D.K.T.E. Society's

TEXTILE & ENGINEERING INSTITUTE

(An Autonomous Institute)
Rajwada, Ichalkaranji – 416115.

Syllabus
of
Data Science (Honors)

(With effect from June 2020)

D.K.T.E. Society's

TEXTILE & ENGINEERING INSTITUTE

(An Autonomous Institute)
Rajwada, Ichalkaranji – 416115.

Department of Computer Science and Engineering

Data Science (Honors)

(With effect from Academic Year 2020-21)

Syllabus Structure

Sr. No.	Course Code	Course Name	Sem	Teaching Scheme Hours/Week				Examination Scheme and Marks						Credits
								Theory			Practical			
				Theory	Tutorial	Practical	Total	SE-I	SE-II	SEE	CIE	SEE	Total	
1	CSL701	Basic Statistics	IV	2	1	-	3	25	25	50	-	-	100	03
2	CSL702	Exploratory Data Analysis and Feature Engineering	V	3	-	-	3	25	25	50	-	-	100	03
3	CSP703	Introduction to Data Science in Python	V	2	-	2	4	-	-	-	50	50	100	03
4	CSL704	Big Data Analytics	VI	3	-	-	3	25	25	50	-	-	100	03
5	CSP705	Applied Text Mining in Python	VI	2	-	2	4	-	-	-	50	50	100	03
6	CSP706	Time Series Analysis	VII	2	-	2	4	-	-	-	50	50	100	03
7	CSD707	Capstone Project	VII	-	-	2	2	-	-	-	50	50	100	02
Total				14	1	8	23	75	75	150	200	200	700	20

DKTES Textile and Engineering Institute, Ichalkaranji

(An Autonomous Institute)

Teaching and evaluation Scheme for year 2020-21

Third Year B. Tech. (Semester – IV) In Data Science Honors for Computer Science and Engineering, Electronics, Electronics and Telecommunication

Sr. No.	Course Code	Course Title	Course Category	Teaching scheme				Course Credits	Evaluation scheme					
									Theory			Practical		
				L	T	P	Contact Hrs./wk.		CIE		SEE	CIE	SEE	TOTAL
									SE-I	SE-II				
1	CSL701	Basic Statistics	BSC	2	1	-	3	3	25	25	50	-	-	100
		Total		**2**	**1**	**-**	**3**	**3**	**25**	**25**	**50**	**-**	**-**	**100**

L- Lecture
T-Tutorial
P-Practical

SE-I : Semester Examination-I
SE-II : Semester Examination-II

CIE – Continuous In Semester Evaluation
SEE- Semester End Examination

Course Category	HSMC (Humanities, Social Science & Management Course)	BSC (Basic Science Course)	ESC (Engineering Science Course.)	PCC (Professional Core Courses)	PEC (Professional Elective Courses)	OEC (Open Elective. Courses)	MC (Mandatory Courses)	PST (Project / Seminar / Ind. Training)
Credits	--	3	--	--	--	--	--	--
Cumulative Sum	--	--	--	--	--	--	--	--

Progressive Total Credits 03 + 00 = 03

DKTES Textile and Engineering Institute , Ichalkaranji
Second Year B. Tech.(Semester – IV)
CSL701: Basic Statistics

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 02 Hrs./Week	03	SE-I: 25 Marks
Tutorials: 01 Hrs./Week		SE-II: 25 Marks
Practicals: 00 Hrs./Week		SEE: 50 Marks

Course Outcomes:

On completion of the course, student will be able to–

- describe the basics of statistics
- illustrate the methods of descriptive statistics.
- explain basics of probability
- illustrate methods of inferential statistics

Course Contents

Unit I	Exploring Data	06 Hours

Cases, variables and levels of measurement, Data matrix and frequency table, Graphs and shapes of distributions, Mode, median and mean, Range, interquartile range and box plot, Variance and standard deviation, Z-scores, Example

Unit II	Correlation and Regression	06 Hours

Crosstabs and scatterplots, ,Pearson's r, Regression - Finding the line, Regression - Describing the line, Regression - How good is the line?, Correlation is not causation, Example contingency table, Example Pearson's r and regression

