# Board of Studies (BOS)

for the

Ph.D. Course Work

in

Computer Science & Engineering

With effective from academic session 2022-23

BOS held on 26/09/2023

Department of Computer Science & Engineering

UNS Institute of Engineering & Technology

VBS Purvanchal University, Jaunpur (U.P.)

26/9/2023 (Prof. D.K. Yadav)

26/9/2023 (Dr. Sanjeev Gangwar)

# Veer Bahadur Singh Purvanchal University, Jaunpur U.P.

Syllabus of Pre-Ph.D. course work as per NEP-2020 guidelines
With effective from **academic session 2022-23**

**Subject prerequisites:** Database Systems, Computer Networks, Compilers, Advanced Algorithms for Graph and Combinatorial Optimization Problems, Advanced Database Theory and Applications, Advanced Image Processing, Advanced Operating Systems, Computer Vision, Cryptography and Network Security, Data Mining, Database Management Systems, Computational Algebra, Information and Coding Theory, Pattern, Recognition and Image Processing, Machine Learning: Theory and Practice, Natural Language Processing.

**Programme outcomes (POs):**

**PO1:** Identify research problems, review the existing literature, and develop a plan to address fundamental questions in the primary area of study.

**PO2:** Develop the ability to solve all aspects of the problem comprehensively, analyze the results to suggest valid solutions.

**PO3:** Apply the theories, methodologies and knowledge to come up acceptable solutions of research problems related to real-life scenarios.

**PO4:** Demonstrate in-depth understanding of the current state of the art in the individual research area, and the ability to exploit new solutions and existing research results in the development of new knowledge and theories in the individual research area.

**PO5:** Design and impart solutions to meet the desired requirements of the public health and safety, and the cultural, societal, and environmental considerations.

**PO6:** Create, select, and apply appropriate techniques, resources, and modern tools to draw substantial conclusions.

**PO7:** Abide by ethical principles and commit to professional ethics, responsibilities and norms of the research practice.

**PO8:** Indulge in life-long learning process to promote research-oriented reasoning and keep on developing innovative products.

**Programme specific outcomes (PSOs):**

**PSO1**: Produce a well-developed research proposal.

**PSO2**: Select an appropriate methodology with which to conduct the research and defend the methodology of their selection.

**PSO3**: Understand the various tasks required to carry out the research.

**PSO4**: Find the resources needed to perform the research process.

**PSO5**: Documentation of its findings in the individual research area.

**PSO6**: Understand the most advanced research in the candidate's specialization area of Computer Science respectively

**PSO7**: Understand of academic theory and the preparation of high-quality research pertinent to the field of study

**PSO8**: Appropriately employ methods and existing research results in the development of new knowledge, theories and presentation of research in the individual research are



26/9/2022 (Dr. Sanjeev Gangwar)

26/09/2022 (Prof. D.K. Yadav)

List of all papers of Pre-Ph.D. course work or Post graduate diploma in Research (PGDR)

| Year | Sem | Course Code | Course Title | Theory/ Research | Credit | Max. Marks |
|---|---|---|---|---|---|---|
| 6 | XI | | Machine Learning | Theory | 6 | 100 [25(CIE)+75(UE)] |
| | | | Data Mining and Data Warehousing | Theory | 6 | 100 [25(CIE)+75(UE)] |
| | | | Research Methodology, Research Publication Ethics and Computer Applications | Theory | 4 | 100 [25(CIE)+75(UE)] |
| | | | Research Project | **Research** | - | 100 [25(CIE)+75(UE)] |

**Credit system:**

- A four (4) credit theory course/paper will have four Lectures/periods (of one hour) in a week. In one full semester the course will be covered in 60 Lectures.
- Similarly, a six (6) credit theory course/paper will have six Lectures/periods (of one hour) in a week. In one full semester the course will be covered in 90 Lectures.

**Continuous Internal Evaluation (CIE) of 25 marks:**

- Continuous internal evaluation will be performed by the teacher/ course coordinator concerned.
- CIE shall be 25% of total assessment in a Theory paper and research project.
- 25 marks shall be distributed as 5 marks for attendance, 5 marks for presentation and assignment and remaining 15 marks for class test.

**Marking system:**

- All papers will have a total maximum mark of 100, including both CIE and University Examination (UE).Maximum marks of 25 will be allotted to CIE and 75 to UE in a theory paper/ research project.
- The CIE of the research project shall be evaluated by the research supervisor and co-supervisor (if any).
- 75 marks of **research project** shall be distributed as 50 marks (project work and presentation) and a viva voce of 25 marks.
- The evaluation (Max Marks 75 UE) of the research project shall be done by internal examiner/s (Supervisor and Co-supervisor (if any)) and one external examiner appointed by the University.

