ACE

Engineering College

Ankushapur, Ghatkesar, Medchal, Hyderabad - 501301, T. S.

(An Autonomous Institution)

Department of CSE (Data Science)

COURSE STRUCTURE & SYLLABUS

Applicable from AY 2022-23 Batch (ACE-R22)

IV YEAR I SEMESTER

S. No.	Course Code	Course Title	L	T	P	Credits
1	DS701PC	Predictive Analytics	3	0	0	3
2	DS702PC	Web and Social Media Analytics	3	0	0	3
3	DS743PE	Natural Language Processing (PE – IV)	3	0	0	3
4	DS753PE	Data Science Applications (PE – V)	3	0	0	3
5	RS&GIS	REMOTE SENSING & GIS (OE – II)	3	0	0	3
6	DS703PC	Predictive Analytics Lab	0	0	2	1
7	DS704PC	Web and Social Media Analytics Lab	0	0	2	1
8	DS705PC	Project Stage – I	0	0	6	3
		Total Credits	**15**	**0**	**10**	**20**

DS701PC: PREDICTIVE ANALYTICS

B.Tech. IV Year I Sem.

L	T	P	C
3	0	0	3

Prerequisite:

1. Data Science/Data analytics

Course Objectives:

- To learn the basics and applications of predictive analytics using different techniques

Course Outcomes:

- Understand the processing steps for predictive analytics
- Construct and deploy prediction models with integrity
- Explore various techniques (machine learning/data mining, ensemble) for predictive analytics.
- Apply predictive analytics to real world examples.

UNIT - I

Introduction – types of analytics, applications of predictive analytics, overview of predictive analytics. Setting up the problem - processing steps, business understanding, objectives, data for predictive modeling, columns as measures, target variables, measures of success for predictive models.

UNIT - II

Prediction effect, deployment of prediction model, ethics and responsibilities The Data effect

UNIT - III

Machine Learning for prediction

Predictive modeling – decision trees, logistic regression, neural network, kNN, Bayesian method,

Regression model

Assessing Predictive models - Batch Approach to Model Assessment, Percent Correct Classification, Rank-Ordered Approach to Model Assessment, Assessing Regression Models

UNIT - IV

Ensemble effect

Model ensembles – motivation, wisdom of crowds, Bagging, Boosting, Random forests, stochastic gradient boosting, heterogeneous ensembles.

UNIT - V

Case studies: Survey analysis, question answering– challenges in text mining, persuasion by the numbers

TEXT BOOKS:

1. Eric Siegel, Predictive analytics- the power to predict who will Click, buy, lie, or die, John Wiley & Sons, 2013.
2. Dean Abbott, Applied Predictive Analytics - Principles and Techniques for the Professional Data Analyst, 2014.

REFERENCE BOOKS:

1. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning-Data Mining, Inference, and Prediction, Second Edition, Springer Verlag, 2009.
2. G. James, D. Witten, T. Hastie, R. Tibshirani-An introduction to statistical learning with applications in R, Springer, 2013.
3. E. Alpaydin, Introduction to Machine Learning, Prentice Hall of India, 2010.

DS702PC: WEB AND SOCIAL MEDIA ANALYTICS

B.Tech. IV Year I Sem.

L	T	P	C
3	0	0	3

Course Objectives:

- Exposure to various web and social media analytic techniques.