Unit III	Probability	06 Hours

Randomness, Probability, Sample space, event, probability of event and tree diagram, Quantifying probabilities with tree diagram, Basic set-theoretic concepts, Practice with sets ,Union, Joint and marginal probabilities, Conditional probability, Independence between random events ,More conditional probability, decision trees and Bayes' Law

Unit IV	Probability Distributions	06 Hours

Random variables and probability distributions, Cumulative probability distribution , The mean of a random variable, Variance of a random variable, Functional form of the normal distribution, The normal distribution: probability calculations, The standard normal distribution, The binomial distribution

Unit V	Sampling Distributions	06 Hours

Sample and population, Sampling, The sampling distribution, The central limit theorem, Three distributions, Sampling distribution proportion, Example

Unit VI	Confidence Intervals and Significance Tests	09 Hours
Confidence Intervals-Statistical inference, CI for mean with known population sd, CI for mean with unknown population sd, CI for proportion, Confidence levels, Choosing the sample size, Example. Significance Tests - Hypotheses, Test about proportion, Test about mean, Step-by-step plan, Significance test and confidence interval, Type I and Type II errors, Example.		
Text Books:		
--		
References Books:		
--		
Useful Links:		
1. https://www.coursera.org/learn/basic-statistics		

DKTES Textile and Engineering Institute, Ichalkaranji

(An Autonomous Institute)

Teaching and evaluation Scheme for year 2020-21

Third Year B. Tech. (Semester – V) In Data Science Honors for Computer Science and Engineering, Electronics, Electronics and Telecommunication

Sr. No.	Course Code	Course Title	Course Category	Teaching scheme				Course Credits	Evaluation scheme					
									Theory			Practical		
				L	T	P	Contact Hrs./wk.		CIE		SEE	CIE	SEE	TOTAL
									SE-I	SE-II				
1	CSL702	Exploratory Data Analysis and Feature Engineering	PCC	3	-	-	3	3	25	25	50	-	-	100
2	CSP703	Introduction to Data Science in Python	PCC	2	-	2	4	3	-	-	-	50	50	100
		Total		**5**	**-**	**2**	**7**	**6**	**25**	**25**	**50**	**50**	**50**	**200**

L- Lecture
T-Tutorial
P-Practical

SE-I: Semester Examination-I
SE-II: Semester Examination-II

CIE – Continuous In Semester Evaluation
SEE- Semester End Examination

Course Category	HSMC (Humanities, Social Science & Management Course)	BSC (Basic Science Course)	ESC (Engineering Science Course.)	PCC (Professional Core Courses)	PEC (Professional Elective Courses)	OEC (Open Elective. Courses)	MC (Mandatory Courses)	PST (Project / Seminar / Ind. Training)
Credits	--	03	--	--	--	--	--	--
Cumulative Sum	--	--	--	06	--	--	--	--

Progressive Total Credits 03 + 06 = 09

DKTES Textile and Engineering Institute, Ichalkaranji

Third Year B. Tech. (Semester – V)

CSL702: Exploratory Data Analysis and Feature Engineering

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 03 Hrs./Week	03	SE-I: 25 Marks
Tutorials: 00 Hrs./Week		SE-II: 25 Marks
Practicals: 00 Hrs./Week		SEE: 50 Marks

Course Objectives:

- To make students to learn data analysis and visualization concepts
- To make students to learn theoretical foundation of Hypothesis Testing and Analysis of Variance
- To make students to learn different methods of feature selection
- To make students to learn how to Reduce feature space in a dataset

Course Outcomes:

On completion of the course, student will be able to–

- Explain exploratory data analysis and visualization techniques
- Explain the theoretical foundation of Hypothesis Testing and Analysis of Variance
- Explain different methods of feature selection
- Explain how to Reduce feature space in a dataset

Course Contents

Unit I	Fundamentals of Exploratory Data Analysis	05 Hours

Exploratory Data Analysis Fundamentals

Understanding data science, Significance of EDA, Making sense of data, Comparing EDA with classical and Bayesian analysis, Software tools available for EDA

