off – Architecture Centric Development Method (ACDM) – requirements documentation and specification

**WEEK 6:** change management – traceability of requirements  
Tutorial: Conduct QAW, elicit, analyze, prioritize, and document requirements using ACDM.

**WEEK 7: CYCLE TEST II**

## **UNIT III ESTIMATION, PLANNING, AND TRACKING**

**WEEK 8:** Identifying and prioritizing risks – risk mitigation plans – estimation techniques – use case points – function points – COCOMO II – top-down estimation.

**WEEK 09:** bottom-up estimation – work breakdown structure – macro and micro plans – planning poker – wideband delphi

**WEEK 10:** documenting the plan – tracking the plan – earned value method (EVM) Tutorial: Estimation, planning, and tracking exercises.

#### **UNIT IV CONFIGURATION AND QUALITY MANAGEMENT**

**WEEK11:** identifying artifacts to be configured – naming conventions and version control – configuration control

#### **WEEK 12: PREMODEL EXAM**

quality assurance techniques – peer reviews – Fegan inspection – unit, integration, system, and acceptance testing

**WEEK 13:** test data and test cases – bug tracking – causal analysis

Tutorial: version control exercises, development of test cases, causal analysis of defects

#### **UNIT V SOFTWARE PROCESS DEFINITION AND MANAGEMENT**

**WEEK 14:** Process elements – process architecture – relationship between elements – process modeling – process definition techniques ETVX (entry-task-validation-exit) – process base lining

**WEEK 15:**– process assessment and improvement CMMI – Six Sigma Tutorial: process measurement exercises, process definition using ETVX

#### **WEEK 16:–REVISION CLASS**

#### **REFERENCES:**

1. Pankaj Jalote, “Software Project Management in Practice”, Pearson, 2002.
2. Chris F. Kemerer, “Software Project Management – Readings and Cases”, McGraw Hill, 1997.
3. Watts S. Humphrey, “PSP: A self-improvement process for software engineers”, Addison-Wesley, 2005.
4. Watts S. Humphrey, “Introduction to the Team Software Process”, Addison-Wesley, 2000.
5. Orit Hazzan and Yael Dubinsky, “Agile software engineering”, Springer, 2008.
6. James R. Persse, “Process Improvement Essentials”, O’Reilly, 2006.



7. Roger S. Pressman, “Software Engineering – A Practitioner’s Approach”, Seventh Edition, McGraw Hill, 2010

## **CP7019 MANAGING BIG DATA**

### **UNIT I UNDERSTANDING BIG DATA**

**WEEK 1:** What is big data – why big data – convergence of key trends – unstructured data – industry examples of big data – web analytics – big data and marketing – fraud and big data – risk and big data

**WEEK 2:** credit risk management – big data and algorithmic trading – big data and healthcare – big data in medicine – advertising and big data – big data technologies – introduction to Hadoop

**WEEK 3:** open source technologies – cloud and big data – mobile business intelligence – Crowd sourcing analytics – inter and trans firewall analytics

**WEEK 4:** CYCLE TEST-I

### **UNIT II NOSQL DATA MANAGEMENT**

**WEEK 5:** Introduction to NoSQL – aggregate data models – aggregates – key-value and document data models – relationships – graph databases – schemaless databases – materialized views – distribution models – sharding – master-slave replication – peer-peer replication – sharding and replication

**WEEK 6:** consistency – relaxing consistency – version stamps – map-reduce – partitioning and combining

**WEEK 7:** composing map-reduce calculations

**CYCLE TEST-II**

### **UNIT III BASICS OF HADOOP**

**WEEK 8:** Data format – analyzing data with Hadoop – scaling out – Hadoop streaming – Hadoop pipes

**WEEK 9:** design of Hadoop distributed file system (HDFS) – HDFS concepts – Java interface – data flow

**WEEK 10:** Hadoop I/O – data integrity – compression- serialization – Avro – file-based data structures

**WEEK 11:** PREMODEL EXAM

## **UNIT IV MAPREDUCE APPLICATIONS**

**WEEK 12:**MapReduce workflows – unit tests with MRUnit – test data and local tests – anatomy ofMapReduce job run- classic Map-reduce – YARN – failures in classic Map-reduce and YARN

