

FACULTY OF ENGINEERING AND TECHNOLOGY

**SYLLABUS FOR THE BATCH FROM
YEAR 2020 TO YEAR 2023**

For

**BACHELOR OF VOCATION (B.VOC.)
(DATA SCIENCE)
(Semester: I – VI)**

EXAMINATION 2020-23



**GURU NANAK DEV UNIVERSITY
AMRITSAR**

**Note: (i) Copy rights are reserved.
Nobody is allowed to print it in any form.
Defaulters will be prosecuted.**

**(ii) Subject to change in the syllabi at any time.
Please visit the University website time to time.**

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 23

SCHEME

SEMESTER – I:

Paper No.	Paper	M. Marks
Paper – I	Descriptive Statistics (Theory)	75
Paper – II	Database Management System (Theory)	75
Paper – III	Data Entry using MS-Word and MS-Excel (Practical)	75
Paper – IV	MS-Access (Practical)	75
Paper – V	Communication Skills in English – I	50
Paper – VI	Punjabi (Compulsory) / ** ਮੁੱਢਲੀ ਪੰਜਾਬੀ / ** Punjab History & Culture (From Earliest Times to C 320)	50

SEMESTER – II:

Paper No.	Paper	M. Marks
Paper – I	Introduction to Data Science (Theory)	75
Paper – II	Basic Mathematics (Theory)	75
Paper – III	Introduction to R (Theory)	75
Paper – IV	Practical Based on Programming in R (Practical)	75
Paper – V	Communication Skills in English – II (Th.35+Pr.15)	50
Paper – VI	Punjabi (Compulsory) / ** ਮੁੱਢਲੀ ਪੰਜਾਬੀ / ** Punjab History & Culture (C 320 TO 1000 B.C.)	50
Paper – VII	* Drug Abuse: Problem, Management and Prevention (Compulsory Paper)	100

Note: * Marks of this Paper will not be included in the Total Marks.

**** (Special Paper in lieu of Punjabi Compulsory)**

(For those students who are not domicile of Punjab)

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 23

SEMESTER – III

	Courses		Hours	Marks
Paper-I	Optimization	Theory	3	50
Paper-II	Business Economics	Theory	3	50
Paper-III	Statistical Inference-I	Theory	3	75
Paper-IV	Data Mining	Theory	3	75
Paper-V	Practical based on SAS	Practical	3	75
Paper-VI	Introduction to Python	Practical	3	75

SEMESTER – IV

	Courses		Hours	Marks
Paper-I	Basics of Linear Algebra and Numerical Analysis	Theory	3	50
Paper-II	Statistical Inference-II	Theory	3	50
Paper-III	Algorithms and Heuristics	Theory	3	75
Paper-IV	Big Data	Theory	3	75
Paper-V	Big Data Analytics using R	Practical	3	75
Paper-VI	Programming Lab based on Numerical Analysis	Practical	3	75
Paper-VII (ESL-221)	* Environmental Studies			100

*** Marks of Paper EVS will not be included in Grand Total.**

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 23

SEMESTER-V

S. No.	Subjects	Marks
	General Education Component:	
1	Regression (Theory)	75
2	Design of Experiments (Theory)	75
	Skill Component	
3	Practical Based on Predictive Analysis using R	75
4	Soft Computing (Theory)	75
5	Machine Learning	100
	Total	400

SEMESTER-VI

S. No.	Subjects	Marks
	General Education Component	
1	Supply Chain Management (Theory)	75
2	Queuing Theory (Theory)	75
	Skill Component:	
3	Operational Research models (Theory)	75
4	Artificial Intelligence (Theory)	75
5	Industrial Training	100
	Total	400

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – I

Paper–I: Descriptive Statistics

Time: 3 Hours

Max. Marks: 75

Instructions for the Paper Setters:-

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION-A

Definition, Scope, Significance, Limitations. Tabulation, Classification and Graphical representation of data (Pie Chart, Bar Diagram, Histogram, Frequency Polygon, Ogive Curve, etc.).

SECTION-B

Measures of Central Tendency – Arithmetic Mean, Median and Mode, Position of averages. Graphical representation of data, Measures of dispersion – range, variance, mean deviation, standard deviation and coeff. of variation, Concepts and Measures of Skewness and Kurtosis.

SECTION-C

Mathematical and Statistical probability, Elementary events, Sample space, Compound events, Types of events, Mutually exclusive, Independent events, addition law of probability, Conditional probability, Multiplication theorem of probability, Baye's Theorem.

SECTION-D

Concept of Random Variable, Probability Mass Function & Density Function, Mathematical Expectation (meaning and properties), Moments, Moment Generating Function and Characteristic Function.

Text/References:

1. Gupta, S.P.: Statistical Methods (1981).
2. Croxton, Cowden & Klein: Applied General Statistics (1973).
3. Kapur and Sexena: Mathematical Statistics (1970)
4. Murry, R. Speigal: Theory and Problems of Statistics (1972)

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – I

Paper–II: Database Management System

Time: 3 Hours

Max. Marks: 75

Instructions for the Paper Setters:-

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION-A

Basic Concepts: A Historical perspective, File Systems vs. DBMS, Characteristics of the Data Base Approach, Abstraction and Data Integration, Database users, Advantages and Disadvantages of DBMS, Implication of Database approach.

SECTION-B

Data Base Systems Concepts and Architecture: Data Models, Schemas and Instances, DBMS architecture and Data Independence, Data base languages & Interfaces, DBMS functions and component modules

SECTION-C

Entity Relationship Model: Entity Types, Entity Sets, Attributes & Keys, Relationships, Relationship Types, Roles and Structural Constraints, Design issues, weak entity types, E-R Diagrams. Design of an E-R Database Schema, Reduction of an E-R Schema to Tables.

Conventional Data Models: An overview of Network and Hierarchical Data Models. Relational Data Model: Relational model concepts, Integrity constraints over Relations, Relational Algebra - Basic Operations.

SECTION-D

Relational Data Base Design: Functional Dependencies, Decomposition, Desirable properties of decomposition, Normal forms based on primary keys (1 NF, 2 NF, 3 NF and BC NF). RDBMS: Terminology, The 12 Rules (Codd's Rule) for an RDBMS. Introduction to Data Mining, Its Applications. Concept of Data ware house, Its Architecture, Introduction to Big Data.

Text/References:

1. C.J. Date, "An Introduction of Database System", The Systems Programming Series, 6/Ed, Addison-Wesley Publishing Company, Inc., 1995.
2. Silberschatz, Korth and Sudarshan, "Database System Concepts", Third Ed. McGraw Hill International Editions, Computer Science Series-1997.

6
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – I

Paper–III: Data Entry Using MS-Word & MS-Excel
(Practical)

Time: 3 Hours

Max. Marks: 75

Note: Practical exam to be conducted by the external examiner.

SECTION-A

MS-Word: Overview, Creating, Saving, Opening, Importing, Exporting & Inserting files. Formatting pages, paragraphs and sections. Indents and outdates. Creating lists and numbering. Heading Styles, Fonts and size editing,

SECTION-B

Using editing and proofing tools, changing layout of a document, positioning & viewing text. Finding & replacing text, inserting page breaks, page numbers, book marks, symbols & dates. Using tabs and tables Header, Footer & Printings. Mail merge

SECTION-C

MS-Excel: Worksheet overview. Entering information in Worksheet. Opening and saving workbook. Formatting number and texts, Protecting cells. Producing Charges and printing operations graphs,

SECTION-D

Creating Different Formulas, 3D formulas, Copying and pasting formulas, conditional formatting and cell styles, creating worksheet charts, sharing workbook, tables, sorting data, filtering data, using what-if analysis, table related functions and making macros.

Text/References:

1. Peter Norton, “Introduction to Computers”, McGraw-Hill, New Delhi.
2. Sanjay Sexana, “A First Course in Computers”, Vikas Publishing House, New Delhi.
3. Rajaraman, V., “Fundamental of Computers”, Prentice Hal India, New Delhi.
4. Srivastava, S.S., “MS-Office” Firewall Media, New Delhi.
5. Alexis Loeon and Matheus Leon, “Introduction to Computers with MS-Office 200”, Tata McGraw-Hill, New Delhi.

Note: The Latest Editions of the books should be followed

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – I

Paper–IV: MS-Access
(Practical)

Time: 3 Hours

Max. Marks: 75

Note: Practical exam to be conducted by the external examiner.

Practical based on MS-Access
SECTION-A

1. Basics of RDBMS

- Introduction to database -What is a Database
- Why use a Relational Database
- Overview of Database Design
- Integrity Rules (Primary/Foreign Key, One-to-Many, Many-to-Many, One-to-One)
- Introduction to MS Access (Objects, Navigation).

2. Working with Table:

- Create a table in MS Access
- Data Types
- Field Properties
- Validation Rules Data Entry
- Add Record Delete Record And Edit Text
- Sort Option
- Find/Replace
- What Is Filter
- Rearrange Columns
- Freeze Columns
- Edit A Tables- Copy, Delete, Import
- Modify Table Structure

SECTION-B

3. Working with Query:

- Introduction of relationship
- How to create a relationship
- Types of relationship
- How to create a relationship
- Set a rule for Referential Integrity

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – I

- Change the join type
- Delete a relationship
- Save relationship
- Queries & filter –difference between queries and filter
- What is query
- Filter using multiple fields and, or
- Advance filter queries
- Create query with one table
- select query
- Find duplicate record with query
- Find unmatched record with query
- Run query
- Save and change query

SECTION-C

4. Working with Forms:

- Introduction to forms
- Types of basic forms: columnar, tabular, datasheet, main/subforms
- Add headers and footers
- Add fields to form
- Tool box
 - Add text to form
 - Use label
 - Use option button
 - Use check box
 - Use combo box
 - Use list box
- Create Form by using Wizard

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – I

SECTION-D

5. Working with Reports:

- Introduction to reports
- Types Of Basic Reports: Single Column, Tabular Report Groups
- Single And Multi Table Report
- Preview And Print Report
- Creating Reports And Labels, Wizard

Text/References:

1. Access 2007 for Starters : The Missing Manual by Mathew Macdonald
2. Access 2013: Bible by Michael Alexander

10
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – I

PAPER–V: COMMUNICATION SKILLS IN ENGLISH–I

Time: 3 Hours

Max. Marks: 50

Instructions for the Paper Setters: -

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

The syllabus is divided in four sections as mentioned below:

SECTION–A

Reading Skills: Reading Tactics and strategies; Reading purposes–kinds of purposes and associated comprehension; Reading for direct meanings.

SECTION–B

Reading for understanding concepts, details, coherence, logical progression and meanings of phrases/ expressions.

Activities:

- Comprehension questions in multiple choice format
- Short comprehension questions based on content and development of ideas

SECTION–C

Writing Skills: Guidelines for effective writing; writing styles for application, personal letter, official/ business letter.

Activities:

- Formatting personal and business letters.
- Organising the details in a sequential order

SECTION–D

Resume, memo, notices etc.; outline and revision.