Visual Aids for Exploratory Data Analysis

Line chart, Bar charts, Scatter plot, Area plot and stacked plot, Pie chart, Table chart, Polar chart, Histogram, Lollipop chart, Choosing the best chart

Unit II	Hypothesis Testing and Analysis of Variance	06 Hours

Descriptive and inferential statistics, Kernel density estimate, Cumulative distribution function, Hypothesis testing-T-test, CHI-squared and Fisher's test, Analysis of Variance (ANOVA), One Way ANOVA, Two Way ANOVA

Unit III	Exploratory Data Analysis	06 Hours

Typical data format and the types of EDA, Univariate non-graphical EDA, Univariate graphical EDA, Multivariate non-graphical EDA, Multivariate graphical EDA, EDA Example, EDA for Text Data

Unit IV	Feature Construction and Feature Selection	06 Hours

Feature Construction

Examining our dataset, Imputing categorical features, Encoding categorical variables, Extending numerical features, Text-specific feature construction

Feature Selection

Importance of Feature Selection in Machine Learning, Goals of Feature Selection, Classes of Feature Selection Methodologies, Effect of Irrelevant Feature, Overfitting to Predictors and External Validation, Greedy Search Methods- Simple Filters, Recursive Feature Elimination, Stepwise Selection

Unit V	Feature Transformations	05 Hours
Dimension reduction – feature transformations versus feature selection versus feature construction, Principal Component Analysis, How centering and scaling data affects PCA, A deeper look into the principal components, Linear Discriminant Analysis, LDA versus PCA – iris dataset		
Unit VI	**Feature Learning**	**06 Hours**
Parametric assumptions of data, Non-parametric fallacy, feature learning algorithms, Reconstructing the data, The Bernoulli RBM, Extracting PCA components from MNIST, Extracting RBM components from MNIST, Using RBMs in a machine learning pipeline, Learning text features – word vectorizations, Word embeddings, Application of word embeddings – information Retrieval		

Textbooks:

1. Suresh Kumar Mukhiya, Usman Ahmed, "Hands-On Exploratory Data Analysis with Python", Packt Publishing, ISBN 978-1-78953-725-3
2. Sinan Ozdemir, Divya Susarla, "Feature Engineering Made Easy", Packt Publishing, ISBN 978-1-78728-760-0
3. Howard J .Seltman, "Experimental Design and Analysis", http://www.stat.cmu.edu/~hseltman/309/Book/Book.pdf
4. Max Kuhn , Kjell Johnson, "Feature Engineering and Selection: A Practical Approach for Predictive Models" 1st Edition, Chapman & Hall/CRC Data Science Series, ISBN 13-978-1-138-07922-9

References Books:

1. John W. Tukey, "Exploratory Data Analysis1st Edition", Pearson Education, ISBN 0134995457, 9780134995458

Useful Links:

1. https://www.coursera.org/learn/exploratory-data-analysis
2. https://www.kaggle.com/pavansanagapati/a-simple-tutorial-on-exploratory-data-analysis
3. https://www.kaggle.com/learn/feature-engineering
4. https://machinelearningmastery.com/discover-feature-engineering-how-to-engineer-features-and-how-to-get-good-at-it/

DKTES Textile and Engineering Institute, Ichalkaranji
Third Year B. Tech. (Semester – V)
CSP703: Introduction to Data Science in Python

Lab Scheme: Practical: 02 Hrs./Week	Credits 01	Evaluation Scheme: CIE: 50 Marks SEE: 50 Marks

Course Outcomes:

On completion of the course, student will be able to–

- Understand techniques such as lambdas and manipulating Comma Separated Files (CSV) files
- Describe common Python functionality and features used for Data Science
- Query Data Frame structures for cleaning and processing
- Explain distributions, sampling, and t-tests

UNIT-I	Fundamentals of Data Manipulation with Python	06 Hours

Python Functions, Python Types and Sequences, Python More on Strings, Python Demonstration: Reading and Writing CSV files, Python Dates and Times, Advanced Python Objects, map(),Advanced Python Lambda and List Comprehensions, Numerical Python Library (NumPy),Manipulating Text with Regular Expression

