INDIAN INSTITUTE OF TECHNOLOGY MANDI

KAMAND, DISTT. MANDI – 175075 (HIMACHAL PRADESH)

Indian Institute of Technology Mandi

MINUTES OF 55TH BOARD OF ACADEMICS MEETING

VENUE	:	GUEST HOUSE (NC)CONFERENCE ROOM + ONLINE
DATE	:	31st May, 2024 (FRIDAY)
TIME	:	10:00 AM

Following members attended the meeting

1.	Dean Academics	Prof. Aniruddha Chakraborty
2.	Associate Dean (Research)	Dr. Amit Jaiswal
3.	Associate Dean (Courses)	Dr. P Anil Kishan
4.	Nominee-1: School of Computing and Electrical Engineering	Dr. Padmanabhan Rajan
5.	Nominee-2: School of Computing and Electrical Engineering	Dr. Gopi Shrikanth Reddy
6.	Nominee-1: School of Biosciences and Bioengineering	Dr. Sumit Murab
7.	Nominee-2: School of Biosciences and Bioengineering	Dr. Kasturi Prasad
8.	Nominee-1: School of Civil and Environmental Engineering	Dr. Maheshreddy Gade
9.	Nominee-2: School of Civil and Environmental Engineering	Dr. Prasanna Rousseau
10.	Nominee-1: Centre of Human Computer Interaction (HCI)	Dr. Dinesh Singh
11.	Nominee-2: Centre of Human Computer Interaction (HCI)	Dr. Deepak Sachan
12.	CnP Advisor	Dr. Kala Venkata Uday
13.	Deputy Registrar (Academics): Secretary	Mr. Suresh Rohilla

Following members could not attend the meeting

Sl. No.		Name	
1.	Chairman Senate Library Committee	Prof. Anjan Kumar Swain	Member
2.	Nominee-1: School of Physical Sciences	Dr. Arko Roy	Member
3.	Nominee-2: School of Physical Sciences	Dr. Gargee Sharma	Member
4.	Nominee-1: School of Chemical Sciences	Dr. Bhaskar Mondal	Member
5.	Nominee-2: School of Chemical Sciences	Dr. Garima Agrawal	Member
6.	Nominee-1: School of Mathematical and Statistical Sciences	Dr. Rajendra K Ray	Member
7.	Nominee-2: School of Mathematical and Statistical Sciences	Dr. Syed Abbas	Member
8.	Nominee-1: School of Mechanical and Materials Engineering	Dr. Gaurav Bhutani	Member
9.	Nominee-2: School of Mechanical and Materials Engineering	Dr. Sudhir Pandey	Member
10.	Nominee-1: School of Humanities & Social Sciences	Dr. Rajeshwari Dutt	Member
11.	Nominee-2: School of Humanities & Social Sciences	Dr. Ramna Thakur	Member
12.	Nominee-1: Centre of AI and Robotics	Dr. Narendra Dhar	Member
13.	Nominee-2: Centre of AI and Robotics	Dr. Jagadeesh	Member
14.	Nominee-1: IKSHMA	Dr. Aniruddha Chakraborty	Member

15.	Nominee-2: IKSHMA	Dr. Sumit Murab	Member
16.	Industry Member – 1	Dr. Nadeem Akhtar	Member
17.	Academic Affairs Secretary	Ms. Dishti Oberai	Member
18.	PG Academic Secretary	Mr. Sudama Babu Singhal	Member
19.	Research Affairs Secretary	Mr. Saurabh Patel	Member

Special Invitee

Sl. No.	**Name**	
1	Dr. Akhaya Nayak	Assoc. Prof., SoM
2	Dr. Shyamasree Dassgupta	Assoc. Prof., SHSS
3	Dr. Sandip Kumar Saha	Assoc. Prof., SCENE

PART-A

(Issues discussed by the Board of Academics when the Student Members were present)

55.1 Confirmation of the minutes of 54th meeting of the Board of Academics:

The minutes of the 54th meeting of the Board of Academics held on 14th May, 2024 were confirmed.

55.2 To consider revision/addition in the minimum eligibility criteria for Ph.D/M.Tech (Research) admissions:

Nominee SCEE presented the proposal revision /addition in the minimum eligibility criteria for Ph.D/M.Tech (Research) Admissions. After due deliberations, the BoA suggested to constitute a committee for the presented proposal and submit its recommendations. The same has been constituted and notified vide Notification No. IIT Mandi/Academics/2024/8100-04 dated 6th June, 2024.

55.3 To consider revision/addition in the minimum eligibility criteria for Ph.D admissions:

Nominee, SoM presented the proposal for revision/addition in the minimum eligibility criteria for Ph.D/M.Tech(Research) admissions. After due deliberations, the BoA suggested to constitute a committee for the presented proposal and submit its recommendations. The same has been constituted and notified vide Notification No. IIT Mandi/Academics/2024/8100-04 dated 6th June, 2024.

55.4 To consider Yearly progress review for PhD student with progress report submission:

Nominee, SCEE presented the proposal of reintroduction of Annual progress Review of the Ph.D. students while submitting Progress reports. After due deliberations, the BoA suggested to constitute a committee for the presented proposal and submit its recommendations. The same has been constituted and notified vide Notification No. IIT Mandi/Academics/2024/8105 dated 6th June, 2024.

55.5 To consider proposal for nominations of examiners for Ph.D. Thesis:

Nominee, SCEE presented the proposal for revision of norms for nominations of examiners for Ph.D Thesis examinations. After due deliberations BoA recommended the following for the consideration of Senate and approval:

"The thesis shall be referred to two examiners, either from India or abroad", chosen by the Chairman, Senate or their nominee from the panel of examiners recommended by the Doctoral Committee at its synopsis meeting.