Activities:

- Converting a biographical note into a sequenced resume or vice-versa
- Ordering and sub-dividing the contents while making notes.
- Writing notices for circulation/ boards

Recommended Books:

- *Oxford Guide to Effective Writing and Speaking* by John Seely.
- *English Grammar in Use* (Fourth Edition) by Raymond Murphy, CUP

11
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – I

PAPER-VI: ਪੰਜਾਬੀ (ਲਾਜ਼ਮੀ)

ਸਮਾਂ : 3 ਘੰਟੇ

ਕੁਲ ਅੰਕ : 50

ਅੰਕ-ਵੰਡ ਅਤੇ ਪਰੀਖਿਅਕ ਲਈ ਹਦਾਇਤਾਂ

1. ਪ੍ਰਸ਼ਨ ਪੱਤਰ ਦੇ ਚਾਰ ਭਾਗ ਹੋਣਗੇ। ਹਰ ਭਾਗ ਵਿਚ ਦੋ ਪ੍ਰਸ਼ਨ ਪੁੱਛੇ ਜਾਣਗੇ।
2. ਵਿਦਿਆਰਥੀ ਨੇ ਕੁੱਲ ਪੰਜ ਪ੍ਰਸ਼ਨ ਕਰਨੇ ਹਨ। ਹਰ ਭਾਗ ਵਿਚੋਂ ਇਕ ਪ੍ਰਸ਼ਨ ਲਾਜ਼ਮੀ ਹੈ। ਪੰਜਵਾਂ ਪ੍ਰਸ਼ਨ ਕਿਸੇ ਵੀ ਭਾਗ ਵਿਚੋਂ ਕੀਤਾ ਜਾ ਸਕਦਾ ਹੈ।
3. ਹਰੇਕ ਪ੍ਰਸ਼ਨ ਦੇ ਬਰਾਬਰ ਅੰਕ ਹਨ।
4. ਪੇਪਰ ਸੈੱਟ ਕਰਨ ਵਾਲਾ ਜੇਕਰ ਚਾਹੇ ਤਾਂ ਪ੍ਰਸ਼ਨਾਂ ਦੀ ਵੰਡ ਅੱਗੋਂ ਵੱਧ ਤੋਂ ਵੱਧ ਚਾਰ ਉਪ-ਪ੍ਰਸ਼ਨਾਂ ਵਿੱਚ ਕਰ ਸਕਦਾ ਹੈ।

ਪਾਠ-ਕ੍ਰਮ ਅਤੇ ਪਾਠ-ਪੁਸਤਕਾਂ

ਸ਼ੈਕਸਨ-ਏ

ਸਰਵੋਤਮ ਪੰਜਾਬੀ ਸਾਹਿਤ (ਸਿੱਖ ਡਾ. ਰਮਿੰਦਰ ਕੌਰ, ਡਾ. ਮੇਘਾ ਸਲਵਾਨ)- (ਕਵਿਤਾ ਅਤੇ ਕਹਾਣੀ ਭਾਗ)

ਸ਼ੈਕਸਨ-ਬੀ

ਇਤਿਹਾਸਕ ਯਾਦਾਂ

(ਜੀਵਨੀ 1 ਤੋਂ 6 ਤਕ)

ਵਿਸ਼ਾ ਵਸਤੂ/ਸਾਰ/ਨਾਇਕ ਬਿੰਬ

ਸ਼ੈਕਸਨ-ਸੀ

(ੳ) ਪੈਰ੍ਹਾ ਰਚਨਾ

(ਅ) ਪੈਰ੍ਹਾ ਪੜ੍ਹ ਕੇ ਪ੍ਰਸ਼ਨਾਂ ਦੇ ਉਤਰ

ਸ਼ੈਕਸਨ-ਡੀ

ਭਾਸ਼ਾ ਵੰਨਗੀਆਂ :

ਭਾਸ਼ਾ ਦਾ ਟਕਸਾਲੀ ਰੂਪ, ਭਾਸ਼ਾ ਅਤੇ ਉਪਭਾਸ਼ਾ ਵਿਚਲਾ ਅੰਤਰ,

ਪੰਜਾਬੀ ਉਪਭਾਸ਼ਾਵਾਂ ਦੇ ਪਛਾਣ ਚਿੰਨ੍ਹ, ਪੰਜਾਬੀ ਭਾਸ਼ਾ - ਨਿਕਾਸ ਤੇ ਵਿਕਾਸ

ਸਹਾਇਕ ਪੁਸਤਕਾਂ

1. ਰਾਮਿੰਦਰਪਾਲ ਸਿੰਘ ਬਰਾੜ, ਪੰਜਾਬੀ ਕਵਿਤਾ ਦਾ ਇਤਿਹਾਸ, ਪੰਜਾਬੀ ਅਕਾਦਮੀ, ਦਿਲੀ।
2. ਬ੍ਰਹਮਜਗਦੀਸ਼ ਸਿੰਘ, ਆਧੁਨਿਕ ਪੰਜਾਬੀ ਕਾਵਿ ਸਿਧਾਂਤ, ਇਤਿਹਾਸ ਅਤੇ ਪ੍ਰਵਿਰਤੀਆਂ, ਵਾਰਿਸ ਸ਼ਾਹ ਫਾਊਂਡੇਸ਼ਨ, ਅੰਮ੍ਰਿਤਸਰ।
3. ਬਲਦੇਵ ਸਿੰਘ ਧਾਲੀਵਾਲ, ਪੰਜਾਬੀ ਕਹਾਣੀ ਦਾ ਇਤਿਹਾਸ, ਪੰਜਾਬੀ ਅਕਾਦਮੀ, ਦਿਲੀ।
4. ਸਤਿੰਦਰ ਸਿੰਘ, ਆਧੁਨਿਕ ਪੰਜਾਬੀ ਵਾਰਤਕ ਦਾ ਇਤਿਹਾਸ, ਪੰਜਾਬੀ ਅਕਾਦਮੀ, ਦਿਲੀ।
5. ਡਾ. ਰਮਿੰਦਰ ਕੌਰ, ਪੰਜਾਬੀ ਕਹਾਣੀ ਦਾ ਸਫ਼ਰ ਤੇ ਸ਼ਾਸਤ੍ਰ ਭਾਗ-1, ਸਿੰਘ ਬ੍ਰਦਰਜ਼, ਅੰਮ੍ਰਿਤਸਰ।
6. ਹਰਕੀਰਤ ਸਿੰਘ, ਭਾਸ਼ਾ ਤੇ ਭਾਸ਼ਾ ਵਿਗਿਆਨ, ਲਾਹੌਰ ਬੁਕ ਸ਼ਾਪ, ਲੁਧਿਆਣਾ।
7. ਹਰਕੀਰਤ ਸਿੰਘ ਤੇ ਗਿਆਨੀ ਲਾਲ ਸਿੰਘ, ਕਾਲਜ ਪੰਜਾਬੀ ਵਿਆਕਰਣ, ਪੰਜਾਬ ਯੂਨੀਵਰਸਿਟੀ, ਚੰਡੀਗੜ੍ਹ।
8. ਬੂਟਾ ਸਿੰਘ ਬਰਾੜ, ਪੰਜਾਬੀ ਵਿਆਕਰਨ : ਸਿਧਾਂਤ ਤੇ ਵਿਹਾਰ, ਚੇਤਨਾ ਪ੍ਰਕਾਸ਼ਨ, ਲੁਧਿਆਣਾ।
9. ਮਿੰਨੀ ਸਲਵਾਨ, ਪੰਜਾਬੀ ਵਿਆਕਰਨ : ਮੁਢਲੇ ਸੰਕਲਪ, ਰਵੀ ਸਾਹਿਤ ਪ੍ਰਕਾਸ਼ਨ, ਅੰਮ੍ਰਿਤਸਰ
10. ਪੰਜਾਬੀ ਭਾਸ਼ਾ ਬੋਧ, ਕਸਤੂਰੀ ਲਾਲ ਐਂਡ ਸੰਨਜ਼, ਅੰਮ੍ਰਿਤਸਰ।

12
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – I

PAPER–VI: ਮੁੱਢਲੀ ਪੰਜਾਬੀ
((In lieu of Compulsory Punjabi))
(For those students who are not domicile of Punjab)

ਸਮਾਂ : 3 ਘੰਟੇ

ਕੁਲ ਅੰਕ: 50

ਅੰਕ-ਵੰਡ ਅਤੇ ਪਰੀਖਿਅਕ ਲਈ ਹਦਾਇਤਾਂ

1. ਪ੍ਰਸ਼ਨ ਪੱਤਰ ਦੇ ਚਾਰ ਭਾਗ ਹੋਣਗੇ। ਹਰ ਭਾਗ ਵਿਚ ਦੋ ਪ੍ਰਸ਼ਨ ਪੁੱਛੇ ਜਾਣਗੇ।
2. ਵਿਦਿਆਰਥੀ ਨੇ ਕੁੱਲ ਪੰਜ ਪ੍ਰਸ਼ਨ ਕਰਨੇ ਹਨ। ਹਰ ਭਾਗ ਵਿਚੋਂ ਇਕ ਪ੍ਰਸ਼ਨ ਲਾਜ਼ਮੀ ਹੈ। ਪੰਜਵਾਂ ਪ੍ਰਸ਼ਨ ਕਿਸੇ ਵੀ ਭਾਗ ਵਿਚੋਂ ਕੀਤਾ ਜਾ ਸਕਦਾ ਹੈ।
3. ਹਰੇਕ ਪ੍ਰਸ਼ਨ ਦੇ ਬਰਾਬਰ ਅੰਕ ਹਨ।
4. ਪੇਪਰ ਸੈੱਟ ਕਰਨ ਵਾਲਾ ਜੇਕਰ ਚਾਹੇ ਤਾਂ ਪ੍ਰਸ਼ਨਾਂ ਦੀ ਵੰਡ ਅੱਗੋਂ ਵੱਧ ਤੋਂ ਵੱਧ ਚਾਰ ਉਪ-ਪ੍ਰਸ਼ਨਾਂ ਵਿੱਚ ਕਰ ਸਕਦਾ ਹੈ।

ਪਠ-ਕ੍ਰਮ
ਸੈਕਸ਼ਨ-ਏ

- ਪੰਜਾਬੀ ਭਾਸ਼ਾ ਤੇ ਗੁਰਮੁਖੀ ਲਿਪੀ
- (ੳ) ਨਾਮਕਰਣ ਤੇ ਸੰਖੇਪ ਜਾਣ-ਪਛਾਣ : ਗੁਰਮੁਖੀ ਵਰਣਮਾਲਾ, ਅਖਰ ਕ੍ਰਮ, ਸਵਰ ਵਾਹਕ (ੳ ਅ ਏ), ਲਗਾਂ-ਮਾਤਰਾਂ, ਪੈਰ ਵਿਚ ਬਿੰਦੀ ਵਾਲੇ ਵਰਣ, ਪੈਰ ਵਿਚ ਪੈਣ ਵਾਲੇ ਵਰਣ, ਬਿੰਦੀ, ਟਿਪੀ, ਅਧਕ।
- (ਅ) ਸਿਖਲਾਈ ਤੇ ਅਭਿਆਸ

ਸੈਕਸ਼ਨ-ਬੀ

ਗੁਰਮੁਖੀ ਆਰਥੋਗ੍ਰਾਫੀ ਅਤੇ ਉਚਾਰਨ : ਸਵਰ, ਵਿਅੰਜਨ : ਮੁਢਲੀ ਜਾਣ-ਪਛਾਣ ਅਤੇ ਉਚਾਰਣ, ਮੁਹਾਰਨੀ, ਲਗਾਂ-ਮਾਤਰਾਂ ਦੀ ਪਛਾਣ।

ਸੈਕਸ਼ਨ-ਸੀ

ਪੰਜਾਬੀ ਸ਼ਬਦ ਜੋੜ : ਮੁਕਤਾ (ਦੋ ਅਖਰਾਂ ਵਾਲੇ ਸ਼ਬਦ, ਤਿੰਨ ਅਖਰਾਂ ਵਾਲੇ ਸ਼ਬਦ), ਸਿਹਾਰੀ ਵਾਲੇ ਸ਼ਬਦ, ਬਿਹਾਰੀ ਵਾਲੇ ਸ਼ਬਦ, ਔਕੜ ਵਾਲੇ ਸ਼ਬਦ, ਦੁਲੈਕੜ ਵਾਲੇ ਸ਼ਬਦ, ਲਾਂ ਵਾਲੇ ਸ਼ਬਦ, ਦੁਲਾਵਾਂ ਵਾਲੇ ਸ਼ਬਦ, ਹੋੜੇ ਵਾਲੇ ਸ਼ਬਦ, ਕਨੌੜੇ ਵਾਲੇ ਸ਼ਬਦ।

ਸੈਕਸ਼ਨ-ਡੀ

ਲਗਾਖਰ (ਬਿੰਦੀ, ਟਿਪੀ, ਅਧਕ ਵਾਲੇ ਸ਼ਬਦ)
ਸ਼ੁਧ, ਅਸ਼ੁਧ (ਪੈਰੂ ਵਿਚ ਲਿਖੇ ਅਸ਼ੁਧ ਸ਼ਬਦਾਂ ਨੂੰ ਸ਼ੁਧ ਕਰਨਾ)

