



Scheme – 2020

Department of Computer Science & Engineering

**G. Pulla Reddy Engineering College (Autonomous):
Kurnool**

Accredited by NBA of AICTE and NAAC of UGC

Affiliated to JNTUA, Anantapuramu

Scheme and Syllabus for
Honors in
COMPUTER SCIENCE & ENGINEERING

(With Effect from the Batch Admitted in 2020-21)

Honors in Computer Science & Engineering (CSE)

Scheme of instruction and examination

(Effective from 2020-2021)

S No.	Semester	Course Title	Credits	Scheme of instruction (periods/ week)		Scheme of Examination		
				L	P	End Exam	Internal Assessment	Total 100
1	IV	Foundations of AI & Machine Learning	4	4	0	60	40	100
2	V	Python for Data Science	4	3	2	60	40	100
3	VI	Social Data Analytics	4	3	2	60	40	100
4	VII	Natural Language Processing	4	3	2	60	40	100
5		MOOC – 1	2	0	0			100
6		MOOC – 2 / Mini Project	2	0	0			100
		Total	20					

MOOCS - 1	
1	Scalable Data Science
2	Information Security
3	Cyber Security
4	Internet of things
MOOCS – 2 / Mini Project	
1	Computer Vision
2	Virtual Reality
3	Database security
4	Fuzzy Logic & Neural Networks

FOUNDATIONS OF AI & MACHINE LEARNING (FAIM)

IV Semester : Honors						Scheme : 2020		
Course Code	Category	Hours/Week			Credits	Maximum Marks		
HCS01	H	L	T	P	C	Continuous Internal Assessment	End Exam	TOTAL
		4	0	-	4	40	60	100
Sessional Exam Duration : 1 ½ Hrs					End Exam Duration: 3 Hrs			
Course Outcomes :At the end of the course the student will be able to								
CO1: Understand the basic concepts of Linear Algebra.								
CO2:Interpret the importance of vectors in Machine learning.								
CO3:Understand the applications of probability and distribution in ML and AI.								
CO4:Understand the types of Machine Learning algorithms and their applications.								
CO5: Describe the foundations of AI and their applications.								
UNIT – I								
Linear Algebra:Systems of Linear Equations, Matrices, Solving Systems of Linear Equations, Vector Spaces, Linear Independence, Basis and Rank, Linear Mappings, Affine Spaces.								
UNIT – II								
Vector Calculus:Differentiation of Univariate Functions, Partial Differentiation and Gradients, Gradients of Vector-Valued Functions, Gradients of Matrices, Useful Identities for Computing Gradients, Backpropagation and Automatic Differentiation, Higher-Order Derivatives, Linearization and Multivariate Taylor Series.								
UNIT – III								
Probability and Distributions:Construction of a Probability Space, Discrete and Continuous Probabilities, Sum Rule, Product Rule and Baye’s Theorem, Summary Statistics and Independence, Gaussians Distribution.								
UNIT – IV								
Introduction to Machine Learning:What is Machine Learning? Examples of Machine Learning Applications – Learning Associations, Classification ,Types of Classification, Simple Linear Regression, Multiple Linear Regression, Unsupervised Learning Examples, Reinforcement Learning								
UNIT – V								
Introduction to AI :WhatIsAI?,TheFoundationsofArtificialIntelligence								
IntelligentAgents:AgentsandEnvironments,GoodBehaviour:TheConceptofRationality,TheNature ofEnvironments,And theStructureofAgents, Example problems.								
Text Books:								
1. Mathematics for Machine Learning, M. P. Deisenroth, A.A. Faisal, C. S. Ong. Cambridge								

