

Scheme – 2020

Department of Computer Science & Business Systems

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 & BUSINESS SYSTEMS

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

Honors in Computer Science & Business Systems (CSBS) Scheme of instruction and examination

(Effective from 2020-2021)

S	Semester	Course Title	Credits	Scheme of instruction (periods/ week)		Scheme of Examination			
110.				L	Р	End Exam	Internal Assessment	Total 100	
1	IV	Introduction to Cyber Security	4	4	0	60	40	100	
2	V	Python for Data Science	4	3	2	60	40	100	
3	VI	Big 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	Introduction to Block Chain Technologies
3	Reinforcement Learning
4	Embedded Systems
	MOOCS – 2 / Mini Project
1	Computer Vision
2	Virtual Reality
3	Database security
4	Fuzzy Logic & Neural Networks

INTRODUCTION TO CYBER SECURITY (ICS)

IV Semester: Honors					Scheme: 2017			
Course Code	Category	Hou	rs/We	eek	Credits	Credits Maximum Mar		
HCB01	Н	L	Т	Р	С	Continuous Internal Assessment	End Exam	TOTAL
		3	0	0	3	40	60	100
Sessional Ex	xam Duration: 1	½ H	rs			End E	xam Duration	:3 Hrs
Course Outc	comes: At the en	d of t	he cou	rse st	udents will	l be able to		
CO1: Cyber	Security archited	ture p	rincip	les				
CO2: Identit	fying different cla	asses	of atta	cks				
CO3: Under	stand about cybe	rcrim	e with	mobil	e and wirel	ess devices		
CO4: Under	stand about the to	ools a	nd met	thods	used in cyb	ercrime.		
CO5: Under	stand about cybe	r secu	rity an	d soci	ial media m	arketing.		
	· · · · ·							
				UN	IT– I			
Introduction to	o Cybercrime							
Introduction, C	ybercrime: Defir	nition	and O	rigins	of the Wor	rd, Cybercrime	and Informatic	n Security,
Who are Cyb	ercriminals? Cla	ssific	ations	of C	ybercrimes	, Cybercrime:	The Legal Pe	erspectives,
Cybercrimes: A	An Indian Perspe	ctive,	Cyber	crime	and the In	dian ITA 2000	, A Global Per	spective on
Cybercrimes, C	Cybercrime Era: S	Surviv	al Ma	ntra fo	or the Netizo	ens.		
				UNI	T– II			
Cyber offenses								
How Criminals	s Plan Them Intro	oducti	on, Ho	ow Cr	iminals Pla	n the Attacks, S	Social Engineer	ring, Cyber
stalking, Cybe	r cafe and Cybe	ercrim	les, B	otnets	: The Fuel	for Cybercrin	ne, Attack Ve	ctor Cloud
Computing.								
				UNI	T– III			
Cybercrime M	obile and Wireles	ss Der	vices					

Introduction, Proliferation of Mobile and Wireless Devices, Trends in Mobility, Credit Card Frauds in Mobile and Wireless Computing Era, Security Challenges Posed by Mobile Devices, Registry Settings for Mobile Devices, Authentication Service Security, Attacks on Mobile/Cell Phones. Mobile Devices: Security Implications for Organizations, Organizational Measures for Handling Mobile, Organizational Security Policies and Measures in Mobile Computing Era, Laptops.

UNIT-IV

Tools and Methods Used in Cybercrime

Introduction, Proxy Servers and Anonymizers, Phishing, Password Cracking, Key loggers and Spywares, Virus and Worms, Trojan Horses and Backdoors, Steganography, DoS and DDoS Attacks, SQL Injection, Buffer Overflow, Attacks on Wireless Networks, Phishing.

UNIT_V

Cyber Security:

Organizational Implications Introduction, Cost of Cybercrimes and IPR issues, Web threats for Organizations, Security and Privacy Implications, Social media marketing: Security Risks and Perils for Organizations, Social Computing and the associated challenges for Organizations.

Text Books:

- 1. Cyber Security: Understanding Cyber Crimes, Computer Forensics and Legal Perspectives, Nina Godbole, SunitBelapure, Wiley.
- 2. Principles of Information Security, MichealE.Whitman and Herbert J.Mattord, Cengage Learning.

Reference Books:

- 1. Information Security, Mark Rhodes, Ousley, MGH.
- 2. Cyber Security Essentials, James Graham, Richard Howard and Ryan Otson, CRC Press

Web References:

1.https://www.tutorialspoint.com/fundamentals_of_science_and_technology/cyber_crime_and_cyber_security.htm

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.