**WEEK 13:**job scheduling – shuffle and sort – task execution – MapReduce types – input formats – outputformats

### **CYCLE TEST-IV**

## **UNIT V HADOOP RELATED TOOLS**

**WEEK 14:**Hbase – data model and implementations – Hbase clients – Hbase examples – praxis.Cassandra– cassandra data model – cassandra examples – cassandra clients – Hadoop integration.

**WEEK 15:**Pig – Grunt – pig data model – Pig Latin – developing and testing Pig Latin scripts.Hive – data types and file formats – HiveQL data definition – HiveQL data manipulation – HiveQLQueries

### **WEEK 16:MODEL EXAM**

### **REFERENCES:**

1. Michael Minelli, Michelle Chambers, and AmbigaDhiraj, "Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses", Wiley, 2013.
2. P. J. Sadalage and M. Fowler, "NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence", Addison-Wesley Professional, 2012.
3. Tom White, "Hadoop: The Definitive Guide", Third Edition, O'Reilley, 2012.
4. Eric Sammer, "Hadoop Operations", O'Reilley, 2012.
5. E. Capriolo, D. Wampler, and J. Rutherglen, "Programming Hive", O'Reilley, 2012.
6. Lars George, "HBase: The Definitive Guide", O'Reilley, 2011.
7. Eben Hewitt, "Cassandra: The Definitive Guide", O'Reilley, 2010.
8. Alan Gates, "Programming Pig", O'Reilley, 2011.

# **CP7024 INFORMATION RETRIEVAL TECHNIQUES**

## **UNIT I INTRODUCTION**

**WEEK 1:** Motivation – Basic Concepts – Practical Issues - Retrieval Process – Architecture – Boolean Retrieval – Retrieval Evaluation – Open Source IR Systems–History of Web Search

**WEEK 2:** Web Characteristics–The impact of the web on IR —IR Versus Web Search–Components of a Searchengine

**WEEK 3: CYCLE TEST I**

## **UNIT II MODELING**

**WEEK 4:** Taxonomy and Characterization of IR Models – Boolean Model – Vector Model - Term Weighting

**WEEK 5:** Scoring and Ranking –Language Models – Set Theoretic Models - Probabilistic Models

**WEEK 6:** Algebraic Models – Structured Text Retrieval Models – Models for Browsing

**WEEK 7: CYCLE TEST II**

## **UNIT III INDEXING**

**WEEK 8:** Static and Dynamic Inverted Indices – Index Construction and Index Compression Searching

**WEEK 9:** Sequential Searching and Pattern Matching. Query Operations -Query Languages

**WEEK 10:** Query Processing - Relevance Feedback and Query Expansion - Automatic Local and Global Analysis Measuring Effectiveness and Efficiency.

**WEEK 11: PREMODEL EXAM**

## **UNIT IV CLASSIFICATION AND CLUSTERING**

**WEEK 12:** Text Classification and Naïve Bayes – Vector Space Classification – Support vector machines and Machine learning on documents.