University Press, 2020.
2. Introduction to Machine Learning, Ethem Alpaydin, MIT Press, Second Edition, 2010.
3. Stuart Russell and Peter Norvig, "Artificial Intelligence: A Modern Approach", Third Edition, 2010. Pearson Education.
Reference Books:
1. Introduction to the Theory of Statistics , A. M. Mood, F.A. Graybill and D.C. Boes, McGraw Hill Education, 2017.
2. A First Course in Probability, S.M Rose, Prentice Hall, 2012.
3. Introduction to Machine Learning with Python, Andreas C. Müller & Sarah Guido, O'Reilly publications, 2016.
4. Artificial Intelligence, Elaine Richie, Kevin Knight, 2008, 3rd Edition, TMH.
Web References:
1. https://www.tutorialspoint.com/machine_learning/machine_learning_tutorial.pdf
2. https://www.w3schools.com/python/python_ml_getting_started.asp
Question Paper Pattern:
<p>Sessional Exam:</p> <p>The question paper for sessional examination shall be for 25 marks, covering half of the syllabus for first sessional and remaining half for second sessional exam. The question paper shall consist of three sections with Two Questions (EITHER/ OR Type) in each section. The student shall answer one question from each section.</p> <p>End Exam:</p> <p>The question paper for End examination shall be for 60 marks. The Question paper shall contain Five Units with Two Questions (Either or Type) from each unit. Each of these questions may contain sub question and the student should answer any one question from each unit. Each Question carries 12 marks..</p>

PYTHON FOR DATA SCIENCE (PDS)

V Semester : Honors CSE/CST/CSBS						Scheme : 2020		
Course Code	Category	Hours/Week			Credits	Maximum Marks		
HCS02	H	L	T	P	C	Continuous Internal Assessment	End Exam	TOTAL
		3		2	4	40	60	100
Sessional Exam Duration : 1 ½ Hrs					End Exam Duration: 3 Hrs			
Course Outcomes : At the end of the course the student will be able to								
CO1: Write python programs using the core concepts like Lists, Dictionaries, sets, tuple, functions and regular expressions.								
CO2: Demonstrate various mathematical operations on arrays using NumPy								
CO3: Analyze and manipulate Data using Pandas								
CO4: Creating static and interactive visualizations using Matplotlib.								
CO5: Enumerate machine learning algorithms, Describe the Classification and Clustering								
UNIT – I								
Introduction to Python: Data Types:Strings, Numbers, Booleans, Date and Time, Lists, Tuples, Dictionaries,Operators,Conditional Statements, Loops, Functions, Modules and packages, Classes and Objects, Regular expressions								
UNIT – II								
Introduction to NumPy: The Basics of NumPy Arrays, Computation on NumPy Arrays: Universal Functions, Aggregations: Min, Max, and Everything in Between, Computation on Arrays: Broadcasting, Comparisons, Masks, and Boolean Logic, Fancy Indexing, Sorting Arrays, Structured Data: NumPy's Structured Arrays								
UNIT – III								
Data Manipulation with Pandas: Introducing Pandas Objects, Data Indexing and Selection, Operating on Data in Pandas, Handling Missing Data, Hierarchical Indexing, Combining Datasets: Concat and Append, Combining Datasets: Merge and Join, Aggregation and Grouping Planets Data, Pivot Tables, Vectorized String Operations, High-Performance Pandas								
UNIT – IV								
Visualization with Matplotlib: Two Interfaces for the Price of One, Simple Line Plots, Simple Scatter Plots, Visualizing Errors, Density and Contour Plots, Histograms, Binnings, and Density, Customizing Plot Legends, Customizing Colorbars, Multiple Subplots, Text and Annotation, Customizing Ticks, Customizing Matplotlib: Configurations and Stylesheets, Three-Dimensional plotting in Matplotlib								
UNIT – V								
Machine Learning: What Is Machine Learning?, Types of machine learning systems, Introducing Scikit-Learn, Feature Engineering, Model development, Linear Regression: Simple Linear Regression, Example of model development.								

Text Books:
1. Python Basics: With Illustrations from the Financial Market,Vivek Krishnamoorthy, Jay Parmar, Mario Pisa Pena,AQuantInsti® Publication, 2020
2. Python Data Science Handbook:Essential Tools for Working with Data,JakeVanderPlas, O'reilly publications, 2016
Reference Books:
1. Python® for Programmers,PaulDeitel,HarveyDeitel, Pearson Education, Inc,2019
2. Data Science & Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data,David Dietrich, Barry Heller, Beibei Yang, Published by John Wiley & Sons, Inc, 2015
Question Paper Pattern:
Sessional Exam: The question paper for sessional examination shall be for 25 marks, covering half of the syllabus for first sessional and remaining half for second sessional exam. The question paper shall consist of three sections with Two Questions (EITHER/ OR Type) in each section. The student shall answer one question from each section.
End Exam: The question paper for End examination shall be for 60 marks. The Question paper shall contain Five Units with Two Questions (Either or Type) from each unit. Each of these questions may contain sub question and the student should answer any one question from each unit. Each Question carries 12 marks.