ਸਹਾਇਕ ਪੁਸਤਕਾਂ

1. ਬ੍ਰਹਮਜਗਦੀਸ਼ ਸਿੰਘ, ਪੰਜਾਬੀ ਭਾਸ਼ਾ ਅਤੇ ਗੁਰਮੁਖੀ ਲਿਪੀ, ਵਾਰਿਸ ਸ਼ਾਹ ਫਾਉਂਡੇਸ਼ਨ, ਅੰਮ੍ਰਿਤਸਰ।
2. ਪ੍ਰੋ. ਸ਼ੈਰੀ ਸਿੰਘ, ਪ੍ਰੋ. ਬ੍ਰਹਮਜਗਦੀਸ਼ ਸਿੰਘ, ਭਾਸ਼ਾ ਵਿਗਿਆਨ : ਪੰਜਾਬੀ ਭਾਸ਼ਾ ਤੇ ਗੁਰਮੁਖੀ ਲਿਪੀ, ਵਾਰਿਸ ਸ਼ਾਹ ਫਾਉਂਡੇਸ਼ਨ, ਅੰਮ੍ਰਿਤਸਰ।
3. ਪੰਜਾਬੀ ਭਾਸ਼ਾ ਬੋਧ, ਕਸਤੂਰੀ ਲਾਲ ਐਂਡ ਸੰਨਜ਼, ਅੰਮ੍ਰਿਤਸਰ।
4. ਮਿੰਨੀ ਸਲਵਾਨ, ਪੰਜਾਬੀ ਵਿਆਕਰਨ : ਮੁਢਲੇ ਸੰਕਲਪ, ਰਵੀ ਸਾਹਿਤ ਪ੍ਰਕਾਸ਼ਨ, ਅੰਮ੍ਰਿਤਸਰ।

13
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – I

PAPER–VI: Punjab History & Culture (From Earliest Times to C 320)
(Special Paper in lieu of Punjabi Compulsory)
(For those students who are not domicile of Punjab)

Time: 3 Hours

Max. Marks: 50

Instructions for the Paper Setters

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION- A

1. Physical features of the Punjab and its impact on history.
2. Sources of the ancient history of Punjab

SECTION- B

3. Harappan Civilization: Town planning; social, economic and religious life of the Indus Valley People.
4. The Indo-Aryans: Original home and settlements in Punjab.

SECTION- C

5. Social, Religious and Economic life during *Rig* Vedic Age.
6. Social, Religious and Economic life during Later Vedic Age.

SECTION- D

7. Teachings and impact of Buddhism
8. Jainism in the Punjab

Suggested Readings

1. L. M Joshi (ed.), *History and Culture of the Punjab*, Art-I, Patiala, 1989 (3rd edition)
2. L.M. Joshi and Fauja Singh (ed.), *History of Punjab*, Vol.I, Patiala 1977.
3. Budha Parkash, *Glimpses of Ancient Punjab*, Patiala, 1983.
4. B.N. Sharma, *Life in Northern India*, Delhi. 1966.
5. Chopra, P.N., Puri, B.N., & Das, M.N.(1974). *A Social, Cultural & Economic History of India*, Vol. I, New Delhi: Macmillan India.

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – II

Paper–I: Introduction to Data Science

Time: 3 Hours

Max. Marks: 75

Instructions for the Paper Setters:-

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION-A

Introduction: What is Data Science? - Big Data and Data Science hype – and getting past the hype - Why now? – Datafication - Current landscape of perspectives - Skill sets needed

Statistical Inference - Populations and samples - Statistical modeling, probability distributions, fitting a model - Intro to R

Exploratory Data Analysis and the Data Science Process - Basic tools (plots, graphs and summary statistics) of EDA - Philosophy of EDA - The Data Science Process - Case Study: Real Direct (online real estate firm)

SECTION-B

Three Basic Machine Learning Algorithms - Linear Regression - k-Nearest Neighbors (k-NN) - k-means

One More Machine Learning Algorithm and Usage in Applications - Motivating application: Filtering Spam - Why Linear Regression and k-NN are poor choices for Filtering Spam - Naive Bayes and why it works for Filtering Spam - Data Wrangling: APIs and other tools for scrapping the Web

Feature Generation and Feature Selection (Extracting Meaning From Data) - Motivating application: user (customer) retention - Feature Generation (brainstorming, role of domain expertise, and place for imagination) - Feature Selection algorithms – Filters; Wrappers; Decision Trees; Random Forests

SECTION-C

Recommendation Systems: Building a User-Facing Data Product - Algorithmic ingredients of a Recommendation Engine - Dimensionality Reduction - Singular Value Decomposition - Principal Component Analysis - Exercise: build your own recommendation system

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – II

Mining Social-Network Graphs - Social networks as graphs - Clustering of graphs - Direct discovery of communities in graphs - Partitioning of graphs - Neighborhood properties in graphs

SECTION-D

Data Visualization - Basic principles, ideas and tools for data visualization 3 - Examples of inspiring (industry) projects - Exercise: create your own visualization of a complex dataset

Data Science and Ethical Issues - Discussions on privacy, security, ethics - A look back at Data Science - Next-generation data scientists

Text Books:

1. Cathy O’Neil and Rachel Schutt. Doing Data Science, Straight Talk from The Frontline. O’Reilly. 2014.
2. Mohammed J. Zaki and Wagner Miera Jr. Data Mining and Analysis: Fundamental Concepts and Algorithms. Cambridge University Press. 2014.
3. Jiawei Han, Micheline Kamber and Jian Pei. Data Mining: Concepts and Techniques, Third Edition.

16
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – II

Paper–II: Basic Mathematics

Time: 3 Hours

Max. Marks: 75

Instructions for the Paper Setters:-

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION-A

Solution of Simultaneous Linear Equations (upto two variable case), Solution of Quadratic Equations. Series: Arithmetic Progression Series, Geometric Progression Series

SECTION-B

Permutations and Combinations, Binomial Theorem, Determinants with simple applications for solution of Linear simultaneous equations using Cramer's Rule, Matrices with simple application for solution of linear simultaneous equations using matrix inversion method.

SECTION-C

Real number systems, constants and variables, functions. Graphical representations of functions, limits and continuity of functions, first principle of differential calculus, derivations of simple algebraic functions and application of derivatives in Economic and Commerce. Maximum and minimum.

SECTION-D

General form of linear Programming, formulating Linear Programming Problems assumptions, Graphic Method, The Standard Maximum and Minimum Problems, Simplex Method, Duality, Dual Linear Programming Problems

Books Recommended:

1. Business Mathematics by Padmalochan Hazarika.
2. Business Mathematics by D.C. Sancheti and V.K. Kapoor.
3. Mathematical Economics by Dowling, T. Edward.
4. Linear Programming by Thomas S. Ferguson

Paper–III: Introduction to R

Time: 3 Hours

Max. Marks: 75

Instructions for the Paper Setters:-

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION - A

Introduction and preliminaries: The R environment , Related software and documentation , R and statistics, R and the window system, Using R interactively, An introductory session ,Getting help with functions and features, R commands, case sensitivity, Recall and correction of previous commands, Executing commands from or diverting output to a file ,Data permanency and removing objects
2 Simple manipulations; numbers and vectors
Vectors and assignment, Vector arithmetic, Generating regular sequences, Logical vectors, Missing values, Character vectors, Index vectors; selecting and modifying subsets of a data set, Other types of objects

Objects, their modes and attributes: Intrinsic attributes: mode and length, Changing the length of an object, Getting and setting attributes, The class of an object

SECTION - B

Ordered and unordered factors: A specific example , The function `tapply()` and ragged arrays, Ordered factors

Arrays and matrices: Arrays, Array indexing. Subsections of an array, Index matrices, The `array()` function, Mixed vector and array arithmetic. The recycling rule, The outer product of two arrays, Generalized transpose of an array, Matrix facilities, Matrix multiplication, Linear equations and inversion, Eigenvalues and eigenvectors, Singular value decomposition and determinants, Least squares fitting and the QR decomposition, Forming partitioned matrices, `cbind()` and `rbind()`,The concatenation function, `c()`, with arrays, Frequency tables from factors.

Lists and data frames: Lists., Constructing and modifying lists, Concatenating lists, Data frames, Making data frames, `attach()` and `detach()`,Working with data frames, Attaching arbitrary lists, Managing the search path

Reading data from files: The `read.table()` function, The `scan()` function., Accessing built in datasets, Loading data from other R packages, Editing data

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – II

SECTION – C

Probability distributions R as a set of statistical tables, Examining the distribution of a set of data, One- and two-sample tests

Grouping, loops and conditional execution, Grouped expressions, Control statements, Conditional execution: if statements, Repetitive execution: for loops, repeat and while

Writing your own functions: Simple examples, Defining new binary operators, Named arguments and defaults, The ‘...’ argument, Assignments within functions, More advanced examples, Efficiency factors in block designs, Dropping all names in a printed array, Recursive numerical integration, Scope, Customizing the environment, Classes, generic functions and object orientation

SECTION – D

Statistical Models in R: Defining statistical models; formulae, Contrasts, Linear models, Generic functions for extracting model information, Analysis of variance and model comparison, ANOVA tables, Updating fitted models, Generalized linear models, Families, The glm() function, Nonlinear least squares and maximum likelihood models, Least squares, Maximum likelihood, Some non-standard models

Graphical Procedures: High-level plotting commands, The plot() function, Displaying multivariate data, Display graphics, Arguments to high-level plotting functions, Low-level plotting commands, Mathematical annotation, Hershey vector fonts, Interacting with graphics, Using graphics parameters, Permanent changes: The par() function, Temporary changes: Arguments to graphics functions, Graphics parameters list, Graphical elements, Axes and tick marks, Figure margins, Multiple figure environment, Device drivers, PostScript diagrams for typeset documents, Multiple graphics devices, Dynamic graphics

Packages, Standard packages: Contributed packages and CRAN, Namespaces

OS facilities: Files and directories, Filepaths, System commands, Compression and Archives.

Text/References:

1. An Introduction to R by W. N. Venables
2. Statistics: An Introduction Using R by Michael J. Crawley
3. R in Action: Data Analysis and Graphics with R by Robert Kabacoff

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – II

**Paper–IV: Practical Based on Programming in R
(Practical)**

Time: 3 Hours

Max. Marks: 75

Note: Practical exam to be conducted by the external examiner.

Practical Based on Programming in R

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – II

PAPER–V: COMMUNICATION SKILLS IN ENGLISH – II

Time: 3 Hours

Max. Marks: 50

Theory Marks: 35

Practical Marks: 15

Instructions for the Paper Setters: -

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

Course Contents:

SECTION–A

Listening Skills: Barriers to listening; effective listening skills; feedback skills.

Activities: Listening exercises – Listening to conversation, News and TV reports

SECTION–B

Attending telephone calls; note taking and note making.

Activities: Taking notes on a speech/lecture

SECTION–C

Speaking and Conversational Skills: Components of a meaningful and easy conversation; understanding the cue and making appropriate responses; forms of polite speech; asking and providing information on general topics.

Activities: 1) Making conversation and taking turns

2) Oral description or explanation of a common object, situation or concept

SECTION–D

The study of sounds of English,

Stress and Intonation,

Situation based Conversation in English,

Essentials of Spoken English.

Activities: Giving Interviews

21
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – II

PRACTICAL / ORAL TESTING

Marks: 15

Course Contents: -

1. Oral Presentation with/without audio visual aids.
2. Group Discussion.
3. Listening to any recorded or live material and asking oral questions for listening comprehension.

Questions: -

1. Oral Presentation will be of 5 to 10 minutes duration (Topic can be given in advance or it can be student's own choice). Use of audio visual aids is desirable.
2. Group discussion comprising 8 to 10 students on a familiar topic. Time for each group will be 15 to 20 minutes.

Note: Oral test will be conducted by external examiner with the help of internal examiner.