UNIT-II	Basic Data Processing with Pandas	06 Hours

Introduction to Pandas, The Series Data Structure, Querying a Series, DataFrame Data Structure, DataFrame Indexing and Loading, Querying a DataFrame, Indexing Dataframes, Missing Values, Example: Manipulating DataFrame

UNIT-III	More Data Processing with Pandas	06 Hours

Merging Dataframes, Pandas Idioms, Group by, Scales, Pivot Table,Date/Time Functionality

UNIT-IV	Answering Questions with Messy Data	06 Hours

Basic Statistical Testing, Other Forms of Structured Data

List of Experiments
(It should consist of 10-12 experiments based on the following topics.)

1	Write a Python program to demonstrate array creation techniques
2	Write a Python program to demonstrate indexing in Numpy array.
3	Write a Python function to find the Max in Numpy array, sum all the numbers in Numpy array, find average of numbers in Numpy array.
4	Write a Python program to demonstrate basic operations on single array and multiple arrays.
5	Write a Python program to demonstrate unary and binary operators in Numpy.
6	Write a Python program to demonstrate lambda technique.
7	Write a Python program to import data from Comma Separated Files (CSV) file, manipulate data, and export data in CSV file.
8	Write a Python program to demonstrate string manipulation and regular expressions.
9	Write a Python program to demonstrate Viewing/Inspecting Data, Selection, Data Cleaning, Filter, Sort, Groupby, Join/Combine, and Statistics in Dataframe.

10	Write a Python program to demonstrate filtering data stored in Dataframe (Single condition filtering, Multiple condition filtering).
11	Write a Python to get a list of the column headers from a Pandas DataFrame, delete DataFrame columns by name or index, add new column to existing DataFrame.
12	Write a Python program to demonstrate cleaning and processing data in Dataframe.
13	Write a Python program to visualize data using data visualization library Matplotlib or Seaborn.
14	Write a Python program to demonstrate One sample t-test, two sampled t-test, Paired sampled t-test.
15	Write a Python program to demonstrate Analysis of Variance (ANOVA).
16	Write a program to generate a normally distributed random variable, Binomial Distribution distributed random variable, and Bernoulli Distribution random variable.

DKTES Textile and Engineering Institute, Ichalkaranji

(An Autonomous Institute)

Teaching and evaluation Scheme for year 2020-21

Third Year B. Tech. (Semester – VI) In Data Science Honors for Computer Science and Engineering, Electronics, Electronics and Telecommunication

Sr. No.	Course Code	Course Title	Course Category	Teaching scheme				Course Credits	Evaluation scheme					
									Theory			Practical		TOTAL
				L	T	P	Contact Hrs./wk.		CIE		SEE	CIE	SEE	
									SE-I	SE-II				
1	CSL704	Big Data Analytics	PCC	3	-	-	3	3	25	25	50	-	-	100
2	CSL705	Applied Text Mining in Python	PCC	2	-	2	4	3	-	-	-	50	50	100
		Total		**5**	**-**	**2**	**7**	**6**	**25**	**25**	**50**	**50**	**50**	**200**

L- Lecture
T-Tutorial
P-Practical

SE-I : Semester Examination-I
SE-II : Semester Examination-II

CIE – Continuous In Semester Evaluation
SEE- Semester End Examination

Course Category	HSMC (Humanities, Social Science & Management Course)	BSC (Basic Science Course)	ESC (Engineering Science Course.)	PCC (Professional Core Courses)	PEC (Professional Elective Courses)	OEC (Open Elective. Courses)	MC (Mandatory Courses)	PST (Project / Seminar / Ind. Training)
Credits	--	--	--	06	--	--	--	--
Cumulative Sum	--	03	--	06	--	--	--	--

Progressive Total Credits 09 + 06= 15

DKTES Textile and Engineering Institute, Ichalkaranji
Third Year B. Tech. (Semester – VI)
CSL704: Big Data Analytics

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 03 Hrs./Week	03	SE-I: 25 Marks
Tutorials: 00 Hrs./Week		SE-II: 25 Marks
Practicals: 00 Hrs./Week		SEE: 50 Marks

Course Outcomes:

On completion of the course, student will be able to–

- Explore the fundamental concepts of big data and its analytics
- Analyze the big data using Hadoop and intelligent techniques
- Apply NoSQL big data management
- Recognize the suitable secure models for building competitive business decisions

Course Contents

Unit I	Importance of Big Data	06 Hours

Classification of Digital Data, Characteristics of Data, Evolution of Big Data, Definition of Big Data, Challenges with Big Data, What is Big Data? Other Characteristics of Data Which are not Definitional Traits of Big Data, Why Big Data? Traditional Business Intelligence (BI) versus Big Data, A Typical Data Warehouse Environment, A Typical Hadoop Environment, What is New Today? What is Changing in the Realms of Big Data? What is Big Data Analytics?, What Big Data Analytics isn't, Classification of Analysis, Challenges that prevent business from capitalizing Big Data, Top challenges facing Big Data, Why is Big Data analytics important?

Unit II	Hadoop Architecture	06 Hours

Hadoop ecosystem, Design of Hadoop distributed file system (HDFS), Anatomy of MapReduce job run, Classic Map-reduce, YARN, Failures in classic Map reduce and YARN, Data format, Analyzing data with Hadoop, Scaling out, Hadoop streaming, Unit tests with MapReduce

Unit III	Hadoop I/O	06 Hours

Data Integrity - Data Integrity in HDFS, LocalFileSystem, ChecksumFileSystem, Compression – Codecs, Compression and Input Splits, Using Compression in MapReduce, Serialization - The Writable Interface, Writable Classes, Implementing a Custom Writable, Serialization Frameworks, Avro File-Based Data Structures – SequenceFile, MapFile

Unit IV	NoSQL Management	06 Hours

Introduction to NoSQL, Impedance mismatch, Emergence of NoSQL, Aggregate data models, Key-value and document data models, Column-family stores, Graph databases, Schemaless databases, Distribution models-sharding, Master-slave replication, Peer-peer replication, Sharding and replication Relaxing consistency – CAP Theorem, Relaxing durability

Unit V	Analytics Framework	06 Hours

Applications on Big Data Using Pig and Hive, Data processing operators in Pig, Hive services – HiveQL, Querying Data in Hive, Fundamentals of HBase and ZooKeeper

Unit VI	Securing Ecosystem	06 Hours

Steps to secure big data, Classifying Data, Protecting–Big Data, Configuring Kerberos for Hadoop, Securing ecosystem components – Pig, Hive, Oozie, Flume

Text Books:

1. Michael Mineli, Michele Chambers, Ambiga Dhiraj, “Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses” (Wiley Publication)
2. Tom White, “Hadoop: The Definitive Guide” (O’Reilly Media)
3. P. J. Sadalage, M. Flower, “NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence” (Addison-Wesley)
4. Sudeesh Narayanan, “Securing Hadoop” (O’Reilly Media)

References Books:

1. Seema Acharya, Subhasini Chellappan, “Big Data Analytics”, Wiley.
2. Chris Eaton,Dirk derooset al., “Understanding Big data”, McGraw Hill.
3. G James, D. Witten, T Hastie, R. Tibshirani, “An Introduction to Statistical Learning: with Applications in R”, Springer.
4. Douglas Eadline, “Hadoop 2 Quick-Start Guide: Learn the Essentials of Big Data Computing in the Apache Hadoop 2 Ecosystem”, Pearson Education.
5. E. Capriolo, D. Wampler, J. Rutherglen, “Programming Hive”, O’ Reilly.
6. Lars George, “HBase: The Definitive Guide”, O’ Reilly.
7. Alan Gates, "Programming Pig”, O’ Reilly

Useful Links:

1. Analytics Vidhya (http://www.analyticsvidhya.com/) ...
2. Dataversity (http://www.dataversity.net/) ...
3. R Bloggers (http://www.r-bloggers.com/) ...
4. SmartData Collective (http://www.smartdatacollective.com/) ...
5. Data Science Central (http://www.datasciencecentral.com/) ...
6. Planet Big Data (http://planetbigdata.com/)