55.6 To consider the guidelines for transfer of Research Scholars from IIT Mandi to another when Guide leaves the Institute:

A Three members committee was constituted with the approval of the Dean (Academics) to frame guidelines regarding transfer of Research Scholars (Ph.D./M.Tech (Research) from IIT Mandi to another Institute when their Guide leaves the Institute to join other Institute or any other reason. Associate Dean (Research) presented the proposal for recommendations of the BoA, After due deliberations, the BoA recommended that the recommendations in these cases will be based on case to case basis.

55.7 To consider proposal for revision of criteria of admission under Project to Ph.D./M.Tech(Research)/MA(Research)/MS(Research) Mode on the basis of 06 Months experience in project:

Associate Dean(Research) presented the proposal for revision of criteria for admission under Project to Ph.D./M.Tech(Research)/MA(Research)/MS(Research) Mode on the basis of 06 Months experience in project. After due deliberations BoA recommended the following for the consideration of Senate and approval:

- If a candidate has applied through an open advertisement and is eligible for Ph.D./M.Tech(Research) as per the general eligibility criteria of IIT Mandi, the 6 months working criteria in the project is not required. The selection procedure will be same as for regular programme. For such candidates the norms of the funding agency will be applicable for the fellowship.
- A candidate meeting the minimum eligibility criteria for Ph.D./M.Tech(Research) can submit application for Ph.D./M.Tech(Research) at any time after completion of the 6 months duration in the project. They may be considered for admission to the Ph.D./M.Tech (Research) /MS(Research)/MA(Research) programme (guided by the same PI only) with strong support of the candidature by the PI. For such candidates the norms of the funding agency will be applicable for the fellowship.

55.8 To consider proposal for revision of norms for Course work requirements for Ph.D. Programme:

Associate Dean (Research) presented the proposal for revision of norms for Course work requirements for Ph.D. Programme. After due deliberations BoA recommended the following for the consideration of Senate and approval:

The students joining for the Ph.D. Programme on the basis of the B.Tech. Degree needs to complete 12 credits+ Research Methodology. A student admitted on the basis of Master's Degree will have to complete 12 credits+ Research Methodology as per existing norms.

Further, students joining for PhD may opt to convert their PhD Degree to MTech(R)+PhD dual degree within first semester of their joining. If approved the norms of MTech(R)+PhD dual degree will be applicable from the date of conversion. The student has to complete 24 credits +Research Methodology in that case. The date of registration will be same as that of the date of joining for the PhD programme. The fellowship will be paid for maximum period of 5 Years from the Initial date of registration in the programme/Institute.

55.9 To consider proposal for revision of format of Provisional Degree certificate:

Associate Dean (Courses) presented the proposal for revision of Provisional Degree Certificate for all programmes. After due deliberations, the BoA recommended the proposal for consideration of the Senate and approval.

The revised format of Provisional Degree certificate is placed as **Annexure – A.**

55.10 To consider the new courses from SOM:

Nominee SoM presented the following courses form SoM. BoA approved the courses and the same shall be reported to the Senate:

Sl.No.	Course No.	Course Name	Credits

1.	MB554	Blockchain for Business	2-0-0-2
2.	MB509	Introduction to Bhagavad Gita	2-0-0-2

The course descriptions are placed as **Annexure- B.**

55.11 Any other item with the permission of the Chair:

(1) To report minor modification in the course from SCEE:

In 54th BoA Digital Signal Processing Course Code was approved with EE-571 code. As this is a core course for EE students and as per the revised EE curriculum, intended for 2nd and 3rd year students, therefore course code is changed from EE-571 to EE-314. BoA reported the same as follows and the same shall be reported to the Senate

Sl.No.	Course No.	Course Name	Credits	Remarks
1.	EE-314	Digital Signal Processing	3-0-2-4	(Course Code changed from EE-571 to EE-314)

The course descriptions are placed as **Annexure- C**.

(2) To Report credit limit for B.Tech. 2023 Students in current semester.

BoA reported that the B.Tech. 2023 students have been permitted to credit more than 22 credits (permissible) during the current Semester (2023-24 Even Semester). Maximum 25 credits are permitted to the students of B.Tech. 2023 for one time. The same will be reported to the Senate.

(3) To Report revision of credit structures for Courses from SCENE

BoA reported the credit structures of following courses from SCENE:

Sl.No.	Course No.	Course Name	Old Credits	New Credits
1.	CE-688P	Post Graduate Project-I	0-0-24-12	0-0-28-14
2.	CE-689P	Post Graduate Project-II	0-0-34-17	0-0-30-15

(4) To consider approval for courses from SCEE for Data Science & Engineering:

Nominee SCEE presented the following courses form SCEE for Data Science Engineering Students. BoA approved the courses and the same shall be reported to the Senate:

Sl.No.	Course No.	Course Name	Credits	Remarks
1.	IC-272	Machine Learning	2-0-2-3	
2.	DS-201	Data handling and visualization	2-0-2-3	DSE 2022 Batch and earlier batches
3.	DS301	Mathematical Foundations of Data Science	3-1-0-4	DSE 2022 Batch and earlier batches
4.	DS302	Computing Systems for Data Processing	2-0-2-3	DSE 2022 Batch and earlier batches
5.	DS303	Statistical Foundations of Data Science	3-0-0-3	DSE Earlier batches 2022
6.	DS313	Statistical Foundations of Data Science	3-1-0-4	DSE 2022 Batch
7.	DS401	Optimization for Data Science	3-0-0-3	DSE Earlier batches 2022
8.	DS402	Matrix Computations for Data Science	2-0-2-3	DSE Earlier batches 2022
9.	DS403	Introduction to Statistical Learning	2-0-2-3	DSE Earlier batches 2022
10.	DS404	Information Security and Privacy	3-0-0-3	DSE Earlier batches 2022

11.	DS411	Optimization for Data Science	3-1-0-4	DSE 2022 Batch
12.	DS412	Matrix Computations for Data Science	3-0-2-4	DSE 2022 Batch
13.	DS413	Introduction to Statistical Learning	3-1-0-4	DSE 2022 Batch

The course descriptions are placed as **Annexure- D**

(5) To consider course numbers for Courses from SBB:

Nominee SBB presented course numbers for the following project courses for the IDD Bioengineering form SBB, the BoA approved the courses and the same shall be reported to the Senate:

Sl.No.	Course No.	Course Name	Credits
1.	BE-609P	M.Tech Project 1	0-0-34-17
2.	BE-610P	M.Tech Project 2	0-0-34-17

-NIL-

PART-B

(Issues discussed by the BoA without the Student Members being present)

-NIL-

The meeting ended with a vote of thanks to and by the Chair.