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – II

PAPER–VI: ਪੰਜਾਬੀ (ਲਾਜਮੀ)

ਸਮਾਂ : 3 ਘੰਟੇ

ਕੁਲ ਅੰਕ : 50

ਅੰਕ-ਵੰਡ ਅਤੇ ਪਰੀਖਿਅਕ ਲਈ ਹਦਾਇਤਾਂ

1. ਪ੍ਰਸ਼ਨ ਪੱਤਰ ਦੇ ਚਾਰ ਭਾਗ ਹੋਣਗੇ। ਹਰ ਭਾਗ ਵਿਚ ਦੋ ਪ੍ਰਸ਼ਨ ਪੁੱਛੇ ਜਾਣਗੇ।
2. ਵਿਦਿਆਰਥੀ ਨੇ ਕੁੱਲ ਪੰਜ ਪ੍ਰਸ਼ਨ ਕਰਨੇ ਹਨ। ਹਰ ਭਾਗ ਵਿਚੋਂ ਇਕ ਪ੍ਰਸ਼ਨ ਲਾਜਮੀ ਹੈ। ਪੰਜਵਾਂ ਪ੍ਰਸ਼ਨ ਕਿਸੇ ਵੀ ਭਾਗ ਵਿਚੋਂ ਕੀਤਾ ਜਾ ਸਕਦਾ ਹੈ।
3. ਹਰੇਕ ਪ੍ਰਸ਼ਨ ਦੇ ਬਰਾਬਰ ਅੰਕ ਹਨ।
4. ਪੇਪਰ ਸੈੱਟ ਕਰਨ ਵਾਲਾ ਜੇਕਰ ਚਾਹੇ ਤਾਂ ਪ੍ਰਸ਼ਨਾਂ ਦੀ ਵੰਡ ਅੱਗੋਂ ਵੱਧ ਤੋਂ ਵੱਧ ਚਾਰ ਉਪ-ਪ੍ਰਸ਼ਨਾਂ ਵਿੱਚ ਕਰ ਸਕਦਾ ਹੈ।

ਪਾਠ-ਕ੍ਰਮ ਅਤੇ ਪਾਠ-ਪੁਸਤਕਾਂ

ਸੈਕਸ਼ਨ-ਏ

ਸਰਵੋਤਮ ਪੰਜਾਬੀ ਸਾਹਿਤ (ਸੰਪਾ. ਡਾ. ਰਮਿੰਦਰ ਕੌਰ, ਡਾ. ਮੇਘਾ ਸਲਵਾਨ)
 (ਨਿਬੰਧ ਅਤੇ ਰੇਖਾ ਚਿਤਰ)
 ਰੇਖਾ ਚਿਤਰ ਦਾ ਨਾਇਕ ਬਿੰਬ, ਵਿਸ਼ਾ ਵਸਤੂ

ਸੈਕਸ਼ਨ-ਬੀ

ਇਤਿਹਾਸਕ ਯਾਦਾਂ
 (ਜੀਵਨੀ 7 ਤੋਂ 12 ਤਕ)
 ਵਿਸ਼ਾ ਵਸਤੂ/ਸਾਰ/ਨਾਇਕ ਬਿੰਬ

ਸੈਕਸ਼ਨ-ਸੀ

- (ੳ) ਸ਼ਬਦ ਬਣਤਰ ਅਤੇ ਸ਼ਬਦ ਰਚਨਾ, ਪਰਿਭਾਸ਼ਾ, ਮੁਢਲੇ ਸੰਕਲਪ
 (ਅ) ਸ਼ਬਦ ਸ਼੍ਰੇਣੀਆਂ

ਸੈਕਸ਼ਨ-ਡੀ

ਦਫ਼ਤਰੀ ਚਿਠੀ ਪਤਰ
 ਮੁਹਾਵਰੇ ਅਤੇ ਅਖਾਣ

ਸਹਾਇਕ ਪੁਸਤਕਾਂ

1. ਸਤਿੰਦਰ ਸਿੰਘ, **ਆਧੁਨਿਕ ਪੰਜਾਬੀ ਵਾਰਤਕ ਦਾ ਇਤਿਹਾਸ**, ਪੰਜਾਬੀ ਅਕਾਦਮੀ, ਦਿਲੀ।
2. ਪ੍ਰੋ. ਪਿਆਰਾ ਸਿੰਘ, **ਪੰਜਾਬੀ ਵਾਰਤਕ : ਸਿਧਾਂਤ ਇਤਿਹਾਸ ਪ੍ਰਵਿਰਤੀਆਂ**, ਨਿਊ ਬੁਕ ਕੰਪਨੀ, ਜਲੰਧਰ।
3. ਇੰਦਰਪ੍ਰੀਤ ਸਿੰਘ ਧਾਮੀ, **ਪੰਜਾਬੀ ਰੇਖਾ ਚਿਤਰ : ਰੂਪ ਤੇ ਪ੍ਰਕਾਸ਼**, ਰਵੀ ਸਾਹਿਤ ਪ੍ਰਕਾਸ਼ਨ, ਅੰਮ੍ਰਿਤਸਰ।
4. ਬਲਬੀਰ ਸਿੰਘ ਦਿਲ, **ਪੰਜਾਬੀ ਨਿਬੰਧ : ਸਰੂਪ, ਸਿਧਾਂਤ ਅਤੇ ਵਿਕਾਸ**, ਪੰਜਾਬੀ ਯੂਨੀਵਰਸਿਟੀ, ਪਟਿਆਲਾ।
5. ਹਰਕੀਰਤ ਸਿੰਘ ਤੇ ਗਿਆਨੀ ਲਾਲ ਸਿੰਘ, **ਕਾਲਜ ਪੰਜਾਬੀ ਵਿਆਕਰਨ**, ਪੰਜਾਬ ਯੂਨੀਵਰਸਿਟੀ, ਚੰਡੀਗੜ੍ਹ।
6. ਡਾ. ਅਮਰ ਕੋਮਲ (ਸੰਪਾ.), **ਚੋਣਵੇਂ ਪੰਜਾਬੀ ਨਿਬੰਧ** (ਭੂਮਿਕਾ), ਨੈਸ਼ਨਲ ਬੁਕ ਟਰਸਟ, ਇੰਡੀਆ।
7. ਅਬਨਾਸ ਕੌਰ, **ਪੰਜਾਬੀ ਰੇਖਾ ਚਿਤਰ**, ਪੰਜਾਬੀ ਯੂਨੀਵਰਸਿਟੀ, ਪਟਿਆਲਾ।
8. ਮਿੰਨੀ ਸਲਵਾਨ, **ਪੰਜਾਬੀ ਵਿਆਕਰਨ : ਮੁਢਲੇ ਸੰਕਲਪ**, ਰਵੀ ਸਾਹਿਤ ਪ੍ਰਕਾਸ਼ਨ, ਅੰਮ੍ਰਿਤਸਰ।
9. ਬੂਟਾ ਸਿੰਘ ਬਰਾੜ, **ਪੰਜਾਬੀ ਵਿਆਕਰਨ : ਸਿਧਾਂਤ ਤੇ ਵਿਹਾਰ**, ਚੇਤਨਾ ਪ੍ਰਕਾਸ਼ਨ, ਲੁਧਿਆਣਾ।

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – II

PAPER–VI ਮੁੱਢਲੀ ਪੰਜਾਬੀ
(In lieu of Compulsory Punjabi)
(For those students who are not domicile of Punjab)

ਸਮਾਂ: 3 ਘੰਟੇ

ਕੁਲ ਅੰਕ: 50

1. ਪ੍ਰਸ਼ਨ ਪੱਤਰ ਦੇ ਚਾਰ ਭਾਗ ਹੋਣਗੇ। ਹਰ ਭਾਗ ਵਿਚ ਦੋ ਪ੍ਰਸ਼ਨ ਪੁੱਛੇ ਜਾਣਗੇ।
2. ਵਿਦਿਆਰਥੀ ਨੇ ਕੁੱਲ ਪੰਜ ਪ੍ਰਸ਼ਨ ਕਰਨੇ ਹਨ। ਹਰ ਭਾਗ ਵਿਚੋਂ ਇਕ ਪ੍ਰਸ਼ਨ ਲਾਜ਼ਮੀ ਹੈ। ਪੰਜਵਾਂ ਪ੍ਰਸ਼ਨ ਕਿਸੇ ਵੀ ਭਾਗ ਵਿਚੋਂ ਕੀਤਾ ਜਾ ਸਕਦਾ ਹੈ।
3. ਹਰੇਕ ਪ੍ਰਸ਼ਨ ਦੇ ਬਰਾਬਰ ਅੰਕ ਹਨ।
4. ਪੇਪਰ ਸੈੱਟ ਕਰਨ ਵਾਲਾ ਜੇਕਰ ਚਾਹੇ ਤਾਂ ਪ੍ਰਸ਼ਨਾਂ ਦੀ ਵੰਡ ਅੱਗੋਂ ਵੱਧ ਤੋਂ ਵੱਧ ਚਾਰ ਉਪ-ਪ੍ਰਸ਼ਨਾਂ ਵਿੱਚ ਕਰ ਸਕਦਾ ਹੈ।

ਪਾਠ-ਕ੍ਰਮ

ਸੈਕਸ਼ਨ-ਏ

ਪੰਜਾਬੀ ਸ਼ਬਦ ਬਣਤਰ : ਧਾਤੂ, ਵਧੇਤਰ (ਅਗੇਤਰ, ਮਧੇਤਰ, ਪਿਛੇਤਰ), ਪੰਜਾਬੀ ਕੋਸ਼ਗਤ ਸ਼ਬਦ ਅਤੇ ਵਿਆਕਰਣਿਕ ਸ਼ਬਦ

ਸੈਕਸ਼ਨ-ਬੀ

- (ੳ) ਸੰਯੁਕਤ ਸ਼ਬਦ, ਸਮਾਸੀ ਸ਼ਬਦ, ਦੋਜਾਤੀ ਸ਼ਬਦ, ਦੋਹਰੇ/ਦੁਹਰਕਤੀ ਸ਼ਬਦ ਅਤੇ ਮਿਸ਼ਰਤ ਸ਼ਬਦ
- (ਅ) ਸਿਖਲਾਈ ਤੇ ਅਭਿਆਸ

ਸੈਕਸ਼ਨ-ਸੀ

ਇਕ-ਵਚਨ, ਬਹੁ-ਵਚਨ, ਲਿੰਗ-ਪੁਲਿੰਗ, ਬਹੁ-ਅਰਥਕ ਸ਼ਬਦ, ਸਮਾਨ-ਅਰਥਕ ਸ਼ਬਦ, ਬਹੁਤੇ ਸ਼ਬਦਾਂ ਲਈ ਇਕ ਸ਼ਬਦ, ਸ਼ਬਦ ਜੋੜ, ਵਿਰੋਧਆਰਥਕ ਸ਼ਬਦ।

ਸੈਕਸ਼ਨ-ਡੀ

ਨਿਤ ਵਰਤੋਂ ਦੀ ਪੰਜਾਬੀ ਸ਼ਬਦਾਵਲੀ : ਖਾਣ-ਪੀਣ, ਸਾਕਾਦਾਰੀ, ਰੁਤਾਂ, ਮਹੀਨਿਆਂ, ਗਿਣਤੀ, ਮੌਸਮ, ਮਾਰਕੀਟ/ਬਾਜ਼ਾਰ, ਵਪਾਰ, ਧੰਦਿਆਂ ਆਦਿ ਨਾਲ ਸੰਬੰਧਿਤ।

ਸਹਾਇਕ ਪੁਸਤਕਾਂ

1. ਜੋਤੀ ਸ਼ਰਮਾ, ਪੰਜਾਬੀ ਵਿਆਕਰਨ, ਵਾਰਿਸ ਸ਼ਾਹ ਫਾਊਂਡੇਸ਼ਨ, ਅੰਮ੍ਰਿਤਸਰ।
2. ਪੰਜਾਬੀ ਭਾਸ਼ਾ ਬੋਧ, ਕਸਤੂਰੀ ਲਾਲ ਐਂਡ ਸੰਨਜ਼, ਅੰਮ੍ਰਿਤਸਰ।
3. ਸਿੰਨੀ ਸਲਵਾਨ, ਪੰਜਾਬੀ ਵਿਆਕਰਨ : ਮੁਢਲੇ ਸੰਕਲਪ, ਰਵੀ ਸਾਹਿਤ ਪ੍ਰਕਾਸ਼ਨ, ਅੰਮ੍ਰਿਤਸਰ।
4. ਰੰਜੂ ਬਾਲਾ, ਅਰਥ ਵਿਗਿਆਨ, ਆਰਸੀ ਪਬਲਿਸ਼ਰਜ਼, ਦਿਲੀ।

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – II

PAPER–VI: Punjab History & Culture (C 320 to 1000 B.C.)
(Special Paper in lieu of Punjabi compulsory)
(For those students who are not domicile of Punjab)

Time: 3 Hours

Max. Marks :50

Instructions for the Paper Setters

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION- A

1. Alexander's Invasion and its Impact
2. Punjab under Chandragupta Maurya and Ashoka.

SECTION- B

3. The Kushans and their Contribution to the Punjab.
4. The Panjab under the Gupta Empire.

SECTION- C

5. The Punjab under the Vardhana Emperors
6. Socio-cultural History of Punjab from 7th to 1000 A.D.

SECTION- D

7. Development of languages and Education with Special reference to Taxila
8. Development of Art & Architecture

Suggested Readings

1. L. M Joshi (ed), *History and Culture of the Punjab*, Art-I, Punjabi University, Patiala, 1989 (3rd edition)
2. L.M. Joshi and Fauja Singh (ed.), *History of Punjab* , Vol.I, Punjabi University, Patiala, 1977.
3. Budha Parkash, *Glimpses of Ancient Punjab*, Patiala, 1983.
4. B.N. Sharma: *Life in Northern India*, Delhi. 1966.