PYTHON FOR DATA SCIENCE (PDS)

Course Code Category Code 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 CO1: Write python programs using the core concepts like Lists, Dictionaries, sets, tuple, functions and regular expressions. CO4: 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: Create static and interactive visualizations using Matplotlib. CO5: Enumerate machine learning algorithms, Describe the Classification and Clustering UNIT – 1 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 – 11 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 – II </th <th colspan="10">V Semester : Honors CSE/CST/CSBS Scheme : 2020</th>	V Semester : Honors CSE/CST/CSBS Scheme : 2020									
HCS02 H L T P C Continuous Internal Assessment Asseston Asseston Assessment Assessment Assessment Asseston Assessment	Course Code	Category	Ho	ours/W	eek	Credits	Max	imum Mar	ks	
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: Create 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 VIII – IV Visualization with Matplotlib: Two Interfaces for the Price of One, Simple Lin	HCS02	Н	L	Т	Р	С	Continuous Internal Assessment	End Exam	TOTAL	
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 COI: 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: Create 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, Histogra			3		2	4	40	60	100	
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: Create 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: Korge 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 Annotaion, 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 Sickit-Learn, Feature Engineering, Model development, Linear Regression: Simple Linear Regression, Example of model development.	Sessional Exam Duration : 1 ½ Hrs End Exam Duration: 3 H									
CO2: Demonstrate various mathematical operations on arrays using NumPy CO3: Analyze and manipulate Data using Pandas CO4: Create 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 Engincering, Model development, Linear Regression: Simple Linear Regression, Example of model development.	Course Out	Course Outcomes : At the end of the course the student will be able to								
regurate expressions. CO2: Demonstrate various mathematical operations on arrays using NumPy CO3: Analyze and manipulate Data using Pandas CO4: Create static and interactive visualizations using Matplotlib. CO5: Enumerate machine learning algorithms, Describe the Classification and Clustering UNIT – 1 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 – II 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	COI: Write	e python programs	s using	the co	re conce	pts like Lis	ts, Dictionaries, se	ets, tuple, fu	nctions and	
 CO3: Definition and evaluation of the analysis of a flag standing stand	regular expl	essions.	athom	ation	noration		using NumDu			
Code: Create static and interactive visualizations using Matplotlib. 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, Custom	CO2: Demo	onstrate various n	ta Data		Dender	is on arrays	using NumPy			
Code: Create static and interactive visualizations using Maphotno. Code: Create static and interactive visualizations using Maphotno. 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	CO3: Analy	yze and manipula		invelie	Pandas	nin a Matula	41:1.			
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.	CO4: Creat	e static and intera		1sualiz	ations u	$\frac{1}{2}$	ound.	Clustoning		
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.		ierate machine lea	arning a	argoriu	lins, De	scribe the C		Clustering		
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.					UN	IT – I				
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.	Introduction Dictionaries Objects, Re	on to Python: I s,Operators,Condi gular expressions	Data T _i itional	ypes:S ⁻ Statem	trings, 1 ents, Lo	Numbers, 1 oops, Funct	Booleans, Date a tions, Modules ar	nd Time, 1 nd packages	Lists, Tuples, , Classes and	
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.					UN	II – II				
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.	Introduction Functions, A Comparison Structured A	on to NumPy: T Aggregations: Mi Is, Masks, and B Arrays	'he Bas n, Max oolean	ics of , and E Logic	NumPy Everythi , Fancy	Arrays, C ng in Betwo Indexing, S	Computation on Neen, Computation Sorting Arrays, S	JumPy Arra on Arrays: tructured D	ys: Universal Broadcasting, ata: NumPy's	
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 PandasUNIT – IVVisualization 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 – VMachine Learning: Scikit-Learn, Feature Engineering, Model development, Linear Regression: Simple Linear Regression, Example of model development.					UN	IT – III				
UNIT – IVVisualization 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 – VMachine 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.	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									
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.		UNIT – IV								
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.	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									
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.	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, Paul Deitel, Harvey Deitel, 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.

	List of Experiments
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.