Course Outcomes:

- Knowledge on decision support systems
- Apply natural language processing concepts on text analytics
- Understand sentiment analysis
- Knowledge on search engine optimization and web analytics

UNIT - I

An Overview of Business Intelligence, Analytics, and Decision Support

Analytics to Manage a Vaccine Supply Chain Effectively and Safely, Changing Business Environments and Computerized Decision Support, Information Systems Support for Decision Making, The Concept of Decision Support Systems (DSS), Business Analytics Overview, Brief Introduction to Big Data Analytics

UNIT - II

Text Analytics and Text Mining

Machine Versus Men on Jeopardy: The Story of Watson, Text Analytics and Text Mining Concepts and Definitions, Natural Language Processing, Text Mining Applications, Text Mining Process, Text Mining Tools

UNIT - III

Sentiment Analysis

Sentiment Analysis Overview, Sentiment Analysis Applications, Sentiment Analysis Process, Sentiment Analysis and Speech Analytics

UNIT - IV

Web Analytics, Web Mining

Security First Insurance Deepens Connection with Policyholders, Web Mining Overview, Web Content and Web Structure Mining, Search Engines, Search Engine Optimization, Web Usage Mining (Web Analytics), Web Analytics Maturity Model and Web Analytics Tools

UNIT - V

Social Analytics and Social Network Analysis

Social Analytics and Social Network Analysis, Social Media Definitions and Concepts, Social Media Analytics

Prescriptive Analytics - Optimization and Multi-Criteria Systems:

Multiple Goals, Sensitivity Analysis, What-If Analysis, and Goal Seeking

TEXT BOOK:

1. Ramesh Sharda, Dursun Delen, Efraim Turban, Business Intelligence and Analytics: Systems for Decision Support, Pearson Education

REFERENCE BOOKS:

1. Rajiv Sabherwal, Irma Becerra- Fernandez," Business Intelligence– Practice, Technologies and Management", John Wiley 2011.
2. Lariss T. Moss, ShakuAtre, "Business Intelligence Roadmap", Addison-Wesley It Service.
3. Yuli Vasiliev, "Oracle Business Intelligence: The Condensed Guide to Analysis and Reporting", SPD Shroff, 2012.

DS743PE: NATURAL LANGUAGE PROCESSING (Professional Elective – IV)

B.Tech. IV Year I Sem.

L	T	P	C
3	0	0	3

Prerequisites:

- Data structures and compiler design

Course Objectives:

- Introduction to some of the problems and solutions of NLP and their relation to linguistics and statistics.

Course Outcomes:

- Show sensitivity to linguistic phenomena and an ability to model them with formal grammars.
- Understand and carry out proper experimental methodology for training and evaluating empirical NLP systems
- Able to manipulate probabilities, construct statistical models over strings and trees, and estimate parameters using supervised and unsupervised training methods.
- Able to design, implement, and analyze NLP algorithms; and design different language modeling Techniques.

UNIT - I

Finding the Structure of Words: Words and Their Components, Issues and Challenges, Morphological Models

Finding the Structure of Documents: Introduction, Methods, Complexity of the Approaches, Performances of the Approaches, Features

UNIT - II

Syntax I: Parsing Natural Language, Treebanks: A Data-Driven Approach to Syntax, Representation of Syntactic Structure, Parsing Algorithms

UNIT – III

Syntax II: Models for Ambiguity Resolution in Parsing, Multilingual Issues

Semantic Parsing I: Introduction, Semantic Interpretation, System Paradigms, Word Sense

UNIT - IV

Semantic Parsing II: Predicate-Argument Structure, Meaning Representation Systems

UNIT - V

Language Modeling: Introduction, N-Gram Models, Language Model Evaluation, Bayesian parameter estimation, Language Model Adaptation, Language Models- class based, variable length, Bayesian topic based, Multilingual and Cross Lingual Language Modeling

TEXT BOOKS:

1. Multilingual natural Language Processing Applications: From Theory to Practice – Daniel M. Bikel and Imed Zitouni, Pearson Publication

REFERENCE BOOK:

1. Speech and Natural Language Processing - Daniel Jurafsky& James H Martin, Pearson Publications.
2. Natural Language Processing and Information Retrieval: Tanvier Siddiqui, U.S. Tiwary.

DS753PE: DATA SCIENCE APPLICATIONS (Professional Elective – V)

B.Tech. IV Year I Sem.