27/6/2024

Chairman, Board of Academics

Secretary, Board of Academics

Anenxure A

Indian Institute of Technology Mandi

Indian Institute of Technology Mandi

No.: IIT Mandi/Academics/B.Tech/Roll No/Year/ DS-______
Date: ____________

PROVISIONAL DEGREE CERTIFICATE

(FOR THE DEGREE OF BACHELOR OF TECHNOLOGY (B.Tech.))

This is to certify that **Mr./Ms. ABC** (Roll No. B0000) has successfully completed all the requirements for the award of the **Bachelor of Technology (B.Tech.) in** ___________of this Institute on _______. The student has secured a CGPA of _______ out of the maximum CGPA of 10.00.

This certificate is issued with approval of Competent Authorities. The original degree shall be conferred during the next Convocation of the Institute. This Certificate can be treated as Degree Certificate till the award of original Degree.

Deputy Registrar
(Academics)

Annexure B

Proposal for New Course		
Course Number	:	**MB554**
Course Name	:	Blockchain for Business
Credits	:	2-0-0-2 (L-T-P-C)[1]
Prerequisites	:	None
Intended for	:	MBA
Distribution	:	Elective
Semester	:	Odd/Even

Preamble

Blockchain as an advent technology is playing a crucial role in the mainstream business applications. Because blockchain has the potential to drive the re-imagination of processes and business models in a distributed and decentralized manner, it can be a transformational technology for many businesses. Many business and technology leaders, however, might overlook its potential usage and value for their business and industry, or associate it primarily with Bitcoin and cryptocurrency applications. This course is going to address the nuances of blockchain and their applications in business. This course will focus on demystifying blockchain concepts, use cases and implementation issues to be faced by the business professionals. This course focuses on the managerial aspects the blockchain to arrive at most appropriate decision.

Objective

On completion of this course, the student should be able to:

- Identify which industry and company will be affected and to what extent.
- Understand how blockchain helps making business decentralized and secure.
- Visualize the concepts of cryptocurrencies as business use cases.
- Acquire the knowledge to discover opportunities in business to drive it through innovation and efficiency.
- Realize the current capabilities and future possibilities of blockchian applicability in business.

[1] L= Lectures per week, T=Tutorials per week – P = Practical/Lab session per week – C = Credits for course

Course Modules with Quantitative lecture hours		
Module 1	**Introduction to Blockchain**	**(4)**
Basic ideas behind blockchain, its purpose, how it is changing the landscape of Business, Enterprise Blockchain, Why Blockchain matters? Public vs Private vs Permissioned Blockchains and use cases, Blockchain Questions from Business and Technology Leaders, Blockchain as a Service.		
Module 2	**Cryptographic Concepts Required**	**(4)**
Confidentiality, Integrity, Authentication, Mathematics, Symmetric key Cryptography-Stream and block ciphers, Asymmetric key Cryptography, Discrete logarithm, RSA, Hash function, Digital Signature - ECDSA.		
Module 3	**Distributed Consensus**	**(4)**
The real need for mining – consensus – Nakamoto consensus, Proof of Work, Proof of Stake, Proof of Burn, Difficulty Level, Byzantine Generals Problem, Byzantine fault tolerance, Sybil Attack, Energy utilization and alternate, Business applications (appropriate case studies, use cases and situation analysis)		
Module 4	**Cryptocurrencies – business use cases of blockchain technology**	**(8)**
Introduction to Cryptocurrency, what is a Bitcoin? Double Spending Problem, Bitcoin Mining, Transactions, The Bitcoin network, Bitcoin payments, Bitcoin ICO- Advantages and Disadvantages, Ethereum and its basic Features, The Ethereum network, Components of the Ethereum ecosystem, Ethereum Virtual Machine, Gas, Applications Built based on Ethereum, ETH, Smart Contracts, Smart contract templates, DApp, Business applications (appropriate case studies, use cases and situation analysis)		
Module 5	**Introduction to Hyperledger**	**(4)**
Projects under Hyperledger, Permissioned Blockchain and use cases, Hyperledger as a protocol, The reference architecture, Privacy and confidentiality, Scalability, Hyperledger Fabric		
Module 6	**Blockchain Use Cases**	**(4)**

Supply Chain Management, E-Governance, Land Registration, Medical Information Systems, and others

Lab Exercises (If applicable):
Nil.