**Research Project Submission**:

- The evaluated research project report in two sets of hard copy (spiral binding) must be prepared. One copy of it shall be submitted to the university if it demands. A second copy of the evaluated research project report must be in the records of the college/research centre.
- The format of university Ph.D. thesis writing guidelines can be used as format of Research project writing guidelines.

2

26/9/23
(Dr. Sanjeev Gangwar)

Hanumendra
26/09/2023
(Prof. D.K. Yadav)

| Programme: Ph.D. in Computer Science & Engineering | Year: Seven (VII) | Semester: XIII (Course Work) |
|---|---|---|
| **Subject: Computer Science & Engineering** | | |
| **Course Code:** | **Course Title: Machine Learning** | |

**Course Outcomes ( Cos) :**

After completion of the course the student will able to :

CO1: **Understand** the basics of machine learning techniques and their performance.

CO2: **Demonstrate** the machine learning algorithms as supervised learning and unsupervised learning and Apply and analyze the various algorithms of supervised and unsupervised learning.

CO3: **Evaluate** a variety of learning algorithms to data.

CO4: **Design** algorithms to a real-world problem, optimize the models learned and report on the expected accuracy that can be achieved by applying the models.

CO5: **Manipulate** an appreciation for what is involved in learning from data.

| **Credits: 6** | **Core Compulsory** |
|---|---|
| **Max. Marks: 25 (CIE) + 75(UE)** | **Min. Passing marks: 55** |

**Total number of lectures:** Lectures-Tutorial-Practical (6 hours in a week) L-T-P: 6-0-0 **(90 hrs)**

| Unit | Topics | No. of Lecture Hrs. |
|---|---|---|
| I | **Introduction:** Defining learning systems, Goals and applications of machine learning. Aspects of developing a learning system: training data, concept representation, function approximation, supervised learning, unsupervised learning, Reinforcement learning, learning algorithms. | 15 |
| II | **Decision Tree Learning:** Representing concepts as decision trees. Recursive induction of decision trees. Picking the best splitting attribute: entropy and information gain. Searching for simple trees and computational complexity. Overfitting, noisy data, and pruning | 15 |
| III | **Ensemble Learning:** Bagging, boosting and Ada-Boost. Experimental Evaluation of Learning Algorithms, Measuring the accuracy of learned hypotheses. Comparing learning algorithms: cross-validation, learning curves and statistical hypothesis testing. | 15 |
| IV | **Rule Learning:** Translating decision trees into rules. Artificial Neural Networks: Neurons and biological motivation. Linear threshold units. Perceptrons: representational limitation and gradient descent training. Multilayer networks and back propagation. Hidden layers and constructing intermediate, distributed representations. Overfitting, learning network structure, recurrent networks | 15 |



26/9/2023
(Dr. Sanjeev Gangwar)

26/09/2023
(Prof. D.K. Yadav)

| | | |
|---|---|---|
| V | **Support Vector Machines**: Maximum margin linear separators. Kernels for learning non-linear functions. Bayesian Learning: theory and Bayes rule. Naive Bayes learning algorithm. Parameter smoothing. Generative vs. discriminative training. Logisitic regression. Bayes nets and Markov nets for representing dependencies. Instance-Based Learning: Constructing explicit generalizations versus comparing to past specific examples. k-Nearest-neighbor algorithm, Case based learning. | 20 |
| VI | Recent trends in various learning techniques of machine learning and classification methods for IOT applications. Various models for IOT applications | 10 |

**Suggested Readings:**


1. David W.Mount, "Bioinformatics, Sequence and Genome Analysis", Cold Spring Harbor Laboratory Press.
2. George J.Klir, BoYuan,"Fuzzy Setsand Fuzzy Logic", PHI
3. Witold Pedrycz and Fernando Gomide. "An Introduction to Fuzzy Sets", PHI
4. Tom M.Mitchell, "Machine Learning", MGH


**Suggestive digital platforms web links:**

- https://emasters.iitk.ac.in/course/masters-in-ai-and-machine-learning

Suggested Continuous Evaluation Methods: Presentation, Sessional Tests, Quiz, and Assignments.
Course prerequisites: Basic concepts in Artificial Intelligence and data mining
Suggested equivalent online courses:

- https://www.cisco.com/c/en/us/training-events/training-certifications.html

26/9/2023
(Prof. D.K. Yadav)

26/9/2023
(Dr. Sanjeev Gangwar)

| Programme: Ph.D. in Computer Science & Engineering. | Year: Seven (VII) | Semester: XIII (Course Work) |
|---|---|---|
| **Subject: Computer Science & Engineering** | | |
| Course Code: | Course Title: **Data Mining and Data Warehousing** | |

**Course Outcomes ( Cos) :**

After completion of the course the student will able to:

CO1: **Understand** the concepts of Data Ware housing and Data Mining Concepts.