L	T	P	C
3	0	0	3

Course Objective:

- To give deep knowledge of data science and how it can be applied in various fields to make life easy.

Course Outcomes:

- Correlate data science and solutions to modern problems.
- Decide when to use which type of technique in data science.

UNIT - I

Data Science Applications in various domains, Challenges and opportunities, tools for data scientists Recommender systems – Introduction, methods, application, challenges.

UNIT - II

Time series data – stock market index movement forecasting. Supply Chain Management – Real world case study in logistics

UNIT - III

Data Science in Education, social media

UNIT - IV

Data Science in Healthcare, Bioinformatics

UNIT - V

Case studies in data optimization using Python.

TEXT BOOKS:

1. Aakanksha Sharaff, G.K. Sinha, “Data Science and its applications “, CRC Press, 2021.
2. Q.A. Menon, S.A. Khoja, “Data Science: Theory, Analysis and Applications”, CRC Press, 2020

CE700OE: REMOTE SENSING & GIS (Open Elective - II)

B.Tech. Civil Engg. IV Year I Sem.	**L**	**T/P/D**	**C**
	3	**0/0/0**	**3**

Course Objectives: The objectives of the course are to

- Know the concepts of Remote Sensing, its interpreting Techniques and concepts of Digital images
- know the concept of Geographical Information System (GIS), coordinate system GIS Data and its types
- Understand the students managing the spatial Data Using GIS.
- Understand Implementation of GIS interface for practical usage.

Course Outcomes: After the completion of the course student should be able to:

- **Describe** different concepts and terms used in Remote Sensing and its data
- Understand the Data conversion and Process in different coordinate systems of GIS interface
- **Evaluate** the accuracy of Data and implementing a GIS
- **Understand the applicability** of RS and GIS for various applications

UNIT – I

Concepts of Remote Sensing Basics of remote sensing- elements involved in remote sensing, electromagnetic spectrum, remote sensing terminology & units, energy resources, energy interactions with earth surface features & atmosphere, atmospheric effects, satellite orbits, Sensor Resolution, types of sensors. Remote Sensing Platforms and Sensors, IRS satellites.

Remote Sensing Data Interpretation Visual interpretation techniques, basic elements, converging evidence, interpretation for terrain evaluation, spectral properties of soil, water and vegetation. Concepts of Digital image processing, image enhancements, qualitative & quantitative analysis and pattern recognition, classification techniques and accuracy estimation.

UNIT- II:

Introduction to GIS: Introduction, History of GIS, GIS Components, GIS Applications in Real life, The Nature of geographic data, Maps, Types of maps, Map scale, Types of scale, Map and Globe, Co- ordinate systems, Map projections, Map transformation, Geo-referencing,

UNIT- III:

Spatial Database Management System: Introduction: Spatial DBMS, Data storage, Database structure models, database management system, entity-relationship model, normalization

Data models and data structures: Introduction, GIS Data model, vector data structure, raster data structure, attribute data, geo-database and metadata,

UNIT- IV:

Spatial Data input and Editing: Data input methods – keyboard entry, digitization, scanning, conversion of existing data, remotely sensed data, errors in data input, Data accuracy, Micro and Macro components of accuracy, sources of error in GIS.

Spatial Analysis: Introduction, topology, spatial analysis, vector data analysis, Network analysis, raster data analysis, Spatial data interpolation techniques

UNIT- V: Implementing a GIS and Applications

Implementing a GIS: Awareness, developing system requirements, evaluation of alternative systems, decision making using GIS

Applications of GIS

GIS based road network planning, Mineral mapping using GIS, Shortest path detection using GIS, Hazard Zonation using remote sensing and GIS, GIS for solving multi criteria problems, GIS for business applications.

TEXT BOOKS

1. Remote Sensing and GIS by Basudeb Bhatta, Oxford University Press, 2nd Edition, 2011.
2. Introduction to Geographic Information systems by Kang-tsung Chang, McGraw Hill Education (Indian Edition), 7th Edition, 2015.
3. Fundamentals of Geographic Information systems by Michael N. Demers, 4th Edition, Wiley Publishers, 2012.