Textbooks:	
Reference Book:	
1	Narayanan, A., Bonneau, J., Felten, E., Miller, A., & Goldfeder, S. (2016). Bitcoin and cryptocurrency technologies: a comprehensive introduction. Princeton University Press.
2	Arun, J. S., Cuomo, J., & Gaur, N. (2019). Blockchain for business. Addison-Wesley Professional.
3	Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. Decentralized business review.
4	Tyagi, S. S., & Bhatia, S. (Eds.). (2021). Blockchain for Business: How it Works and Creates Value. John Wiley & Sons.
5	Forouzan, B. A., & Mukhopadhyay, D. (2015). Cryptography and network security (Vol. 12). New York, NY, USA: Mc Graw Hill Education (India) Private Limited.
6	Bashir, I. (2017). Mastering blockchain. Packt Publishing Ltd.
7	
8	
9	

IIT Mandi
Proposal for a New Course

Course number : **MB-509**
Course Name : **Introduction to Bhagavad Gita**
Credit : **2**
Distribution : **2-0-0-2**
Intended for : **MBA**
Prerequisite : **NA**
Mutual Exclusion : (Specify the equivalent courses in other schools. These *Courses (with high similarity) are not allowed to credit by the students after or along with this course.*)

---

1. Preamble:

Bhagavad Gita explores at length the five major branches of knowledge such as Individual conscious entity, Supreme conscious entity, Material nature, Time, and Actions. It also contains enriching discussion between Arjuna and Krishna on topics such as dilemmas of life, identity conflict, theory of actions (forbidden action, recommended action, Inactions), importance of controlled mind, qualities that builds a divine personality, bondage and freedom, etc. It is considered as a manual of life which outlines essential principles of leading a fulfilling life. Off late, it has garnered ample attention from management scholars and volumes of scholarly literature in form of journals papers, books, and conference proceedings are published on this subject. It is, therefore, pertinent for the aspiring managers and business leaders to understand the principles of Bhagavad Gita to aid them in decision making. This course is an earnest attempt to make the students aware of the time tasted principles of Bhagavad Gita which will help them deal with adversaries of life and equip them with better decision making power in the interest of themselves, organization, community, society, and other important stakeholders.

2. Course Modules with quantitative lecture hours:

Unit/Topic 1: Bhagavad Gita: The timeless science **(2 Hours)**

Unit/Topic 2: Krishna: As He is **(2 Hours)**

Unit/Topic 3: Description of atomic soul **(3 Hours)**

Unit/Topic 4: Matter and consciousness **(3 Hours)**

Unit/Topic 5: Yoga systems **(2 Hours)**

Unit/Topic 6: Transcendental knowledge **(2 Hours)**

Unit/Topic 7: Mindfulness from Bhagavad Gita **(3 Hours)**

Unit/Topic 8: The topmost yoga system **(2 Hours)**

Unit/Topic 9: Modes of material nature and professional excellence (**3 hours**)

Unit/Topic 10: Perfection of renunciation (**2 hours**)

Laboratory/practical/tutorial Modules:

NA

3. Text books:

(Latest, Only 2)

1. Prabhupada, ACBS, Bhagavad Gita: As it is, Bhaktivedanta Book Trust, Mumbai, 1972
2. Das, L. P., Five Aspects of the Absolute Truth - a Bhagavad Gita Study Guide, India, 2022

4. References:

(No limit on numbers, relevant)
Standard format can be followed

1. Chatterjee, D. (2012). *Timeless leadership: 18 leadership sutras from the Bhagavad Gita*. John Wiley & Sons.
2. Nayak, A. K. (2018). Effective leadership traits from Bhagavad Gita. *International Journal of Indian Culture and Business Management, 16*(1), 1-18.
3. Krishnan, R., Jain, R., & Maheshwari, A. K. (2023). Development of Consciousness-Based Leadership from Bhagavad Gita and Yoga Sutras. In *Consciousness-Based Leadership and Management, Volume 2: Organizational and Cultural Approaches to Oneness and Flourishing* (pp. 97-113). Cham: Springer International Publishing.

5. Similarity with the existing courses:
(Similarity content is declared as per the number of lecture hours on similar topics)

S. No.		Course Code	Similarity Content	Approx. % of Content
1.				

6. Justification of new course proposal if cumulative similarity content is >30%:

Approvals:

Faculty interested in teaching this course: – Dr. Akhaya Kumar Nayak

Proposed by: Dr. Akhaya Kumar Nayak **School: School of Management**

Signature: **Date:**

The following faculty (at least 3 faculty) discussed on..............and approved the proposal on.............

Sl. No	Faculty Name	Signature
1	Dr. Laxmidhar Behera	
2	Dr. Amit Shukla	
3	Akhaya Kumar Nayak	

School Chair:

School:

Date:

This proposal is reported in ....th Board of Academics on ....

Dean Academics

Date:

Note: School is responsible for the Course Code. Academic Office provides the IC Course Code.

Annexure C

IIT Mandi
Proposal for a revision of an existing course

Course Number : EE314

Course Name: Digital Signal Processing

Distribution: 3-0-2-4

Intended for: B.Tech. 2nd, 3rd year students

Prerequisites: EEXXX (Signals & Systems)

Mutual Exclusion: None

Preamble: This is an introductory course in discrete-time signal processing. The primary objective of this course is to introduce the students to time and frequency domain analysis of discrete-time signals and applications of these techniques in designing various linear time-invariant discrete-time systems. This course will also serve as a prerequisite for advanced courses on signal processing, e.g., Image processing, advanced signal processing and speech processing. The theory sessions must be accompanied by the corresponding lab sessions where students will learn how to visualize discrete-time signals in both time and frequency domains as well as the implementations of various DSP algorithms in Python or MATLAB.

Course modules:

Unit 1: Review of discrete-time signals and systems (4 hours)

LTI systems and their properties, impulse response and convolution, Difference equations, Eigen-functions of LTI systems

Unit 2: Fourier and Z-transforms (10 hours)

Discrete-time Fourier transform and its properties; Z-transform and its properties; Discrete Fourier Transform and its properties; Fast Fourier Transform

Unit 3: Sampling (8 hours)

Time domain and frequency domain representation, Nyquist theorem, Signal reconstruction, Discrete-time processing of continuous-time signals, Continuous-time processing of discrete-time signals, Changing the sampling rate, Multi-rate signal processing, Sub-Nyquist sampling and its applications

Unit 4: Filtering (10 hours)

Discrete-time frequency selective filtering, Phase distortion and delay, Characterization with difference equations, Stability and Causality, Frequency response of rational system functions, All pass and minimum phase systems, Basics of filter design, Z-transform characterization of IIR filters, Window functions for FIR filters, Filter structures for IIR and FIR filters,

Unit 5: Introduction to wavelets and their applications (10 hours)

Haar expansions, Wavelets in continuous time, Discrete Wavelet Transform using Haar basis, Construction of wavelets using Fourier techniques, Multiresolution analysis and construction of the wavelet.