<i>List of Experiments</i>
1. Data Types: Operations on Number data type, sets and tuples.
2. Operators: Example programs using all operators.
3. Conditional Statements: Usage of if-else -if statement, break, continue and pass.
4. Loop Statements: Armstrong number, Palindrome.
5. Loop Statements: Bubble sort using lists.
6. Functions: Calculator program.
7. Functions: Programs on anonymous functions.
8. NumPy: Arithmetic Operations on Arrays.

SOCIAL DATA ANALYTICS (SDA)

VI Semester : Honors					Scheme : 2020			
Course Code	Category	Hours/Week			Credits	Maximum Marks		
HCS03	H	L	T	P	C	Continuous Internal Assessment	End Exam	TOTAL
		3		2	4	40	60	100
Sessional Exam Duration : 1 ½ Hrs					End Exam Duration: 3 Hrs			
Course Outcomes : At the end of the course the student will be able to								
CO1: Describe updated social media landscape, key figures and the methods connecting the popular social networks.								
CO2: Describe Facebook API ecosystem, data extraction methodology, text mining techniques : stopword removal, stemming using NLTK and device detection customized cleaning								
CO3: Demonstrate Youtube campaigns and channel measurement using the combination of traffic and sentiment data and discovery of users interesting projects and technologies on Github.								
CO4: Extract conversational topics from Internet Forums and users interests from Pinterest through network analysis.								
CO5: Write and execute Python ready to run scripts and Spark jobs on AWS								
UNIT – I								
Introduction to the Latest Social Media Landscape and Importance: Introducing social graph, Delving into social data, Understanding the process, Working environment, Getting the data, Analyzing the data, Visualizing the data.								
Harnessing Social Data - Connecting, Capturing, and Cleaning: APIs in a nutshell, Introduction to authentication techniques, Parsing API outputs. Basic cleaning techniques, MongoDB to store and access social data								
UNIT – II								
Uncovering Brand Activity, Popularity, and Emotions on Facebook: Facebook brand page, Project planning, Analysis, Keywords, Noun phrases, Detecting trends in time series, Uncovering emotions.								
Analysing Twitter Using Sentiment Analysis and Entity Recognition: Getting the data, Sentiment analysis: Customized sentiment analysis, Named entity recognition, Combining NER and sentiment analysis.								
UNIT – III								
Campaigns and Consumer Reaction Analytics on YouTube – Structured and Unstructured: Getting the data, Data pull, Data processing, Data analysis								
The Next Great Technology – Trends Mining on GitHub: Getting the data, Data pull, Data processing, Textual data, Numerical data, Data analysis, Comparison of technologies in terms of forks, open issues, size, and watchers count								
UNIT – IV								

Scraping and Extracting Conversational Topics on Internet Forums: Getting the data, Data pull and pre-processing, Data analysis.
Demystifying Pinterest through Network Analysis of Users Interests: Getting the data, Data pull and pre-processing, Data analysis
UNIT – V
Social Data Analytics at Scale – Spark and Amazon Web Services: Different scaling methods and platforms, Topic models at scale, Spark on the Cloud – Amazon Elastic MapReduce
Text Books:
1. Python Social Media Analytics, Siddhartha Chatterjee, Michal Krystyanczuk, Packt Publishing Ltd, 2017
Reference Books:
1. Mastering Social Media Mining with Python, Marco Bonzanini, Packt Publishing Ltd, 2016
2. Thoughtful Data Science, David Taieb, Packt Publishing, June 2018.
Question Paper Pattern:
Sessional Exam: The question paper for sessional examination shall be for 25 marks, covering half of the syllabus for first sessional and remaining half for second sessional exam. The question paper shall consist of three sections with Two Questions (EITHER/ OR Type) in each section. The student shall answer one question from each section.
End Exam: The question paper for End examination shall be for 60 marks. The Question paper shall contain Five Units with Two Questions (Either or Type) from each unit. Each of these questions may contain sub question and the student should answer any one question from each unit. Each Question carries 12 marks.