BIG DATA ANALYTICS (BDA)

VI Semeste	r : Honors						S	cheme : 2020	
Course Code	Category	Hours/Week			Credits	Max	imum Mar	ks	
НСВ02	Н	L	Т	Р	С	Continuous Internal Assessment	End Exam	TOTAL	
		3	-	2	4	40	60	100	
Sessional E	xam Duration :	1½ H	rs			En	d Exam Du	ration: 3 Hrs	
Course Out	Course Outcomes : At the end of the course the student will be able to								
CO1: Unde	rstand the basics	of Big	Data A	nalytics	, Hadoop.				
CO2: Desig	n Map Reduce p	rogram	s for a	given p	roblem.				
CO3: Write	Pig Scripts on H	adoop	that wo	orks on	arge datase	ts.			
CO4: Perfo	rm Data Querying	g Opera	ations i	using Aj	pache Hive.				
CO5: Imple	ement Data Mana	gement	using	NoSQL	Databases				
				UN	I – TI				
Big Data A What is Big Analytics, T Introduction Introducing and Applica	Big Data Analytics: What is Big Data Analytics, Why this Sudden Hype Around Big Data Analytics? Classification of Analytics, Top Challenges Facing Big Data, Few Top Analytics Tools, Introduction to Hadoop: Introducing Hadoop: HDFS, HDFS Commands, Processing Data with Hadoop, Managing Resources and Applications with Hadoop VAPN. Internating with Hadoop EaoSystem								
				UN	IT – II				
Understand The Map Ro While Desig	ling Map Reduc educe Frameworl gning Map Reduc	e & YA c Conce e, YAF	ARN: ept, De RN Bac	evelopin ekgroun	g Simple M d and YARI	Iap Reduce Appli N Architecture.	cation, Poir	ts to consider	
				UN	IT – III				
Analyzing I Introducing	Data with Pig: PIG, Running PI	G, Gett	ing sta	rted wit	h pig Latin,	Working with op	erators in pi	lg,	
				UN	IT – IV				
Understand Introducing	ling HIVE: Hive,Hive Servic	es, Bu	ilt in fu	inctions	in Hive, Hi	ve DDL, Data Ma	anipulation i	n Hive.	
UNIT – V									
NoSQL Data Management: Introduction to NoSQL, Characteristics of NoSQL, Types of NoSQL Data Models, Schema-less Databases.									
Text Books:									
1. Big Data Black Book: Covers Hadoop 2, Map Reduce, Hive, YARN, Pig, R and Data Visualization by DreamTech, 2015.									

2. Big Data and Analytics by Seema Acharya, Wiley Publication, 2015..

Reference Books:

1. Data Science & Big Data Analytics: Discovering, Analyzing, Presenting Data Visualizing.

- 2. Hadoop: The Definitive Guide, 3rd Edition, By Tom White, O'reilly Media.
- 3. Big Data Now: 2012 Edition Publisher: O'Reilly Media.
- 4. Too Big to Ignore: The Business Case for Big Data (Wiley and SAS Business Series) By Phil Simon, Wiley 1e.

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.

	List of Experiments							
1.	Perform Hadoop Setup in Local and Pseudo mode and monitor through Web Based UI.							
2.	Implementation of Hadoop Shell Commands on files.							
3.	Implementation of word count Example using Hadoop Map Reduce.							
4.	Write a Map Reduce Program that works on Gutenberg data.							
5.	Write Pig Latin Scripts on Describe, for each and order by operator							
6.	Perform DDL Operations on Hive							

NATURAL LANGUAGE PROCESSING (NLP)

VII Semester :	Honors CSE/C	ST/C	SBS				Sc	heme: 2020		
Course Code	Course Category	Hou	rs/We	ek	Credits	Ma	aximum Mar	[•] ks		
HCS04	Н	L	T	P	C	Continuous Internal Assessment	End Exam	Total		
Sossional Exam	Duration: 1 1/	J Uma	0	2	4	40 Fn	d Exam Dur	ation: 3 Hrs		
Sessional Exam Duration: 1 72 Hrs End Exam Duration: 5 Hrs										
Course Outc	omes: At the end	ofthe	- cours	e stud	ents will be	able to				
CO1: Unders	stand the pre proc	essing	techni	aues of	f natural lang	guage processin	g			
CO2: Under	stand models for	buildi	ng NLI	P Appl	ications		0			
CO3: Use C	lassification and	Clust	ering C	Concer	ots on Textu	al data				
CO4: Apply	text tokenization	, featu	ire extr	raction	on WordNe	et database				
CO5: Relate	other topics relat	ed to	Natura	l Lang	uage Proces	ssing				
	*				<u> </u>	<u> </u>				
				UN	IT–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										
classifier, Entrop	y classifier.					-8	,,			
				UNI	T– V					
Topics in Brief: Translation, Que	Search and Inforestion-Answering	matio Syste	n Retri ms	ieval, '	Topic Mode	ling, Text Sum	imarization, N	Machine		
TextBooks :										
1. Natural La Nitin Hard	nguage Processin leniya by Packt 2	g: Pyt 016.	hon an	d NL7	TK, Deepti (Chopra, Jacob I	Perkins, and			
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.

List of Experiments

- 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