Lab: The exercises/projects should cover various real-world applications of DSP. A few suggested topics: image processing and compression; Speech processing; Compressed sensing; Signal processing in communication systems

Text Books:

1. Alan V. Oppenheim, Ronald W. Schafer, John R. Buck., "Discrete-Time Signal Processing," Second edition, Pearson, 1999.
2. Martin Vetterli, Jelena Kovacevic and Vivek Goyal, "Fourier and Wavelet Signal Processing", Cambridge University Press, 2014

Reference Books:

1. John G. Proakis, Dimitris G. Manolakis., "Digital Signal Processing – Principles, Algorithms, and Applications," Fourth Edition, Pearson 2007.
2. Stephen Mallat, A Wavelet Tour of Signal Processing The Sparse Way, Elsevier, 2009

5. Similarity with the existing courses:

(Similarity content is declared as per the number of lecture hours on similar topics)

This is a core course for undergraduate students in electrical engineering stream and also revision of an existing course.

6. Justification of new course proposal if cumulative similarity content is >30%:

Not applicable

Approvals:

Faculty interested in teaching this course: Siddhartha Sarma, Arnav Bhavsar, Sneha Singh

Proposed by: Siddhartha Sarma, Arnav Bhavsar School: SCEE

Signature: Date: 3 May 24

The following faculty (at least 3 faculty) discussed on **2 May 24** and approved the proposal on **3 May 24**.

Sl. No	Faculty Name	Signature
1	Arnav Bhavsar	
2	Siddhartha Sarma	
3	Sneha Singh	

School Chair:

School:

Date:

This proposal is reported in ……………..th Board of Academics on …………..

Annexure D

IIT Mandi
Proposal for a New Course

Course Number	: IC272
Course Name	: Machine Learning
Credits	: 2-0-2-3 (L-T-P-C)
Prerequisites	: IC111 – Linear Algebra, IC152 – Computing and Data Science, IC252 - Probability and Statistics (Data Science 2)
Intended for	: Second year B.Tech
Distribution	: Institute core for B. Tech
Semester	: Even/Odd

Preamble: Data science involves using the scientific methods to process the data to extract knowledge and insights from data in various forms, both structured and unstructured. Vast amounts of data are being generated in many fields, and the data analyst's job is to make sense of it all: to extract important patterns and trends. This involves learning from data. This course intends to provide fundamental ideas and techniques in learning from data.

Objective: After data science - 1 (computation & data science), and data science – 2 (probability and statistics), the students acquire familiarity with statistical and mathematical properties for analyzing the data and computation involved with it. This introductory course aims to provide foundations to different techniques for learning from useful information/patterns and trends from data with practical applications in various fields. This course also aims at providing practical implementation of different techniques used for learning from data and making predictions on the real world data.
Primary objective of this course are:

- Understand various machine learning (ML) algorithms such as supervised and unsupervised algorithms.
- Learn to analyse data to gain insights using an appropriate ML algorithm under a given task and context.
- Learn various learning algorithms for tasks such as regression, classification and clustering
- Analyse & solve problems from different areas and disciplines such as finance, environment, agriculture, networking, manufacturing, sports, bioinformatics, healthcare etc using ML algorithms.
- Learn to use python-based toolbox scikit-learn, a simple and efficient tool for machine learning and data modeling

Course Contents:

1. **Data preprocessing**: Data cleaning – missing values, noisy data; Data integration and transformation – normalization; Data reduction – dimension reduction and principal component analysis (PCA) [5 Lecture]

2. **Introduction to machine learning:** Supervised and unsupervised learning [1 Lecture]

3. **Supervised learning with applications in classification problems**: Bayes classifier with unimodal and multimodal density - maximum likelihood estimation, expectation-maximization (EM) algorithm (only at idea level), K-nearest neighbor methods, decision trees, neural networks [8 Lectures]

4. **Supervised learning - regression**: Linear regression, polynomial regression, regression using neural networks [8 Lectures]

5. **Unsupervised Learning Algorithms - Clustering**: K-means and fuzzy-K-means clustering, density based clustering (DBSCAN) [6 Lectures]

Lab Exercises:

- Lab to be conducted on a 3-hour slot. It will be conducted in tandem with the theory course so the topics for problems given in the lab are already initiated in the theory class. The topics taught in the theory course should be appropriately be sequenced for synchronization with the laboratory.
    - Lab1 : Data Preprocessing – data cleaning and normalization
    - Lab2 : Dimension reduction and PCA
    - Lab3 : Classification using Bayes classifier with unimodal density
    - Lab4 : Classification using Bayes classifier with multimodal density
    - Lab5 : Classification using K-nearest neighbor methods and decision trees
    - Lab6: Classification using neural networks
    - Lab7: Linear regression
    - Lab8: Polynomial regression
    - Lab9: Neural network based regression
    - Lab10: Clustering using K-means and fuzzy-K-means
    - Lab11: Clustering using DBSCAN

Textbooks:

1. C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
2. C. Muller and S. Guido, *Introduction to Machine Learning with Python: A Guide for Data Scientists,* O'Reilly, 2017

Reference books:

1. J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Third Edition, Morgan Kaufmann Publishers, 2011
2. S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, Academic Press, 2009.
3. T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Second Edition, Springer, 12th Reprint, 2017
4. R. O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification*, John Wiley, 2001.