REFERENCE BOOKS

1. Remote Sensing and Image Interpretation by Thomas M. Lillesand and Ralph W. Kiefer, Wiley Publishers, 7th Edition, 2015.\
2. Geographic Information systems – An Introduction by Tor Bernhardsen, Wiley India Publication, 3rd Edition, 2010.
3. Advanced Surveying: Total Station, GIS and Remote Sensing by Satheesh Gopi, R. Sathi Kumar, N. Madhu, Pearson Education, 1st Edition, 2007.
4. Textbook of Remote Sensing and Geographical Information systems by M. Anji Reddy,

DS703PC: PREDICTIVE ANALYTICS LAB

B.Tech. IV Year I Sem. **L T P C**
0 0 2 1

Course Objectives:

- To learn the basics and applications of predictive analytics using different techniques

Course Outcomes:

- Understand the processing steps for predictive analytics
- Construct and deploy prediction models with integrity
- Explore various techniques (machine learning/data mining, ensemble) for predictive analytics.
- Apply predictive analytics to real world examples.

List of Experiments: Following experiments to be carried out using Python/SPSS/SAS/R/Power BI

1. Simple Linear regression
2. Multiple Linear regression
3. Logistic Regression
4. CHAID
5. CART
6. ARIMA – stock market data
7. Exponential Smoothing
8. Hierarchical clustering
9. Ward's method of clustering
10. Crowdsource predictive analytics- Netflix data

TEXT BOOKS:

1. Eric Siegel, Predictive analytics- the power to predict who will Click, buy, lie, or die, John Wiley & Sons, 2013.
2. Dean Abbott, Applied Predictive Analytics - Principles and Techniques for the Professional Data Analyst, 2014.

REFERENCE BOOKS:

1. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning-Data Mining, Inference, and Prediction, Second Edition, Springer Verlag, 2009.
2. G. James, D. Witten, T. Hastie, R. Tibshirani-An introduction to statistical learning with applications in R, Springer, 2013
3. E. Alpaydin, Introduction to Machine Learning, Prentice Hall of India, 2010

DS704PC: WEB AND SOCIAL MEDIA ANALYTICS LAB

B.Tech. IV Year I Sem.

L	T	P	C
0	0	2	1

Course Objectives:

- Exposure to various web and social media analytic techniques.

Course Outcomes:

- Knowledge on decision support systems
- Apply natural language processing concepts on text analytics
- Understand sentiment analysis
- Knowledge on search engine optimization and web analytics

List of Experiments

1. Preprocessing text document using NLTK of Python
   a. Stopword elimination
   b. Stemming
   c. Lemmatization
   d. POS tagging
   e. Lexical analysis
2. Sentiment analysis on customer review on products
3. Web analytics
   a. Web usage data (web server log data, clickstream analysis)
   b. Hyperlink data
4. Search engine optimization- implement spamdexing
5. Use Google analytics tools to implement the following
   a. Conversion Statistics
   b. Visitor Profiles
6. Use Google analytics tools to implement the Traffic Sources.

Resources:

1. Stanford core NLP package
2. GOOGLE.COM/Analytics

TEXT BOOK:

1. Ramesh Sharda, Dursun Delen, Efraim Turban, Business Intelligence and Analytics: Systems for Decision Support, Pearson Education

REFERENCE BOOKS:

1. Rajiv Sabherwal, Irma Becerra-Fernandez, “Business Intelligence–Practice, Technologies and Management”, John Wiley 2011.
2. Lariss T. Moss, Shaku Atre, “Business Intelligence Roadmap”, Addison-Wesley It Service.
3. Yuli Vasiliev, “Oracle Business Intelligence: The Condensed Guide to Analysis and Reporting”, SPD Shroff, 2012.