CO2: **Demonstrate** or Characterize the kinds of patterns that can be discovered by association rule mining.

CO3: **Evaluate** redundancy and incomplete data from the dataset using data pre-processing methods.

CO4: **Design** interesting patterns from large amounts of data to analyze for predictions and classification.

CO5: **Manupulate** a data mining application for data analysis using various tools.

| Credits: 6 | Core Compulsory |
|---|---|
| Max. Marks: 25 (CIE) + 75(UE) | Min. Passing marks: 55 |

**Total number of lectures:** Lectures-Tutorial-Practical (6 hours in a week) L-T-P: 6-0-0 **(90 hrs)**

| Unit | Topics | No. of Lecture Hrs. |
|---|---|---|
| I | **Introduction:** Fundamentals of data mining, Data Mining Functionalities, Classification of Data Mining systems, Data Mining Task Primitives, Integration of a Data Mining System with a Database or a Data Warehouse System, Major issues in Data Mining. **Data Pre-processing:** Overview, Need for Pre-processing; the Data, Data Cleaning, Data Integration and Transformation, Data Reduction, Data Transformation and Data Discretization by Binning, Discretization by Histogram Analysis, Discretization by cluster, Decision Tree and correlation Analysis, concept Hierarchy generation for Nominal data | 15 |
| II | **Data Warehouse and OLAP Technology for Data Mining:** Data Warehouse, Data Warehouse Architecture, Data Warehouse Implementation, Further Development of Data Cube Technology, From Data Warehousing to Data Mining Data Cube Computation and Data Generalization: Efficient Methods for Data Cube Computation, Further Development of | 15 |



16/09/2023
(Dr. Sanjeev Gangwar)

26/09/2023
(Prof. D.K. Yadav)

| | | |
|---|---|---|
| | Data Cube and OLAP Technology. Demand for Online Analytical Processing, Major Features and Functions, OLAP Models, OLAP Implementation Consideration | |
| III | **Mining Frequent Patterns, Associations and Correlations:** Basic Concepts, Frequent itemset Mining methods: Apriori Algorithm, Generate Association rules from Frequent itemsets, Improving the efficiency of Apriori, A pattern-growth approach for mining frequent itemsets, using frequent itemset using Vertical data format, Mining closed and max. patterns, Pattern Evaluation Methods, Advanced Pattern Mining: A Road map, Pattern mining in Multilevel, Multidimensional space, Constraint Based Frequent Mining, Classification: Basic Concepts, Decision Tree induction, Bayes Classification Method, Rule based Classification, Model evaluation & selection, techniques to improve classification accuracy, Classification Advanced Methods: Bayesian Belief networks, Classification by Back Propagation, Support Vector Method, Classification using frequent Patterns, lazy learners, other classification methods, Cluster Analysis: Basic Concepts & Methods, Cluster Analysis, partitioning methods, Hierarchical Methods, Density based Methods, Grid based Methods, Evaluation of Clustering, Advanced Cluster Analysis: Probabilistic Model based Clustering, Clustering High Dimensional Data, Clustering Graph & Network data, Clustering & Constraints | **20** |
| IV | **Classification and Prediction:** Issues Regarding Classification and Prediction, Classification by Decision Tree Induction, Bayesian Classification, Rule-Based Classification, Classification by Back propagation, Support Vector Machines, Associative Classification, Lazy Learners, Other Classification Methods, Prediction, Accuracy and Error measures, Evaluating the accuracy of a Classifier or a Predictor, Ensemble Methods. | **15** |
| V | **Cluster Analysis Introduction:** Types of Data in Cluster Analysis, A Categorization of Major Clustering Methods, Partitioning Methods, Hierarchical Methods, Density-Based Methods, Grid-Based Methods, Model-Based Clustering Methods, Clustering High-Dimensional Data, Constraint-Based Cluster Analysis, Outlier Analysis - Mining Streams, Time Series and Sequence Data: Mining Data Streams, Mining Time-Series Data, Mining Sequence Patterns in Transactional Databases, Mining Sequence Patterns in Biological Data, Graph Mining, Social Network Analysis and Multi-relational Data Mining: | **15** |