Similarity Content Declaration with Existing Courses:

Introductory institute core meant for undergraduate engineering students.

Justification for new course proposal if cumulative similarity content is > 30%: NA

Approvals:

Other faculty interested in teaching this course: Varun Dutt, Manoj Thakur

Proposed by: A. D. Dileep School: SCEE

Signature: Date: 20 June 2019

Recommended/Not Recommended, with comments

Chairperson, Course proposal committee (CPC) Date:

Approved / Not Approved

Chairperson, Senate Date:

IIT Mandi

Proposal for a New Course

Course Number : DS201
Course Name : Data handling and visualization
Credits : 2-0-2-3 (L-T-P-C)
Prerequisites : None
Intended for : UG
Distribution : Discipline core for B. Tech. Data Science and Engineering, Discipline elective for B. Tech. Computer Science and Engineering, Electrical Engineering, Free elective for other B. Tech. disciplines

Semester : Even/Odd

Preamble: This course will provide undergraduate students with an understanding of the data collection, data representation and data visualization. Furthermore, this course will provide a general framework that allows students to identify errors associated with different data sources. Students will learn some metrics to quantify each potential error, and thus they will be able to describe the quality of a data source. The course will introduce different large-scale data collection efforts undertaken by industry and government agencies, and review the learned concepts through case-studies and examples. This course is suitable for students who are looking for a general framework to use data sources.

Objective: With the help of various examples students will learn how to identify which data sources likely matches research questions, how to turn research questions into measurable pieces, and how to think about an analysis plan.

On completion of this course, the student should be able to

- describe what makes a good or bad visualization
- demonstrate familiarity with various data sources
- Identify the functions that are best for particular problems
- Understand best practices for creating basic charts
- Learn data analysis tools and software

Course modules with Quantitative lecture hours:

1. **Data sources and collection:** **(6 lectures)**

   This module walks you through the process of data collection. Starting with a review of existing structured and unstructured data sources, we cover data collection techniques using sensors, surveys, and different instruments. This includes data collection and storing for different domains such as IoT, Audio and Video, Web and Social Networks etc.

2. **Data Pre-processing:** **(7 lectures)**

Highlight the importance of data correction and discuss some basic features that can affect your data analysis when dealing with sample data. Issues of data access and resources for access are introduced in this module. Descriptive data summarization, data cleaning, normalization, data integration and transformation, data reduction.

3. **Data representation:** **(9 lectures)**

Importance of data representations, Extracting salient features from data, Examples include MFCC from audio signals, histogram representation for text, feature representations for images, encoded representations, Spatial data representation: cartography, GIS paper maps to ArcGIS ArcMap symbolizing, Time-series data representations and curve fitting.

4. **Basic charting and data visualization:** **(6 lectures)**

Basic charting, examples with real world weather data, extract and manipulate the data to display the maximum information, various types of graphs like pie chart, bar graphs, 3-D plots using Matlab and R. Procedure of composite charts by overlaying a scatter plot of record breaking data for a given year, Visualization of high dimensional data e.g. TSNE plot, histogram etc. Also, dynamic data representations and visualization of data using D3.j

Lab Exercises:

- Lab to be conducted on a 2-hour slot. It will be conducted in tandem with the theory course so the topics for problems given in the lab are already initiated in the theory class. The topics taught in the theory course should be appropriately be sequenced for synchronization with the laboratory.

Textbooks:

1. Yau, Nathan. *Visualize this: the Flowing Data guide to design*, visualization, and statistics. John Wiley & Sons, 2011.

2. Tufte, Edward R. *The visual display of quantitative information*. Vol. 2. Cheshire, CT: Graphics press, 2001.

Reference books:

1. Janert, Philipp K. *Data analysis with open source tools: a hands-on guide for programmers and data scientists*. O'Reilly Media, Inc., 2010.

2. Zhu, Xuan. *GIS for environmental applications: a practical approach*. Routledge, 2016.

IIT Mandi

Proposal for a New Course

Course Number : DS301

Course Name : Mathematical Foundations of Data Science

Credit s : 3-1-0-4

Prerequisites : IC110-Engineering Mathematics, IC111-Linear Algebra

Students intended for : UG

Elective or core : Discipline core for B. Tech. Data Science and Engineering, Discipline elective for B. Tech. Computer Science and Engineering, Electrical Engineering, Free elective for other B. Tech. disciplines

Semester : Odd/Even

Preamble: A large amount of work done in the area of data science and engineering span through range of very simple to mathematically elegant methods. A number of topics in mathematical analysis are of key interests to data science researchers. The mathematical analysis components offered in a typical mathematics program are often scattered in many specialised and dense courses. On the other hand mathematics courses offered in a typical computer science curriculum generally do not cover these important concepts that are very crucial for understanding of the core of many areas of machine learning. This introductory course aims to provide a graspable introduction to mathematical analysis required for data science and engineering practitioner.

On completion of this course, students should be able to demonstrate their understanding of and apply various concepts of mathematical analysis in DSE and serves as the input to many concepts in other courses on numerical linear algebra, optimization, and machine learning.

In particular, students should be able to

- acquire the adequate depth in relevant topics in mathematical analysis that adjoins to the knowledge of a data science practitioner.
- understand concepts related to notion of convergence in norm which is important ingredient in many of the techniques such as regression, classification and clustering requiring approximation different kind of functions.
- understand concepts of projection, orthogonality and their properties that are essential ideas for many machine learning techniques.

Course modules with Quantitative lecture hours:

Module 1: Definition of metric spaces, Examples, Open sets, Closed sets, Dense sets, Compact sets, Connectedness, Closure and interior of the sets, Metric subspace.

(10 lectures)

Module 2: Cauchy sequences, Convergent sequences, Complete metric space, , Continuous functions, Continuity of composite functions, Continuity and inverse image of open and closed sets.