25
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – II

**PAPER–VII DRUG ABUSE: PROBLEM, MANAGEMENT AND PREVENTION
(Compulsory Paper)**

Time: 3 Hours

Max. Marks: 100

Instructions for the Paper Setters:-

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

Section – A

Meaning of Drug Abuse:

- 1) Meaning, Nature and Extent of Drug Abuse in India and Punjab.
- 2) Consequences of Drug Abuse for:
 - Individual : Education, Employment, Income.
 - Family : Violence.
 - Society : Crime.
 - Nation : Law and Order problem.

Section – B

Management of Drug Abuse:

- (i) Medical Management: Medication for treatment and to reduce withdrawal effects.
- (ii) Psychiatric Management: Counselling, Behavioural and Cognitive therapy.
- (iii) Social Management: Family, Group therapy and Environmental Intervention.

Section – C

Prevention of Drug abuse:

- (i) Role of family: Parent child relationship, Family support, Supervision, Shaping values, Active Scrutiny.
- (ii) School: Counselling, Teacher as role-model. Parent-teacher-Health Professional Coordination, Random testing on students.

Section – D

Controlling Drug Abuse:

- (i) Media: Restraint on advertisements of drugs, advertisements on bad effects of drugs, Publicity and media, Campaigns against drug abuse, Educational and awareness program
- (ii) Legislation: NDPs act, Statutory warnings, Policing of Borders, Checking Supply/Smuggling of Drugs, Strict enforcement of laws, Time bound trials.

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – II

References:

1. Ahuja, Ram (2003), *Social Problems in India*, Rawat Publication, Jaipur.
2. Extent, Pattern and Trend of Drug Use in India, Ministry of Social Justice and Empowerment, Government of India, 2004.
3. Inciardi, J.A. 1981. *The Drug Crime Connection*. Beverly Hills: Sage Publications.
4. Kapoor. T. (1985) *Drug epidemic among Indian Youth*, New Delhi: Mittal Pub.
5. Kessel, Neil and Henry Walton. 1982, *Alcoholism*. Harmond Worth: Penguin Books.
6. Modi, Ishwar and Modi, Shalini (1997) *Drugs: Addiction and Prevention*, Jaipur: Rawat Publication.
7. National Household Survey of Alcohol and Drug abuse. (2003) New Delhi, Clinical Epidemiological Unit, All India Institute of Medical Sciences, 2004.
8. Ross Coomber and Others. 2013, *Key Concept in Drugs and Society*. New Delhi: Sage Publications.
9. Sain, Bhim 1991, *Drug Addiction Alcoholism, Smoking obscenity* New Delhi: Mittal Publications.
10. Sandhu, Ranvinder Singh, 2009, *Drug Addiction in Punjab: A Sociological Study*. Amritsar: Guru Nanak Dev University.
11. Singh, Chandra Paul 2000. *Alcohol and Dependence among Industrial Workers*: Delhi: Shipra.
12. Sussman, S and Ames, S.L. (2008). *Drug Abuse: Concepts, Prevention and Cessation*, Cambridge University Press.
13. Verma, P.S. 2017, “*Punjab’s Drug Problem: Contours and Characteristics*”, Economic and Political Weekly, Vol. LII, No. 3, P.P. 40-43.
14. World Drug Report 2016, United Nations office of Drug and Crime.
15. World Drug Report 2017, United Nations office of Drug and Crime.

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – III

Paper – I: Optimization

Time: 3 Hours

Max. Marks: 50

Instructions for the Paper Setters:-

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION-A

Meaning, significance, limitation and scope. Introduction to linear programming, formation of Linear Programming Problem, Graphical method Simplex Method.

SECTION-B

Two Phase Simplex Method , Duality in Linear Programming, Definition of Dual Problem, general rules of converting primal into its dual.

SECTION-C

Transportation Problem, Definition of Assignment Model, Hungarian Method for solution of Assignment Problems, Travelling Salesman Problem.

SECTION-D

Games Theory: Two persons zero sum games, pure strategies, mixed strategies, Dominance.

Text/References:

1. Swaroop, K., Gupta, P.K. and Manmohan, "Operations Research", 2013, 18th Edition, Sultan Chand & Sons, New Delhi.
2. Gupta, P.K. and Hira, D.S., "Operations Research", 2009, S. Chand & Co., New Delhi.

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – III

Paper – II: Business Economics

Time: 3 Hours

Max. Marks 50

Instructions for Paper Setter:

The whole syllabus has been divided into four sections (A-D). Eight questions of equal marks (two from each section) in all will be set. Candidates are required to answer total five questions, selecting at least one question from each section. The fifth question may be from any section. Questions may have sub-parts, not more than four.

SECTION-A

Consumer Behaviour:

Theory of Demand- Meaning, types; Law of Demand. Price Elasticity of Demand and its measurement.

Theories of Consumer Behaviour – Cardinal Utility Analysis- Brief outline of Law of DMU, EMU and consumer equilibrium.

Ordinal Utility Analysis- Meaning, properties and consumer equilibrium. Brief outline of Price Effect, Income effect and Substitution effect.

SECTION-B

Producer Behaviour:

Theory of Production- Law of Variable Proportions and Laws of Returns to Scale.

Cost & Cost Curves- The Traditional theory of Cost (short run and long run).

Concept of Revenue- Total Revenue, Average Revenue and Marginal Revenue; their inter-relationship.

SECTION-C

Market Forms:

Perfect Competition- Meaning, Features. Conditions of Equilibrium; Price and Output determination of firm and industry.

Monopoly- Meaning, Features. Price and Output determination under Monopoly.

Monopolistic Competition- Meaning, Features. Price and Output determination under Monopolistic Competition.

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – III

SECTION-D

National Income:

Meaning, Definition and Importance of National Income.

Important Aggregates of National Income- GDP at Market Price, GNP at Market Price, NDP at Market Price, NNP at Market Price, GNP at Factor Cost, NNP at Factor Cost, Personal Income, Disposable Income.

Measurement of National Income- Product Method, Income Method and Expenditure Method.

Problems in the measurement of National Income, particularly in UDCs.

Suggested Readings:

1. Ahuja, H.L., “Modern Micro Economics”, (2009), Sultan Chand and Co., New Delhi.
2. Dwivedi, D.N., “Managerial Economics”, 7th Edition, Vikas Publication.
3. Koutsoyiannis, A., “Modern Micro Economics”, 2nd Edition, McMillan House, New Delhi.
4. Froyen, R., “Macroeconomics”, 9th Edition (2008), Pearson Publication.

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – III

Paper – III: Statistical Inference-I

Time: 3 Hours

Max. Marks: 75

Instructions for the Paper Setters:-

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION-A

Cumulative distribution function, Two dimensional random variables, joint distribution, marginal and conditional distributions, Stochastic independence, Introduction to function of random variables.

SECTION-B

Mathematical expectations and moments, moment generating function and its properties Chebyshev's inequality and its application, central limit theorem (Laplace Theorem)

SECTION-C

Discrete Probability Distributions: Binomial, Poisson, Geometric, Continuous probability distributions: Uniform, Exponential, Gamma, Beta, Normal distributions.

SECTION-D

Sampling Distributions: Chi-square, t and F-distributions with their properties, distribution of sample mean and variance. Introduction to Estimators, Types of Estimators

Text/References:

1. Hogg R.V., Mckean, J.W. and Craig A.T. : Introduction to Mathematical Statistics
2. Gupta S.C. and Kapoor V.K. : Fundamentals of mathematical statistics
3. Goon,A.M.,Gupta M.K. & Dasgupta B. : Fundamental of statistic, Vol. I
4. Goon,A.M.,Gupta M.K. & Dasgupta B. : An outline of statistical theory, Vol. I

31
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – III

Paper–IV: Data Mining

Time: 3 Hours

Max. Marks: 75

Instructions for the Paper Setters:-

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION A

Data Warehousing: Architecture of a data warehouse; Differences between Online Transaction Processing (OLTP) and Online Analytical Processing (OLAP).

SECTION B

From Data Warehousing to Data Mining, Why Data Mining, What Is Data Mining, What Kinds of Data Can Be Mined, What Kinds of Patterns Can Be Mined, Which Technologies Are Used, Which kinds of applications are Targeted, Fundamentals of data mining, Data Mining Functionalities, Classification of Data Mining systems.

SECTION C

Major issues in Data Mining: Data preprocessing Descriptive data mining: characterization and comparison.

Data mining techniques: Association rule analysis Cluster analysis: Types of data – Clustering Methods – Partitioning methods – Model based clustering methods

Outlier Detection: Outliers and Outlier Analysis, Outlier Detection Methods: Supervised, Semi-Supervised, and Unsupervised Methods, Statistical Methods, Proximity-Based Methods, and Clustering-Based Methods. Statistical Approaches, Parametric Methods, Nonparametric Methods, **Proximity-Based Approaches:** Distance-Based Outlier Detection and a Nested Loop Method, A Grid-Based Method, Density-Based Outlier Detection, Clustering-Based Approaches, Classification-Based Approaches, Mining Contextual and Collective Outliers, Outlier Detection in High-Dimensional Data.

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – III

SECTION D

Classification – Decision Tree Induction – Bayesian Classification – Prediction – Back Propagation **Case studies** in Data Mining applications.

Data Mining Trends and Research Frontiers, Mining Complex Data Types, Mining Sequence Data: Time-Series, Symbolic Sequences, and Biological Sequences, Mining Graphs and Networks, Mining Other Kinds of Data, **Other Methodologies** of Data Mining, Statistical Data Mining, Views on Data Mining Foundations, Visual and Audio Data Mining.

References:

1. Data Mining: Concepts and Techniques, Jiawei Han, Micheline Kamber, Morgan Kaufmann;
2. 2nd Edition (2006)
3. Data Mining Introductory and Advanced Topics –Margaret H Dunham, Pearson Education
4. Data Warehousing in the Real World – Sam Anahory & Dennis Murray. Pearson Edn Asia.
5. Data Warehousing Fundamentals – Paulraj Ponnaiah Wiley Student Edition.

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – III

**Paper–V: Practical Based on SAS
(PRACTICAL)**

Time: 3 Hours

Max. Marks: 75

Note: Eight Programs of equal marks (Specified In the syllabus) are to be set, two in each of the four sections(A-D). Candidates are requested to attempt four programs, selecting at least one question one question from each section.

SECTION–A

SETTING UP THE SAS SOFTWARE ENVIRONMENT

What does SAS do? What is your preparation of SAS? Let's get started with your free version of SAS, history of SAS interfaces.