**WEEK 13:** Flat Clustering – Hierarchical Clustering –Matrix decompositions and latent semantic indexing – Fusion and Meta learning

**CYCLE TEST IV**

### **UNIT V SEARCHING AND RANKING**

**WEEK 14:** Searching the Web –Structure of the Web –IR and web search – Static and Dynamic Ranking -Web Crawling and Indexing – Link Analysis

**WEEK 15:** XML Retrieval Multimedia IR: Models and Languages – Indexing and Searching Parallel and Distributed IR – Digital Libraries

**WEEK 16: MODEL EXAM**

### **REFERENCES:**

1. Ricardo Baeza – Yates, BerthierRibeiro – Neto, Modern Information Retrieval: The concepts and Technology behind Search (ACM Press Books), Second Edition 2011
2. Christopher D. Manning, PrabhakarRaghavan, HinrichSchutze, Introduction to InformationRetrieval, Cambridge University Press, First South Asian Edition 2012
3. Stefan Buttcher, Charles L. A. Clarke, Gordon V. Cormack, Information RetrievalImplementing and Evaluating Search Engines, The MIT Press, Cambridge, Massachusetts London, England, 2010

## **CP7029 INFORMATION STORAGE MANAGEMENT**

### **UNIT I INTRODUCTION TO STORAGE TECHNOLOGY**

**WEEK 1:** Review data creation and the amount of data being created and understand the value of data to a business

**WEEK 2:** challenges in data storage and data management, Solutions available for data storage

**WEEK 3:** Core elements of a data center infrastructure, role of each element in supporting business activities.

## **UNIT II STORAGE SYSTEMS ARCHITECTURE**

**WEEK 3:** Hardware and software components of the host environment, Key protocols and concepts used by each component  
,Physical and logical components of a connectivity environment  
,Major physical components of a disk drive and their function

### **WEEK 4: CYCLE TEST-I**

logical constructs of a physical disk, access characteristics, and performance Implications, Concept of RAID and its components, Different RAID levels and their suitability for different application environments:

**WEEK 5:**RAID 0, RAID 1, RAID 3, RAID 4, RAID 5, RAID 0+1, RAID 1+0, RAID 6

**WEEK 6:**Compare and contrast integrated and modular storage systems ,High-level architecture and working of an intelligent storage system.

## **UNIT III INTRODUCTION TO NETWORKED STORAGE**

### **WEEK 7: CYCLE TEST II**

Evolution of networked storage, Architecture, components, and topologies of FC-SAN, NAS, and IP-SAN

**WEEK 8 :**Benefits of the different networked storage options, understand the need for long-term archiving solutions and describe how CAS full fill the need

**WEEK 9:**understand the appropriateness of the different networked storage options for different application environments

**WEEK10:**

## **UNIT IV INFORMATION AVAILABILITY, MONITORING & MANAGING DATACENTER**

List reasons for planned/unplanned outages and the impact of downtime, Impact of downtime - Differentiate between business

continuity (BC) and disaster recovery (DR) ,RTO and RPO, Identify single points of failure in a storage infrastructure and list solutions to mitigate these failures

**WEEK 11.**Architecture of backup/recovery and the different backup/recovery topologies, replication technologies and their role in ensuring information availability and business continuity, Remote replication technologies and their role in providing disaster recovery and business continuity capabilities

### **WEEK 12 :PREMODEL EXAM**

Identify key areas to monitor in a data center, Industry standards for data center monitoring and management, Key metrics to monitor for different components in a storage infrastructure, Key management tasks in a data center

## **UNIT V SECURING STORAGE AND STORAGE VIRTUALIZATION 9**

**WEEK 13:** block-level and file-level virtualization technologies and processes

**WEEK 14:** Information security, Critical security attributes for information systems, Storage security domains

### **WEEK 15:**

List and analyzes the common threats in each domain, Virtualization technologies

### **WEEK 16: :REVISION CLASS AND MODEL EXAM**

#### **REFERENCE BOOKS:**

1. EMC Corporation, Information Storage and Management, Wiley, India.
2. Robert Spalding, “Storage Networks: The Complete Reference“, Tata McGraw Hill ,Osborne, 2003.
3. Marc Farley, “Building Storage Networks”, Tata McGraw Hill ,Osborne, 2001.

4. Additional resource material on [www.emc.com/resource-library/resource-library.esp](http://www.emc.com/resource-library/resource-library.esp)

\*\*\*\*\*