(10 lectures)

Module 3: Normed linear spaces, Linear subspaces of normed linear spaces, Banach spaces, Riesz lemma, Continuity of linear maps, Bounded linear maps, Norm equivalence.

(10 lectures)

Module 4: Hilbert spaces, Cauchy -Schwarz inequality, Parallelogram law, Orthogonality, Pythagorean Theorem, Orthogonal projection, orthogonal complement and projection theorem, Orthonormal sets, Orthonormal basis, Gram-Schmidt process, Examples of orthonormal basis.

(12 lectures)

Text Books:

1. Apostol, T., Mathematical Analysis, 2nd ed., Narosa Publishers, 2002.
2. Limaye, B. V., Functional Analysis, 2nd ed., New age international Publishers, 2009.
3. Dan Simovici Mathematical Analysis for Machine Learning and Data Mining, World Scientific, 2018
4. Rudin, W., Principles of Mathematical Analysis, 3rd ed., McGraw-Hill, 2013.

Reference Books:

1. Stein, E. M. and Shakarchi, M., Real Analysis, Princeton Lectures.
2. Tao, T, Analysis I and II, Trim, Hindustan book agency.
3. Kreyszig, E., Introductory Functional Analysis with Applications, Reprint 2017
4. Naylor, A. C. and Sell, G. R., Linear Operator Theory in Engineering and Science.

IIT Mandi

Proposal for a New Course

Course Number : DS302
Course Name : Computing Systems for Data Processing
Credits : 2-0-2-3 (L-T-P-C)
Prerequisites : None
Intended for : UG
Distribution : Discipline core for B. Tech. Data Science and Engineering, Not allowed to B.Tech. CSE, Discipline elective for B. Tech. Electrical Engineering, Free elective for other B. Tech. disciplines.

Semester : Even/Odd

Preamble:

One of the important problems in data science is how to design computing systems that can deal with various sources of data, and can scale to huge volumes of data. This course is designed to give basic understanding of storage and computing systems at the logical and physical level of abstraction. Relevant topics from Operating Systems, Computer Organization and Databases are fused in this course to fulfil the above course objective. On completion of the course, students should be able to understand how data is represented in the computer, the different memory hierarchies and their performance, trade-offs between using processes and threads for applications, concepts of scheduling and different data models like Relational and NoSQL.

Course Modules with Quantitative lecture hours:

Basics of Computer Organization **(8 lectures)**

Data representation, machine code, computer arithmetic, code compilation, memory organization and management, memory and run time performance optimization

Introduction to Operating Systems **(8 lectures)**

Responsibilities of OS; process & thread management: process model, states and its structure, process creation and termination, thread models and issues, User/Kernel level threads; inter-process communication, process synchronization, and process scheduling; file systems.

Introduction to Databases **(8 lectures)**

Information modeling, ACID properties and Transactions, Trade-offs between Relational Databases and NoSQL, Data Manipulation Language SQL

Scalable data processing using MapReduce **(4 lectures)**

Hadoop Distributed File System, MapReduce programming model

Text book:

1. Stallings, William. Computer *Organization and Architecture*, Global Edition. Pearson Education Limited, 2015.
2. Silberschatz, Abraham, Greg Gagne, and Peter B. Galvin. *Operating system concepts*. Wiley, 2018.
3. Elmasri, Ramez, and Sham Navathe. *Fundamentals of database systems*. Pearson, 2017.

Reference Books:

1. Forouzan, Behrouz, Catherine Coombs, and Sophia Chung Fegan. *Introduction to data communications and networking*. McGraw-Hill, Inc., 1997.
2. Ramakrishnan, Raghu, and Johannes Gehrke. *Database management systems*. McGraw Hill, 2000.
3. Carl Hamacher, V., Zvonko G. Vranesic, and Safwat G. Zaky. *Computer organization*. 2004.

IIT Mandi

Proposal for New Course

Course Number : DS303

Course Name : Statistical Foundations of Data Science

Credits : 3-0-0-3 (L-T-P-C)

Prerequisites : IC110-Engineering Mathematics, IC252-Data Science 2

Intended for : UG

Distribution : Discipline core for B. Tech. Data Science and Engineering, Discipline elective for B. Tech. Computer Science and Engineering, Electrical Engineering, Free elective for other B. Tech. disciplines

Semester : Even/Odd

In the context of data science, probability theory is the one of most important instrument for predictive modelling. It assists in the study of data and helps in making inferential decisions about it. Most of the commercial and open source software/packages available currently for the purpose of data analysis and prediction rely upon fundamental concepts of probability theory and statistics. A practitioner typically uses these concepts as a black box, but a data scientist capable of analyzing and designing new algorithms must have a clear understanding of the underlying notions. This helps in choosing/designing better tools for problem solving and analyzing the data. This course is intended to impart a strong background in probability theory to the students so that they are in a position to grasp notions involving probability in advanced topics such as statistical analysis, machine learning etc.

On completion of this course, students should be able to exhibit their understanding of various concepts of probability and statistics that are required in DSE for subsequent courses such as Statistical Learning and Pattern Recognition.

In particular, students should be able to

- understand the notion of random variables, their distribution properties and certain key results related to convergence and limit theorems.

- acquire adequate knowledge of various random processes and their properties.

- have a thorough understanding of various sampling techniques of practical importance and graphical models for machine learning and data science

Course modules with Quantitative lecture hours:

Sample space, Sigma field, axiomatic definition of probability, conditional probability and independence, Bayes Rule.

(2 lectures)

Review: Random variables – discrete and continuous, probability mass function, probability density function, some standard (important) pdfs, independence, expectation, variance, conditional distribution, conditional expectation, covariance and correlation, Functions of random variables.

(7 lectures)

Probability generating function, moment generating function and characteristic functions – properties and applications.