SAS Studio web-based GUI: Describing the rest of SAS Studio

(SAS Studio section- Server files and folders, tasks and utilities, snippets, libraries, file shortcuts)

SAS programming languages: first SAS data step program, first use of a SAS PROC, Saving a SAS program, Creating a new SAS program, The AUTOEXEC file, visual programmer versus SAS Programmer, what's in the SAS University Edition? Different levels of the SAS analytic platform

SAS data storage: the SAS dataset, the SAS scalable performance data engine, the scalable performance data server, SAS HDAT, SAS formats and informats, date and time data

SECTION–B

WORKING WITH DATA USING SAS SOFTWARE

Preparing data for analytics: making data in SAS, Data step to make data, PROC SQL to make data

Working with external data: data step code for importing external data, PROC IMPORT,

Referencing external files directly referencing external files, indirectly referencing external files.

Specialty PROCs for working with external data: PROC HADOOP AND PROC HDMD, PROC JSON

Specialty PROCs for working with computer languages: PROC GROOVY, LUA

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – III

SECTION–C

DATA PREPARATION USING SAS DATA STEP AND SAS PROCEDURES

Data preparation for analytics: Creating indicators for the first and last observation in a by group transposing: PROC TRANSPOSE, SAS studio transpose data task. Statistical and Mathematical data transformations: PROC MEANS, Imputation, Identifying missing values, Characterizing data, List Tables Attributes.

SAS macro facility: macro variables, macros.

SECTION–D

ANALYSIS WITH SAS SOFTWARE Analytics

Description and predictive analysis, descriptive analysis: PROC FREQ, CORR, UNIVARIATE.

Predictive analysis: regression analysis, PROC REG.

Forecasting analysis: PROC TIMEDATA, ARIMA.

Optimization analysis: SASA/IML, Interacting with the R programming language, PROC IML

35
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – III

**PAPER-VI: INTRODUCTION TO PYTHON
(PRACTICAL)**

Time: 3 Hours

Max. Marks: 75

Note:- Eight programs of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Candidates are required to attempt four programs, selecting at least one question from each Section.

SECTION-A

Python Programming Language Data and Expressions: Literals, Variables and Identifiers, Operators, Expressions, Statements and Data Types

Control Structures: Boolean Expressions (Conditions), Logical Operators, Selection Control, Nested conditions

Lists: List Structures, Lists (Sequences) in Python, Iterating Over Lists (Sequences) in Python

SECTION-B

Functions: Fundamental Concepts, Program Routines, Flow of Execution, Parameters & Arguments

Iteration: While statement, Definite loops using For, Loop Patterns, Recursive Functions

SECTION-C

Dictionaries: Dictionaries and Files, Looping and dictionaries

Files: Opening Files, Using Text Files, String Processing, Exception Handling

SECTION-D

Objects and Their Use: Introduction to Object Oriented Programming

Modular Design: Modules, Top-Down Design, Python Modules

Using Databases and SQL: Database Concepts, SQLite Manager Firefox Add-on, SQL basic summary, Basic Data modeling, Programming with multiple tables

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – III

Reference Books:

1. Python for Informatics, Charles Severance, version 0.0.7
2. Introduction to Computer Science Using Python: A Computational Problem-Solving Focus, Charles Dierbach, Wiley Publications, 2012, ISBN: 978-0-470-91204-1
3. Introduction to Computation And Programming Using Python, GUTTAG JOHN V, PHI, 2014, ISBN-13: 978-8120348660
4. Introduction to Computating & Problem Solving Through Python, Jeeva Jose and Sojan P. Lal, Khanna Publishers, 2015, ISBN-13: 978-9382609810
5. Introduction to Computing and Programming in Python, Mark J. Guzdial, Pearson Education, 2015, ISBN-13: 978-9332556591

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – IV

Paper–I: Basics of Linear Algebra & Numerical Analysis

Time: 3 Hours

Max. Marks: 50

Instructions for the Paper Setters:-

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION-A

Definition of groups, rings and fields with examples. Definition of a vector space, subspaces with examples. Linear dependence, Linear independence of vectors. Linear combination of vectors

SECTION-B

Solution of non-linear equations, Bisection Method, Method of false position, Secant Method, Newton's Method. Solution of linear system of equation; Gauss elimination Method, Gauss Jordan Method, Gauss Seidel Method.

SECTION –C

Numerical Differentiation, Numerical Integration: Trapezoidal rule, Simpson's 1/3 rule, Simpson's 3/8 rule

SECTION-D

Interpolation: Lagrangian Interpolation, Newton's Methods: Forward Difference Method, Backward Difference Method, Divided Difference Method Curve Fitting: Method of Least Square, Fitting of Straight Line, Fitting of a Polynomial

Books Recommended:

1. V. Rajaraman: Computer Oriented Numerical Methods, Prentice Hall of India Private Ltd., New Delhi.
2. B.S. Grewal, Numerical Methods for Engineering, Sultan Chand Publication.
3. K.Hoffman & R. Kunze: Linear Algebra, 2nd Edition, Prentice Hall, New Jersey, 1971.
4. Surjit Singh: Linear Algebra, 1997.
5. S.S. Sastry: Introductory Methods of Numerical Analysis, 2003 (3rd Edition), Prentice Hall of India.
6. A. Maritava Gupta and Subash Ch. Bose: Introduction to Numerical Analysis

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – IV

Paper – II: Statistical Inference-II

Time: 3 Hours

Max. Marks: 50

Instructions for the Paper Setters:-

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION-A

Basics of Estimators: Properties of unbiasedness, consistency, sufficiency, efficiency, completeness, uniqueness (Without Proofs)

SECTION-B

Applications of Sampling Distributions: Test of mean and variance in the normal distribution,
Tests of single proportion and equality of two proportions, Chi-square test, t-test, F-test.

SECTION-C

Statistical Hypothesis: Null hypothesis, Alternate hypothesis, Level of Significance, simple and composite hypothesis Steps in solving Testing of hypothesis problem, Neyman Pearson Lemma.

SECTION-D

Introduction to ANOVA (Analysis of variance), One way Analysis of variance, Two way Analysis of variance. Problem based on ANOVA.

Text/References:

1. Hogg R.V., Mckean, J.W. and Craig A.T.: Introduction to Mathematical Statistics
2. Gupta S.C. and Kapoor V.K.: Fundamentals of mathematical statistics
3. Goon,A.M.,Gupta M.K. & Dasgupta B. : Fundamental of statistic, Vol. I
4. Goon,A.M.,Gupta M.K. & Dasgupta B. : An outline of statistical theory, Vol. I

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – IV

Paper – III: Algorithm and Heuristics

Time: 3 Hours

Max. Marks: 75

Instructions for the Paper Setters:-

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION-A

Introduction: Introduction to Algorithms and Heuristics, Comparison of Algorithms and Heuristics

Retrieval Strategies: Vector Space Model, Probabilistic Retrieval Strategies, Language Models, Inference Networks, Genetic Algorithms, Fuzzy Set Retrieval

SECTION-B

Retrieval Utilities: Relevance Feedback, Clustering, Passage Based Retrieval, N-Grams, Regression Analysis, Semantic Networks, Parsing

Cross- Language Information Retrieval: Crossing the Language barrier, Cross-Language Retrieval Strategies, Cross Language Utilities

SECTION-C

Efficiency: Inverted Index, Query Processing, Signature Files, Duplicate Document Detection
Integrating Structured Data and Text: Review of the Relational Model, Information retrieval as a Relational Application, Multidimensional Data Model

SECTION-D

Parallel Information Retrieval: Parallel Text Scanning, Parallel Indexing, Clustering and Classification

Distributed Information Retrieval: A Theoretical model of Distributed retrieval, Web Search, Result Fusion, Peer-to-Peer Information Systems

References:

1. Information Retrieval systems : Theory and Implementation by Gerald Kowalski
2. Cross- Language Information Retrieval: by Gregory Grefenstette
3. Text Retrieval and Filtering: Analytics Models of Performance by Robert M. Losee

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – IV

Paper–IV: Big Data

Time: 3 Hours

Max. Marks: 75

Instructions for the Paper Setters:-

Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION–A

Getting an Overview of Big Data: What is Big Data? History of Data Management – Evolution of Big Data, Structuring Big Data, Elements of Big Data, Big Data Analytics, Careers in Big Data, Future of Big Data.

Exploring the Use of Big Data in Business Context: Use of Big Data in Social Networking, Use of Big Data in Preventing Fraudulent Activities, Use of Big Data in Detecting Fraudulent Activities in Insurance Sector, Use of Big Data in Retail Industry.

Introducing Technologies for Handling Big Data: Distributed and Parallel Computing for Big Data, Introducing Hadoop, Cloud Computing and Big Data, In-Memory Computing Technology for Big Data.

SECTION–B

Understanding Hadoop Ecosystem: Hadoop Ecosystem, Hadoop Distributed File System, MapReduce, Hadoop YARN, Introducing HBase, Combining HBase and HDFS, Hive, Pig and Pig Latin, Sqoop, ZooKeeper, Flume, Oozie.

Understanding MapReduce Fundamentals and HBase: The MapReduce Framework, Techniques to Optimize MapReduce Jobs, Uses of MapReduce, Role of HBase in Big Data Processing.

Understanding Big Data Technology Foundations: Exploring the Big Data Stack, Virtualization and Big Data, Virtualization Approaches, Summary, Quick Review

SECTION–C

Storing Data in Databases and Data Warehouses: RDBMS and Big Data, Non-Relational Database, Polyglot Persistence, Integrating Big Data with Traditional Data Warehouses, Big Data Analysis and Data Warehouse, Changing Deployment Models in Big Data Era.

Processing Your Data with MapReduce: Recollecting the Concept of MapReduce Framework, Developing Simple MapReduce Application, Points to Consider while Designing MapReduce.

SECTION–D

Customizing MapReduce Execution and Implementing MapReduce Program:

Controlling MapReduce Execution with Input Format, Reading Data with Custom Record Reader, Organizing Output Data with Output Formats, Customizing Data with Record Writer, Optimizing MapReduce Execution with Combiner, Controlling Reducer Execution with Partitioners, Customizing the MapReduce Execution in Terms of YARN, Implementing a MapReduce Program for Sorting Text Data.

Testing and Debugging MapReduce Applications: Debugging Hadoop MapReduce Locally, Performing Unit Testing for MapReduce Applications, Performing Local Application Testing with Eclipse, Logging for Hadoop Testing, Application Log Processing, Defensive Programming in MapReduce.

References:

1. Michael Minelli, Michelle Chambers, and Ambiga Dhiraj, "Big Data, Big Analytics, "Emerging Business Intelligence and Analytic Trends for Today's Businesses", Wiley.
2. Big-Data Black Book, DT Editorial Services, Wiley India
3. Massive Online Open Courses (MOOCS): Big Data University, Udacity and Coursera.
4. P. J. Sadalage and M. Fowler, "NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence", Addison-Wesley Professional, 2012.
5. Tom White, "Hadoop: The Definitive Guide", Third Edition, O'Reilley, 2012.
6. Eric Sammer, "Hadoop Operations", O'Reilley, 2012.

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – IV

Paper–V: Big Data Analytics using R
(PRACTICAL)

Time: 3 Hours

Max. Marks: 75

Note: Eight Programs of equal marks (Specified In the syllabus) are to be set, two in each of the four sections (A-D). Candidates are requested to attempt four programs, selecting at least one question one question from each section.

SECTION–A

Understanding Analytics and Big Data: Comparing Reporting and Analysis, Types of Analytics, Points to Consider during Analysis, Developing an Analytic Team, Understanding Text Analytics,

Analytical Approaches and Tools to Analyze Data: Analytical Approaches, History of Analytical Tools, Introducing Popular Analytical Tools, Comparing Various Analytical Tools, Installing R, Installing RStudio.

SECTION B

Exploring R: Exploring Basic Features of R, Exploring RGUI, Exploring RStudio, Handling Basic Expressions in R, Variables in R, Working with Vectors, Storing and Calculating Values in R, Creating and Using Objects, Interacting with Users, Handling Data in R Workspace, Executing Scripts, Creating Plots, Accessing Help and Documentation in R.

SECTION C

Reading Datasets and Exporting Data from R: Using the c() Command, Using the scan() Command, Reading Multiple Data Values from Large Files, Reading Data from RStudio, Exporting Data from R.