<i>List of Experiments</i>
1. Extract Data from Twitter Using Python and Twitter APIs
2. Extract Data from Face book Using Python and Face book APIs
3. Analyze Twitter Data using Sentiment Analysis and Entity Recognition.
4. Analyze the YouTube data by ‘Sentiment by weekday’, ‘Comments by weekday’ after pre-processing
5. Extract Data from Twitter Using Python and Twitter APIs

NATURAL LANGUAGE PROCESSING (NLP)

VII Semester : Honors CSE/CST/CSBS						Scheme: 2020		
Course Code	Course Category	Hours/Week			Credits	Maximum Marks		
HCS04	H	L	T	P	C	Continuous Internal Assessment	End Exam	Total
		3	0	2	4	40	60	100
Sessional Exam Duration: 1 ½ Hrs					End Exam Duration: 3 Hrs			
Course Outcomes: At the end of the course students will be able to CO1: Understand the pre processing techniques of natural language processing CO2: Understand models for building NLP Applications CO3: Use Classification and Clustering Concepts on Textual data CO4: Apply text tokenization, feature extraction on WordNet database CO5: Relate other topics related to Natural Language Processing								
UNIT- I								
Introduction to Natural Language Processing: Why Learn NLP, Driving into NLTK, Text Wrangling and Cleaning, Tokenization ,Stemming ,Lemmatization ,Stop word removal ,Rare word removal, Spell correction.POS Tagging and Named Entity Recognition								
UNIT- II								
NLP Application: Building your first NLP Application, Other NLP Applications-Machine Translation, Vector Space Model, Probabilistic Model.								
UNIT- III								
Text Classification: Machine Learning , Sampling, Naïve Bayes, Decision Trees, Stochastic gradient Descent, Logistic Regression,SVM, Random forest algorithm ,Text clustering-K Means.								
UNIT- IV								
Tokenizing Text and WordNet Basics: Introduction, Tokenizing text into sentences, Tokenizing sentences into words, Tokenizing sentences using regular expressions, Training a sentence Tokenizer, Filtering stop words in a tokenized sentence, Looking up Synsets for a word in WordNet , Looking up lemmas and synonyms in WordNet.								
Feature Extraction: BOW, TF-IDF Feature Extraction. Training a Naive Bayes classifier, Decision tree classifier, Entropy classifier.								
UNIT- V								
Topics in Brief: Search and Information Retrieval, Topic Modeling, Text Summarization, Machine Translation, Question-Answering Systems								
TextBooks :								
1. Natural Language Processing: Python and NLTK, Deepti Chopra, Jacob Perkins, and Nitin Hardeniya by Packt 2016.								
2. Practical Natural Language Processing: A Comprehensive Guide to Building Real-World NLP Systems, Bodhisattwa Majumder, Anuj Gupta, Sowmya Vajjala, Harshit Surana,Published by O'Reily Media , Inc, 2020.								
Question Paper Pattern:								

Sessional Exam:

The question paper for sessional examination shall be for 25 marks, covering half of the syllabus for first sessional and remaining half for second sessional exam. The question paper shall consist of three sections with Two Questions (EITHER/ OR Type) in each section. The student shall answer one question from each section.

End Exam:

The question paper for End examination shall be for 60 marks. The Question paper shall contain Five Units with Two Questions (Either or Type) from each unit. Each of these questions may contain sub question and the student should answer any one question from each unit. Each Question carries 12 marks.

<i>List of Experiments</i>
1. Perform Text pre processing using NLP techniques
2. Program using POS Tagging and Named Entity Recognition
3. Program using Machine Translation.
4. Perform text classification using Machine Learning classifiers
5. Apply Feature Extraction techniques on textual dataset