DKTES Textile and Engineering Institute, Ichalkaranji
Third Year B. Tech. (Semester – VI)
CSP705: Applied Text Mining in Python

Lab Scheme: Lecture: 01 Hrs./Week Practical: 02 Hrs./Week	Credits 03	Evaluation Scheme: CIE: 50 Marks SEE: 50 Marks

Course Outcomes:

On completion of the course, student will be able to–

- Understand how text is handled in Python
- Apply basic natural language processing methods
- Write code that groups documents by topic
- Describe the NLTK framework for manipulating text

Unit I	Working with Text in Python	06 Hours

Introduction to Text Mining, Handling Text in Python, Regular Expressions
Demonstration: Regex with Pandas and Named Groups, Internationalization and Issues with Non-ASCII Characters

Unit II	Basic Natural Language Processing	06 Hours

Basic Natural Language Processing, Basic NLP tasks with NLTK, Advanced NLP tasks with NLTK

Unit III	Classification of Text	06 Hours

Text Classification, Identifying Features from Text, Naive Bayes Classifiers, Naive Bayes Variations, Support Vector Machines, Learning Text Classifiers in Python, Demonstration: Case Study - Sentiment Analysis

Unit IV	Topic Modeling	06 Hours

Semantic Text Similarity, Topic Modeling, Generative Models and LDA, Information Extraction

List of Experiments
(It should consist of 10-12 experiments based on the following topics.)

1	Write a Python program to demonstrate Pattern matching in Python with Regex.
2	Write a Python program to demonstrate use of regular expression methods match(), search(), findall().
3	Write a Python program to Validate phone numbers, email address and social security number.
4	Write a Python program to parse HTML document and find all hyperlinks.
5	Write a Python program to discover abstract "topics" that occur in a collection of documents.
6	Write a Python program to classify document in specified category.
7	Write a Python program to determine Similarity between Documents.
8	Write a Python program to predict if a message is spam or not.
9	Write a Python program to demonstrate Basic NLP Tasks with NLTK
10	Write a Python program to cluster documents using clustering algorithms.
11	Write a Python program to identify Named Entities (NE) from document corpus.

12	Write a Python for spelling recommender function that uses nltk to find words similar to the misspelling.
13	Write a Python program to generate word embedding for documents corpus using Word2vec algorithm.
14	Write a Python program to identify sentiment of given text.

DKTES Textile and Engineering Institute, Ichalkaranji

(An Autonomous Institute)

Teaching and evaluation Scheme for year 2020-21

Third Year B. Tech. (Semester – VII) In Data Science Honors for Computer Science and Engineering, Electronics, Electronics and Telecommunication

Sr. No.	Course Code	Course Title	Course Category	Teaching scheme				Course Credits	Evaluation scheme					
									Theory			Practical		TOTAL
				L	T	P	Contact Hrs./wk.		CIE		SEE	CIE	SEE	
									SE-I	SE-II				
1	CSP706	Time Series Analysis	PCC	2	-	2	4	3	-	-	-	50	50	100
2	CSD707	Capstone Project	PCC	-	-	2	2	2	-	-	-	50	50	100
		Total		**-**	**-**	**4**	**6**	**5**	**-**	**-**	**-**	**100**	**100**	**200**

L- Lecture
T-Tutorial
P-Practical

SE-I: Semester Examination-I
SE-II: Semester Examination-II

CIE – Continuous In Semester Evaluation
SEE- Semester End Examination

Course Category	HSMC (Humanities, Social Science & Management Course)	BSC (Basic Science Course)	ESC (Engineering Science Course.)	PCC (Professional Core Courses)	PEC (Professional Elective Courses)	OEC (Open Elective. Courses)	MC (Mandatory Courses)	PST (Project / Seminar / Ind. Training)
Credits	--	--	--	05	--	--	--	--
Cumulative Sum	--	03	--	12	--	--	--	--

Progressive Total Credits 15 + 05= 20

DKTES Textile and Engineering Institute, Ichalkaranji
Final Year B. Tech. (Semester – VII)
CSP706: Time Series Analysis

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 02 Hrs./Week Tutorials: 00 Hrs./Week Practicals: 02 Hrs./Week	03	CIE: 50 Marks SEE: 50 Marks