Manipulating and Processing Data in R: Selecting the Most Appropriate Data Structure, Creating Data Subsets, Merging Datasets in R, Sorting Data, Putting Your Data into Shape, Managing Data in R Using Matrices, Managing Data in R Using Data Frames

SECTION D

Working with Functions and Packages in R: Using Functions Instead of Scripts, Using Arguments in Functions, Built-in Functions in R, Introducing Packages, Working with Packages

Performing Graphical Analysis in R: Using Plots, Saving Graphs to External Files, Advanced Features of R

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – IV

References:

1. Michael Minelli, Michelle Chambers, and Ambiga Dhiraj, "Big Data, Big Analytics, “Emerging Business Intelligence and Analytic Trends for Today’s Businesses”, Wiley.
2. Big-Data Black Book, DT Editorial Services, Wiley India
3. Massive Online Open Courses (MOOCS): Big Data University, Udacity and Coursera.
4. Tom White, "Hadoop: The Definitive Guide", Third Edition, O'Reilley, 2012.
5. Eric Sammer, "Hadoop Operations", O'Reilley, 2012.

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – IV

**PAPER-VI: PROGRAMMING LAB BASED ON NUMERICAL ANALYSIS
(PRACTICAL)**

Time: 3 Hours

M. Marks: 75

Note :- Eight Programs are to be set, two in each of the four Sections (A-D). Candidates are required to attempt four programs, selecting at least one question from each Section.

SECTION–A

Solution of Non–linear Equations: Bisection Method, False position method, Secant Method, Newton Raphson Method.

SECTION–B

Solution of System of Linear Equations: Gauss Elimination Method, Gauss Jordan Method, Gauss Seidel method

SECTION–C

Numerical Integration: Trapezoidal Rule, Simpson’s 1/3 Rule, Simpson’s 3/8 Rule.
Numerical Differentiaton: Function tabulated at equal intervals, Function tabulated at unequal intervals.

SECTION-D

Interpolation: Lagrangian Interpolation, Newton’s Methods: Forward Difference Method, Backward Difference Method, Divided Difference Method.
Curve Fitting: Method of Least square, Fitting Straight line, Fitting a polynomial

Books Recommended:

1. V. Rajaraman: Computer Oriented Numerical Methods, Prentice Hall of India Private Ltd., New Delhi.
2. B.S. Grewal, Numerical Methods for Engineering, Sultan Chand Publication

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – IV

PAPER VII : ESL-221 : Environmental Studies (Compulsory Paper)

Time: 3 Hrs.

Max. Marks: 100

Teaching Methodologies

The Core Module Syllabus for Environmental Studies includes class room teaching and field work. The syllabus is divided into 8 Units [Unit-1 to Unit-VII] covering 45 lectures + 5 hours for field work [Unit-VIII]. The first 7 Units will cover 45 lectures which are class room based to enhance knowledge skills and attitude to environment. Unit-VIII comprises of 5 hours field work to be submitted by each candidate to the Teacher in-charge for evaluation latest by 15 December, 2020.

Exam Pattern: **End Semester Examination- 75 marks**
 Project Report/Field Study- 25 marks [based on submitted report]
 Total Marks- 100

The structure of the question paper being:

Part-A, Short answer pattern with inbuilt choice – 25 marks

Attempt any five questions out of seven distributed equally from Unit-1 to Unit-VII.
Each question carries 5 marks. Answer to each question should not exceed 2 pages.

Part-B, Essay type with inbuilt choice – 50 marks

Attempt any five questions out of eight distributed equally from Unit-1 to Unit-VII. Each question carries 10 marks. Answer to each question should not exceed 5 pages.

Project Report / Internal Assessment:

Part-C, Field work – 25 marks [Field work equal to 5 lecture hours]

The candidate will submit a hand written field work report showing photographs, sketches, observations, perspective of any topic related to Environment or Ecosystem. The exhaustive list for project report/area of study are given just for reference:

1. Visit to a local area to document environmental assets: River / Forest/ Grassland / Hill / Mountain / Water body / Pond / Lake / Solid Waste Disposal / Water Treatment Plant / Wastewater Treatment Facility etc.
2. Visit to a local polluted site – Urban / Rural / Industrial / Agricultural
3. Study of common plants, insects, birds
4. Study of tree in your areas with their botanical names and soil types
5. Study of birds and their nesting habits
6. Study of local pond in terms of wastewater inflow and water quality
7. Study of industrial units in your area. Name of industry, type of industry, Size (Large, Medium or small scale)
8. Study of common disease in the village and basic data from community health centre
9. Adopt any five young plants and photograph its growth
10. Analyze the Total dissolved solids of ground water samples in your area.
11. Study of Particulate Matter (PM_{2.5} or PM₁₀) data from Sameer website. Download from Play store.

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – IV

12. Perspective on any field on Environmental Studies with secondary data taken from Central Pollution Control Board, State Pollution Control Board, State Science & Technology Council etc.

Unit-I

The multidisciplinary nature of environmental studies

Definition, scope and importance, Need for public awareness

(2 lectures)

Unit-II

Natural Resources: Renewable and non-renewable resources:

Natural resources and associated problems.

- (a) Forest resources: Use and over-exploitation, deforestation, case studies. Timber extraction, mining, dams and their effects on forests and tribal people.
- (b) Water resources: Use and over-utilization of surface and ground water, floods, drought, conflicts over water, dams-benefits and problems.
- (c) Mineral resources: Use and exploitation, environmental effects of extracting and using mineral resources, case studies.
- (d) Food resources: World food problems, changes caused by agriculture and overgrazing, effects of modern agriculture, fertilizer-pesticide problems, water logging, salinity, case studies.
- (e) Energy resources: Growing energy needs, renewable and non-renewable energy sources, use of alternate energy sources, case studies.
- (f) Land resources: Land as a resource, land degradation, man induced landslides, soil erosion and desertification.
 - Role of an individual in conservation of natural resources.
 - Equitable use of resources for sustainable lifestyles.

(8 Lectures)

Unit-III

Ecosystems

- Concept of an ecosystem
- Structure and function of an ecosystem
- Producers, consumers and decomposers
- Energy flow in the ecosystem
- Ecological succession
- Food chains, food webs and ecological pyramids
- Introduction, types, characteristic features, structure and function of the following ecosystem: Forest ecosystem, Grassland ecosystem, Desert ecosystem, Aquatic ecosystems (ponds, streams, lakes, rivers, ocean estuaries)

(6 Lectures)

Unit-IV

Biodiversity and its conservation

- Introduction – Definition: genetic, species and ecosystem diversity
- Biogeographical classification of India
- Value of biodiversity: consumptive use, productive use, social, ethical aesthetic and option values
- Biodiversity at global, national and local levels
- India as a mega-diversity nation
- Hot-spots of biodiversity

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – IV

- Threats to biodiversity: habitat loss, poaching of wildlife, man wildlife conflicts
 - Endangered and endemic species of India
 - Conservation of biodiversity: In-situ and Ex-situ conservation of biodiversity
- (8 Lectures)**

Unit-V

Environmental Pollution

Definition

- Causes, effects and control measures of Air pollution, Water pollution, Soil pollution, Marine pollution, Noise pollution, Thermal pollution, Nuclear pollution
- Solid waste management: Causes, effects and control measures of urban and industrial wastes.
- Role of an individual in prevention of pollution
- Pollution case studies
- Disaster management: floods, earthquake, cyclone and landslides

(8 Lectures)

Unit-VI

Social Issues and the Environment

- From unsustainable to sustainable development
- Urban problems and related to energy
- Water conservation, rain water harvesting, watershed management
- Resettlement and rehabilitation of people; its problems and concerns. Case studies.
- Environmental ethics: Issues and possible solutions
- Climate change, global warming, acid rain, ozone layer depletion, nuclear accidents and holocaust. Case studies.
- Wasteland reclamation
- Consumerism and waste products
- Environmental Protection Act, 1986
- Air (Prevention and Control of Pollution) Act, 1981
- Water (Prevention and control of Pollution) Act, 1974
- Wildlife Protection Act
- Forest Conservation Act
- Issues involved in enforcement of environmental legislation
- Public awareness

(7 Lectures)

Unit-VII

Human Population and the Environment

- Population growth, variation among nations
- Population explosion – Family Welfare Programmes
- Environment and human health
- Human Rights
- Value Education
- HIV / AIDS
- Women and Child Welfare
- Role of Information Technology in Environment and Human Health
- Case Studies

(6 Lectures)

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – IV

Unit-VIII

Field Work

- Visit to a local area to document environmental assets
river/forest/grassland/hill/mountain
- Visit to a local polluted site – Urban / Rural / Industrial / Agricultural
- Study of common plants, insects, birds
- Study of simple ecosystems-pond, river, hill slopes, etc

(Field work equal to 5 lecture hours)

ADVISORY FOR PUSHPA GUJRAL SCIENCE CITY, KAPURTHALA :

The Under Graduate students studying Environmental Studies (Compulsory Paper for All UG College Courses) may be taken to Pushpa Gujral Science City, Kapurthala in lieu of Field study report of 25 marks.

Although students will submit a hand written reports with pictures/ graphs/ tables related to biodiversity, ecology, health, biotechnology, energy, water etc. in about 10 pages to the teacher in-charge.

Above advisory is issued to promote scientific temperament in undergraduate classes and is optional. Further, the report will only be considered if there will be a minimum strength of 25 students along with deputed teacher by the college.

References:

1. Bharucha, E. 2005. Textbook of Environmental Studies, Universities Press, Hyderabad.
2. Down to Earth, Centre for Science and Environment, New Delhi.
3. Heywood, V.H. & Waston, R.T. 1995. Global Biodiversity Assessment, Cambridge House, Delhi.
4. Joseph, K. & Nagendran, R. 2004. Essentials of Environmental Studies, Pearson Education (Singapore) Pte. Ltd., Delhi.
5. Kaushik, A. & Kaushik, C.P. 2004. Perspective in Environmental Studies, New Age International (P) Ltd, New Delhi.
6. Rajagopalan, R. 2011. Environmental Studies from Crisis to Cure. Oxford University Press, New Delhi.
7. Sharma, J. P., Sharma. N.K. & Yadav, N.S. 2005. Comprehensive Environmental Studies, Laxmi Publications, New Delhi.
8. Sharma, P. D. 2009. Ecology and Environment, Rastogi Publications, Meerut.
9. State of India's Environment 2018 by Centre for Sciences and Environment, New Delhi
10. Subramanian, V. 2002. A Text Book in Environmental Sciences, Narosa Publishing House, New Delhi.

49
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – V

Paper I: Regression

Time: 3 Hours

Max. Marks: 75

Instructions for the Paper Setters: - Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION–A

Slope of a line, Various forms of equations of a line: parallel to axis, point-slope form, slope-intercept form, two-point form, intercept form and normal form, general equation of a line.

Derivative introduced as rate of change both as that of distance function and geometrically.

SECTION–B

Correlation Analysis: Introduction, Importance, Karl-Pearson's Coefficient of Correlation, Spearman's Rank Correlation Coefficient

SECTION–C

Simple Regression Analysis; Difference between Correlation and Regression, Lines of Regression, Properties of Correlation and Regression Coefficients (Stress on numerical examples).

SECTION–D

Multiple Linear Regression: Concepts, Estimation and Applications (without derivations) of: Partial and Multiple Correlation. Non-Linear Regression: Quadratic and Exponential; Estimation of Fitting of Various Growth Curves (Modified Exponential, Gompertz and Logistic).

Text/References:

1. Gupta, S.P.: Statistical Methods (1981).
2. Yamane, Taro: Mathematics for Economists.
3. Allen, R.G.D.: Mathematical Analysis for Economists.
4. Croxton, Cowden & Klein: Applied General Statistics (1973).
5. Kapur and Sexena: Mathematical Statistics (1970)
6. Murry, R. Speigal: Theory and Problems of Statistics (1972)

50
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – V

Paper II: Design of Experiment

Time: 3 Hours

Max. Marks: 75

Instructions for the Paper Setters: Eight questions of equal marks are to be set, two in each of the four sections (A-D). Questions may be subdivided into parts (not exceeding four). All the candidates are required to answer **Five** questions in all, selecting at least one from each section. The fifth question may be attempted from any section.