26/9/2023
(Dr. Sanjeev Gangsar)

26/9/2023
(Prof. D.K. Yadav)

| | | |
|---|---|---|
| VI | **Mining Object, Spatial, Multimedia, Text and Web Data**: Multidimensional Analysis and Descriptive Mining of Complex Data Objects, Spatial Data Mining, Multimedia Data Mining, Text Mining, Mining the World Wide Web.<br>**Applications and Trends in Data Mining:** Data Mining Applications, Data Mining System Products and Research Prototypes, Additional Themes on Data Mining and Social Impacts of Data Mining. | **10** |

**Suggested Readings:**


1. Pang-Ning Tan, Michael Steinbach and Vipin Kumar, "Introduction to Data Mining " Pearson education.
2. Sam Aanhory & Dennis Murray, "Data Warehousing in the Real World", Pearson Edn Asia.
3. Paulraj Ponnaiah, "Data Warehousing Fundamentals", Wiley student Edition.
4. Margaret H Dunham, "Data Mining Introductory and advanced topics", Pearson education.
5. Arun K Pujari, "Data Mining Techniques", University Press.


**Suggestive digital platforms web links**

- https://onlinecourses.nptel.ac.in/noc22_cs91/preview
- http://epgp.inflibnet.ac.in/epgpdata/uploads/epgp_content/S000007CS/P001063/M013953/ET/1454415596ET.pdf
- https://www.inflibnet.ac.in/about/database.php

Suggested Continuous Evaluation Methods: Presentation, Sessional Tests, Quiz, and Assignments.

Course prerequisites: Basic concepts about Data structures and Databases.

Suggested equivalent online courses:

- https://www.coursera.org/courses?query=database%20management
- https://www.udemy.com/topic/database-management/

26/9/2023 (Dr. Sanjeev Gangwar)

26/09/2023 (Prof. D.K. Yadav)

| Programme: PH.D. in Computer Science & Engineering | Year: Seven (VII) | Semester: XIII(Course Work) |
|---|---|---|
| **Subject: Computer Science & Engineering** | | |
| Course Code: | Course Title: Research Methodology, Research Publication Ethics and Computer Applications | |

**Course Outcomes ( Cos):**

After completion of the course the student will able to:

**CO1:** With the help of this course, students will be able to decide the research field, topic, design, and pros and cons of research, sampling, and data collection techniques.

**CO2:** The student will be able to understand the research process and acquire the skill of writing research articles.

**CO3:** The course will enable you to execute the best practices, morals, and ethical values in scientific conduct and avoid publication misconduct.

**CO4:** With the help of this course, students will be able to learn about the standards of journals for good-quality publications of their research work.

**CO5:** After this course, the students will be able to learn how to use computers and different application software for manuscript writing.

**CO6:** This course will enable the students to learn about reference management and the maintenance of academic integrity using scientific tools. They will be familiar with the protection of the machines from computer hazards.

| **Credits: 4** | **Core Compulsory** |
|---|---|
| **Max. Marks: 25 (CIE) + 75(UE)** | **Min. Passing marks: 55** |

**Total number of lectures:** Lectures-Tutorial-Practical (4 hours in a week) L-T-P: 4-0-0 **(60 hrs.)**

| Unit | Topics | No. of Lecture Hrs. |
|---|---|---|
| I | **Research Methodology** Definition, and Objectives, Motivation and Significance of Research, Types of Research, Truth and Facts of Research, Similarity and Contrast in Literary Research and Scientific Research, Research and Criticism, Research Problem and Research Design, Sampling Design and Methods of Data Collection. | 12 |
| II | **Research standards:** Layout of the Research Report, Research Process: subject Selection, Outline of the Research, Review of Literature, Material Collection; Testing and Classification, Analysis, Discussion and Conclusions, Precautions in Writing Synopsis/Research Paper/Thesis/Research Report. | 12 |
| III | **Philosophy, Ethics, Scientific Conducts and misconducts** Moral Philosophy, Nature of Moral Judgments and Reactions, | 08 |