Course Outcomes:

On completion of the course, student will be able to–

- Explain different software engineering processes.
- Describe various concepts of software engineering problem domain.
- Describe basics concepts of software engineering solution domain.
- Apply the software engineering principles to the give problem

Course Contents

Unit I	**Basic Statistics**	**03 Hours**

Getting Started in R: Download and Install R on Windows, Getting Started in R: Using Packages Concatenation, Five-number summary, Standard Deviation, Histogram in R, Scatterplot in R, Simple Linear Regression, More Linear Regression, Inference

Unit II	**Visualizing Time Series, and Beginning to Model Time Series**	**04 Hours**

Introduction, Time Plots, Autocovariance function, Autocovariance coefficient, Autocorrelation function, Introduction to Moving Averages, Simulating MA(2) process

Unit III	**Stationarity, MA(q) and AR(p) processes**	**05 Hours**

Stationarity - Intuition and Definition, Stationarity - First Examples...White Noise and Random Walks Stationarity - First Examples ACF of Moving Average, Series and Series Representation, Backward shift operator, Introduction to Invertibility, Duality, Mean Square Convergence, Autoregressive Processes - Definition, Simulation, and First Examples Autoregressive Processes - Backshift Operator and the ACF, Difference equations, Yule - Walker equations

Unit IV	**AR(p) processes, Yule-Walker equations, PACF**	**05 Hours**

Partial Autocorrelation and the PACF First Examples, Partial Autocorrelation and the PACF - Concept Development, Yule-Walker Equations in Matrix Form, Yule Walker Estimation - AR(2) Simulation
Yule Walker Estimation - AR(3) Simulation, Recruitment data - model fitting, Johnson & Johnson-model fitting

Unit V	Akaike Information Criterion (AIC), Mixed Models, Integrated Models	05 Hours
Akaike Information Criterion and Model Quality, ARMA Models, ARMA Properties and Examples ARIMA Processes, Q-Statistic, Daily births in California in 1959		
Unit VI	**Seasonality, SARIMA, Forecasting**	**05 Hours**
SARIMA processes, ACF of SARIMA models, SARIMA fitting: Johnson & Johnson, SARIMA fitting: Milk production, SARIMA fitting: Sales at a souvenir shop, Forecasting Using Simple Exponential Smoothing Double Exponential Smoothing, Triple Exponential Smoothing Concept Development, Triple Exponential Smoothing Implementation		

Text Books:

--

References Books:

--

Useful Links:

1. https://www.coursera.org/learn/practical-time-series-analysis?ranMID=40328&ranEAID=vedj0cWlu2Y&ranSiteID=vedj0cWlu2Y-EHCYZFT7gt_kCfSbJHQ6DA&siteID=vedj0cWlu2Y-EHCYZFT7gt_kCfSbJHQ6DA&utm_content=10&utm_medium=partners&utm_source=linkshare&utm_campaign=vedj0cWlu2Y
2. https://onlinecourses.nptel.ac.in/noc21_ch28/preview

DKTES Textile and Engineering Institute, Ichalkaranji
Final Year B. Tech. (Semester – VII)
CSP707: Capstone Project

Lab Scheme:	Credits	Evaluation Scheme:
Practical: 02 Hrs./Week	02	CIE: 50 Marks SEE: 50 Marks

Course Outcomes:

On completion of the course, student will be able to–

- A team of student will analyze the problem statement
- A team of student will build the SRS and design document
- A team of student will develop the code according to the design
- A team of student will test the developed software
- A team of student will write the report.

Student will form the group for the capstone project. The group will submit the completed project work to the department at the end of semester VII as mentioned below.

1. The workable project.
2. The project report in all respect with the following : -

i. Problem specifications
ii. System definition – requirement analysis.
iii. System design – dataflow diagrams, database design
iv. System implementation – algorithm, code documentation
v. Test results and test report.
vi. In case of object oriented approach – appropriate process be followed.

CIE will be jointly assessed by a panel of teachers appointed by head of the institution. SEE examination will be conducted by internal and external examiners as appointed by the CoE.