SECTION–A

Data Preparation & Data Screening: Identification of Outliers, Checking Normalcy of data, Checking duplicity.

Data Transformation: Conversion of available variable into some other variables using various statistical operations.

Data Analysis & Interpretation.

SECTION–B

Review of Literature: Objectives; Types; Sources.

SECTION–C

Measurement Concepts & Levels of Measurement: Nominal, Ordinal, Interval and Ratio.

Attitude Measurement: Comparative & Non-Comparative Scaling Techniques.

SECTION–D

Research Designs: Exploratory, Descriptive & Diagnostic and Causal Research Designs. (Basic Designs- After only, After along with Control Group, Before-After with Control Group; Time Series Designs)

Suggested Readings:

1. Anderson, D.R.; Sweeney, D.J. and Williams, T.A., “*Statistics for Business and Economics*”, 2nd Edition, Thompson, New Delhi.
2. Cooper, D.R. and Schindler, P.S., “*Business research Methods*”, 9th Edition, Tata McGraw Hill, New Delhi.
3. Kothari, C.R., “*Research Methodology*”, 2nd Edition, New Age International.

51
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – V

Paper III: Practical Based on Predictive Analysis using R

Time: 3 Hours

Max. Marks: 75

Predictive Analytics, Business analytics and its types, Applications, technology and big data influences, analytical techniques, tools,

Predictive modeling, propensity models, cluster models, presenting and using the results of a predictive model, applications, Fundamental limitations of predictive model based on data fitting, software – open source and commercial, introduction to R

Modeling techniques, statistical modeling, machine learning, empirical bayes method, introduction, point estimation, software, examples using R,

Naves bayes classifier, introduction, probabilistic model, parameter estimation and event models, examples, software, examples using R

Decision tree leaning: General, types, metrics, decision tree advantages, limitations, extensions, software, examples using R

Random forests: history, algorithm, bootstrap aggregating, from bagging to random forest, random sub space method , relationship to nearest neighbors, variable importance, variants, software, examples using R

Multivariate adaptive regression splines: basics, mars model, hinge functions, the model building process, pros and cons, software, examples using R

Ordinary least squares: Linear model, estimation, alternative derivations, finite sample properties, large sample properties, example with real data, software, examples using R

References:

1. Jeffery Strickland, Predictive Analytics Using R
2. Eric Mayor, Learning Predictive Analytics With R
3. James Dmiller, Mastering Predictive Analytics With R

52
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – V

Paper IV: Soft Computing

Time: 3 Hours

Max. Marks: 75

Instructions for the Paper Setters: - Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION-A

Introduction to Evolutionary Computing & Genetic Algorithms. Introduction to Genetic Algorithms, Goals of Optimization, How Genetic Algorithms work, A Simple Genetic Algorithm's

Computer Implementation highlighting Reproduction by Selection Crossover, Mutation

SECTION-B

Advanced GA Techniques Mapping Objective Function to Fitness Form, Fitness scaling, discretization, Different types of Selection and Crossover techniques. A case study of Travelling Salesman Problem using GA Techniques. Introduction to other Evolutionary Techniques: PSO, Simulated Annealing and Ant Colony Optimization

SECTION-C

Basics of Neural Networks: Fundamental concept, Model of an Artificial Neuron, Neural Network Architectures, Various types of Learning and Activation Functions.

SECTION-D

Supervised Learning: Perceptron learning, Linear Separability, Delta Rule or Widrow Hoff Rule, Back Propagation algorithm. Introduction to Fuzzy Logic and Fuzzy Sets, Fuzzy Relations, Fuzzification, Defuzzification.

References:

1. David E. Goldberg, Genetic Algorithms in Search Optimization and Machine Learning, Pearson Education.
2. S. N. Sivanandam, S. N. Deepa, Principles of Soft Computing, Wiley Publications.
3. How to Solve It: Modern Heuristics, by Zbigniew Michalewicz, David B.Fogel, Second Edition Springer Verilag-2004, ISBN- 3-540-22494-7.
4. Gallant Stephen I, Neural Network Learning & Extent Systems, MIT Press, 1993.
5. Aleksander & Morton, Neural Computing, Chapman & Hall,

BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – V

Paper V: Machine Learning

Time: 3 Hours

Max. Marks: 100

Instructions for the Paper Setters: - Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION-A

Introduction: Well-Posed Learning Problems, Designing a Learning System, Perspectives and Issues in Machine Learning.

Concept Learning and the General-to-Specific Ordering: Introduction, a Concept Learning Task, Concept Learning as Search. FIND-S: Finding a Maximally Specific Hypothesis, Version Spaces and the CANDIDATE-ELIMINATION Algorithm, Remarks on Version Spaces and CANDIDATE-ELIMINATION, Inductive Bias

SECTION-B

Decision Tree Learning: Introduction, Decision Tree Representation, Appropriate Problems for Decision Tree Learning, The Basic Decision Tree Learning Algorithm Hypothesis Space Search in Decision Tree Learning, Inductive Bias in Decision Tree Learning, Issues in Decision Tree Learning.

Artificial Neural Networks: Introduction, Neural Network Representations, Appropriate Problems for Neural Network Learning, Perceptrons, Multilayer Networks and the BACKPROPAGATION Algorithm, Remarks on the BACKPROPAGATION Algorithms, An Illustrative Example: Face Recognition.

SECTION-C

Bayesian Learning: Introduction, Bayes Theorem, Bayes Theorem and Concept Learning, Maximum Likelihood and Least-Squared Error Hypotheses, Maximum Likelihood Hypotheses for Predicting Probabilities, Minimum Description Length Principle, Bayes Optimal Classifier, Gibbs Algorithm, Naive Bayes Classifier, An Example: Learning to Classify Text, Bayesian Belief Networks, The EM Algorithm

Instance-Based Learning: Introduction, k-NEAREST NEIGHBOR LEARNING, Locally Weighted Regression, Radial Basis Functions, Case-Based Reasoning, Remarks on Lazy and Eager Learning

SECTION-D

Genetic Algorithms: Motivation, Genetic Algorithms, An Illustrative Example, Hypothesis Space Search, Genetic Programming, Models of Evolution and Learning, Parallelizing Genetic Algorithms

Learning Sets of Rules: Introduction, Sequential Covering Algorithms, Learning Rule Sets, Learning First-Order Rules, Learning Sets of First-Order Rules: FOIL, Induction as inverted Deduction, Inverting Resolution

References:

1. Introduction to Machine Learning, Ethem Alpaydin, 3RD Edition, 2015, PHI Learning Pvt. Ltd- New Delhi
2. Machine Learning, Tom M. Mitchell, McGraw Hill Education; First edition, 2013

55
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – VI

PAPER-I: Supply Chain Management

Time: 3 Hours

Max. Marks: 75

Instructions for the Paper Setters:-

Instructions for the Paper Setters: - Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

The student can use only Non-programmable & Non-storage type Calculator.

SECTION-A

SCM – Definition – Objectives – Evolution – Need – Issues involved in developing SCM Framework – Types. SCM activities – Constituents – Organisation.

SECTION-B

Forecasting- Importance, Different kinds of Forecasting Techniques in Estimating Demands, Methods used to Determine accuracy of forecast, Sourcing and vendor selection.

SECTION-C

Inventory Management, EOQ, Minimum Level and Safety Stock, Reordering level, Maximum level, Rational of Discounts of Bulk Purchase, Uncertainty and Inventory management. Innovations in Supply chain management.

SECTION-D

Procurement and Outsourcing: Outsourcing, Benefits and Risks, Framework for make/ buy decision, E -Procurement, Framework of E -procurement. Dimension of Customer Value, Conformance of Requirement, Product Selection, Price and Brand, Value added Services, Pricing, and Customer Value Measures.

References:

1. Supply Chain Management, Kulkarni Sarika& Sharma Ashok
2. Supply Chain Management concept and cases, Rahul V

56
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – VI

Paper-II: Queueing Theory

Time: 3 Hours

Max. Marks: 75

Instructions for the Paper Setters: - Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

SECTION-A

Queueing Theory: Introduction, Queueing System, elements of queueing system, operating characteristics of a Queueing system, distributions of arrivals, inter arrivals, departure and service times.

SECTION-B

Classification of queueing models, single service queueing model with infinite capacity (M/M/1): (∞ /FIFO), Queueing Models:(M/M/1): (N/FIFO), Generalized Model: Birth-Death Process.

SECTION-C

(M/M/C): (∞ /FIFO), (M/M/C) (N/FIFO), (M/M/R) (KIGD), Power supply model.

SECTION-D

Simulation: Introduction, Need of simulation, methodology of simulation, Simulation of queueing systems

References:

1. Handley, G. : Mathematical Programming
2. Kambo, N.S. : Mathematical Programming
3. Panneerselvam, R. : Operations Research
4. Taha, H.A. : Operations Research
5. Kanti Sawrup, Gupta,; Operations Research P.K. and Manmohan

57
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – VI

Paper III: Operations Research

Time: 3 Hours

Max. Marks: 75

Instructions for the Paper Setters: - Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each section. The fifth question may be attempted from any Section. **The student can use only Non-programmable & Non-storage type Calculator.**

SECTION-A

Basics of Operational Research – Development, Definition Characteristics, Necessity, Scope, Limitation. Linear Programming – Introduction, Application, Formulation of Linear Programming Problem, General Linear Programming Problem.

SECTION-B

Graphical Method of Solution, Simplex method, Big-M Method.

SECTION-C

Transportation Model – Assumption, Formulation and Solution of transportation Models, Trans-shipment Problems, Definition of Assignment Model, Hungarian Method for solution of Assignment Problems, Traveling Salesman Problem.

SECTION-D

Game Theory – Theory of Games, Characteristics of Games, Rules – Look for a pure Strategy, Reduce Game by Dominance, Mixed Strategies (2 x 2 Games, 2 x n Games or m x 2 Games).

References:

1. V.K. Kapoor, “Operation Research”.
2. N.D. Vohra, “Quantitative Techniques in Management”.
3. Narinder K. Sethi, “Operation Management”.
4. Gordon/Pressman, “Quantitative Decision Making for Business”.
5. C.R. Kothari, “Quantitative Methods”.
6. Barry Shore, “Quantitative Method for Research”.

58
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – VI

Paper IV: Artificial Intelligence

Time: 3Hrs.

Max. Marks :75

Instructions for the Paper Setters: - Eight questions of equal marks (Specified in the syllabus) are to be set, two in each of the four Sections (A-D). Questions may be subdivided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each Section. The fifth question may be attempted from any Section.

The student can use only Non-programmable & Non-storage type Calculator

SECTION–A

Introduction: Artificial Intelligence, Its applications & importance. Knowledge Representation: Definition & importance of Knowledge, Approaches & issues, predicate logic, production rules, semantic networks, frames, conceptual graphs, Object Oriented representation.

SECTION–B

State Space & Search: Defining problem as state space search, problem characteristics, Blind search, Heuristic search, hill climbing, Best-first, constraint satisfaction.

Natural Language Processing: Grammars & Languages, Paring techniques, RTN, ATN.

SECTION–C

Expert System: Shells, Architectures, Knowledge Acquisition, Case Study MYCIN. Non – Monotonic & Probabilistic Reasoning: Truth maintenance system, Baysean Networks, Fuzzy logic.

SECTION–D

Pattern Recognition: Introduction, recognition & Classification process, clustering.

References:

1. E. Rich & K. Knight, Artificial Intelligence, Tata McGraw Hill, 1991.
2. D.W. Pattern, Introduction to Artificial Intelligence & Expert Systems.
3. G.F. Luger & W.A. Stubblefield, Artificial Intelligence - Structures & Strategies for Complex Problem solving, Second ed. 1993. Benjamin / Cummings
4. Rebert J. Schalkoff, Artificial Intelligence, An Engineering Approach, McGraw Hill, 1990.

59
BACHELOR OF VOCATION (B.VOC.) DATA SCIENCE
Batch from Year 2020 to Year 2023
SEMESTER – VI

Paper IV: Industrial Training

Max. Marks: 100