26/9/2023

26/09/2023

(Prof. O.K. Yadav)

| | | |
|---|---|---|
| | Publication Ethics, Best Practices/Standards Setting Initiatives and Guidelines: Committee on Publication Ethics (COPE), World Association of Medical Editors (WAME) etc., Intellectual Honesty and Research Integrity: Falsification, Fabrication and Plagiarism (FFP), Open Access Publishing, and Publication Misconduct. | |
| **IV** | **Databases and Research Metrics**<br>Databases: Indexing Databases, Citation Databases: Web of Science, Scopus etc., Research Metrics: Impact Factor of Journal as Per Journal Citation Report, SNIP, SJR, IPP, Cite Score; Metrics: h-Index, g-Index, i-10 Index, and Altimetric. | **08** |
| **V** | **Fundamentals of Computers and application Softwares**<br>Types Of Computers, Computer Peripherals and internal component, Types of Operating Systems, Web Browser, Web Search Engine, Spreadsheet Processing, Presentation (MS Power Points Preparation or Beamer or Libre Office (Optional), Project/Thesis/Report writing, Using MS-Word or LaTeX or Libre Office documentation style Labelling, References Style, Footnotes etc. | **12** |
| **VI** | **Scientific Softwares**<br>Use of Reference Management Software Like Mendeley, Zotero, Reference Manager, Endnote, Authorea Etc. Anti-Plagiarism Software Like Turnitin, iAuthenticate, Urkund, Ebooks and Virtual Library, UGC-Infonet, Computer Hazards and Security | **08** |

**Suggested Readings:**


1. C.R. Kothari, *Research methodology Methods and Techniques*, 4th Edition, New Age International (P) Ltd. Publisher, 2014.
2. W. Creswell, *Research Design, Qualitative, Quantitative and mixed method approaches*, 3rd Edition, Sage Publications, Inc.
3. D.B. Resnik, (2011) What is ethics in research & Why is it important. National institute of Environmental Health Science, 1-10 Retrieved from https://www.niehs.nih.gov/research/resources/bioethics/whatis/index.cfm
4. Indian National Science Academy (INSA), Ethics in Science Education, Research and Governance (2019), ISBN:978-81-939482-1-7. https://www.insaindia.res.in/pdf/Ethics_Book.pdf
5. Reema Thareja (2019) Fundamentals Of Computers (2nd Edition), Oxford University Press
6. Microsoft Office 365 : A complete Guide to Master Word, Excel, and PowerPoint 365 for Beginners, Matt Vic
7. Leslie Lamport, LaTeX, A Document Preparation System, 2nd Edition, Addison-Wesley Professional Publisher, July, 1994.
8. Latex tutorials https://www.tug.org/twg/mactex/tutorials/ltxprimer-1.0.pdf
9. Libre Office tutorial: www.documentation.libreoffice.org/en/english-documentation


Suggested equivalent online courses: https://epgp.inflibnet.ac.in/



26/9/2023

(Prof. D.K. Yadav)

26/9/23

(Dr. Sanjeev Gargoos)

| Programme: Ph.D. in Computer Science & Engineering | Year: Seven (7) | Semester: XIII |
|---|---|---|
| **Subject: Computer Science & Engineering** | | |
| **Course Code:** | **Course Title: Research Project** | |
| **Course Outcomes (Cos):**<br>After completion of the course the student will able to:<br>1. Demonstrate a sound technical knowledge of their selected research area.<br>2. Undertake problem identification, formulation and solution.<br>3. Design engineering solutions to complex problems utilising a systems approach.<br>4. Conduct a research.<br>5. Communicate with related community at large in written an oral forms. | | |
| **Credits: Non -Credit** | **Core Compulsory** | |
| **Max. Marks: 25 (CIE) + 75(UE)** | **Min. Passing marks: 55** | |

Suggested Readings:


1. Sally Fincher, "Computer Science Education Research", Taylor & Francis, Netherlands 2004.
2. Holder, Allen, and Eichholz, Joseph, "An Introduction to Computational Science", Springer International Publishing, Germany, 2019.


**The format of the question paper and evaluation will be as follows –**

The duration of each question paper is 3 hours

| Types of Question | Total No of Questions | Questions to be Attempted | Maximum Marks = 75 (UE) (Questions x marks) |
|---|---|---|---|
| Very Short Answer Type Questions (50 words) | 10 | 10 | 10 x 2 = 20 |
| Short Answer Type (200 words) | 8 | 5 | 5 x 7 = 35 |
| Longs Answer Type (500 words) | 4 | 2 | 2 x 10 = 20 |
| | | | = 75 (Maximum Marks) |



Approved
26/09/2023
(Prof. O.K. Yadav)

26/9/2023
(Dr. Sanjeev Gargava)