(3 lectures)

Convergence of random variables – basic results, inequalities (Markov and Chebyshev), law of large numbers (weak and strong), central limit theorem.

(5 lectures)

Random vectors and covariance and correlation matrix, Random processes – stationarity, WSS, Autocorrelation, cross correlation, power spectral density, Ergodicity. Wiener processes, Markov processes, Poisson Process.

(8 lectures)

Sampling methods: Inverse transforms sampling, Rejection sampling, adaptive rejection sampling, importance sampling, Markov chains and MCMC

(8 lectures)

Graphical models: ML and MAP estimation, directed and undirected models, Bayesian networks, CRF, Learning and Inference method (ML, MAP, Sampling)

(9 lectures)

Textbook:

1. Grimmett, Geoffrey, and David Stirzaker. *Probability and random processes*. Oxford university press, 2001.
2. Bishop, Christopher M. *Pattern recognition and machine learning*. Springer, 2006.

Reference:

1. Ross, Sheldon. *A first course in probability*. Pearson, 2014.
2. Stark, Henry, and John William Woods. *Probability, statistics, and random processes for engineers*. Pearson, 2012.
3. Papoulis, Athanasios, and S. Unnikrishna Pillai. *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.

IIT Mandi

Proposal for New Course

Course Number : DS313

Course Name : Statistical Foundations of Data Science

Credits : 3-1-0-4 (L-T-P-C)

Prerequisites : IC110-Engineering Mathematics, IC252-Data Science 2

Intended for : UG

Distribution : Discipline core for B. Tech. Data Science and Engineering, Discipline elective for B. Tech. Computer Science and Engineering, Electrical Engineering, Free elective for other B. Tech. disciplines

Semester : Even/Odd

In the context of data science, probability theory is the one of most important instrument for predictive modelling. It assists in the study of data and helps in making inferential decisions about it. Most of the commercial and open source software/packages available currently for the purpose of data analysis and prediction rely upon fundamental concepts of probability theory and statistics. A practitioner typically uses these concepts as a black box, but a data scientist capable of analyzing and designing new algorithms must have a clear understanding of the underlying notions. This helps in choosing/designing better tools for problem solving and analyzing the data. This course is intended to impart a strong background in probability theory to the students so that they are in a position to grasp notions involving probability in advanced topics such as statistical analysis, machine learning etc.

On completion of this course, students should be able to exhibit their understanding of various concepts of probability and statistics that are required in DSE for subsequent courses such as Statistical Learning and Pattern Recognition.

In particular, students should be able to

- understand the notion of random variables, their distribution properties and certain key results related to convergence and limit theorems.

- acquire adequate knowledge of various random processes and their properties.

- have a thorough understanding of various sampling techniques of practical importance and graphical models for machine learning and data science

Course modules with Quantitative lecture hours:

Sample space, Sigma field, axiomatic definition of probability, conditional probability and independence, Bayes Rule.

(2 lectures)

Review: Random variables – discrete and continuous, probability mass function, probability density function, some standard (important) pdfs, independence, expectation, variance, conditional distribution, conditional expectation, covariance and correlation, Functions of random variables.

(7 lectures)

Probability generating function, moment generating function and characteristic functions – properties and applications.

(3 lectures)

Convergence of random variables – basic results, inequalities (Markov and Chebyshev), law of large numbers (weak and strong), central limit theorem.

(5 lectures)

Random vectors and covariance and correlation matrix, Random processes – stationarity, WSS, Autocorrelation, cross correlation, power spectral density, Ergodicity. Wiener processes, Markov processes, Poisson Process.

(8 lectures)

Sampling methods: Inverse transforms sampling, Rejection sampling, adaptive rejection sampling, importance sampling, Markov chains and MCMC

(8 lectures)

Graphical models: ML and MAP estimation, directed and undirected models, Bayesian networks, CRF, Learning and Inference method (ML, MAP, Sampling)

(9 lectures)

Textbook:

3. Grimmett, Geoffrey, and David Stirzaker. *Probability and random processes*. Oxford university press, 2001.
4. Bishop, Christopher M. *Pattern recognition and machine learning*. Springer, 2006.

Reference:

4. Ross, Sheldon. *A first course in probability*. Pearson, 2014.
5. Stark, Henry, and John William Woods. *Probability, statistics, and random processes for engineers*. Pearson, 2012.
6. Papoulis, Athanasios, and S. Unnikrishna Pillai. *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.

IIT Mandi

Proposal for a New Course

Course Number : DS401

Course Name : Optimization for Data Science

Credits : 3-0-0-3 (L-T-P-C)

Prerequisites : IC111-Linear Algebra, IC110-Engineering Mathematics, DS301-Mathematical Foundations of Data Science I

Intended for : UG

Distribution : Discipline core for B. Tech. Data Science and Engineering, Discipline elective for B. Tech. Computer Science and Engineering, Electrical Engineering, Free elective for other B. Tech. disciplines

Semester : Even

Preamble

Optimization lies at the heart of data science. Most of the problems in data science and machine learning are formulated as optimization problems. Techniques such as regression, estimation, learning etc. used in data science when applied to large scale data sets need efficient optimization techniques to find the solution. This course is meant as a first course in optimization for the undergraduate data science and engineering program. The first part of this course gives a strong foundation on convex optimization. An exposure to typical algorithms and acceleration methods used in data science as well as machine learning are provided in the second part. An expertise in these topics will prepare the student for dealing with advanced optimization methods used in data science and machine learning.

On completion of this course, students should be able to demonstrate their understanding of and apply various concepts of optimization in DSE and subsequent courses in statistics, Pattern recognition, and Deep learning.

In particular, students should be able to

- understand concepts related with convex optimization and optimality conditions.
- understand some of the modern powerful and efficient algorithms to solve problems originating in data science.
- have familiarity with algorithms dealing with problems in data science having non convex optimization formulation

Course modules with Quantitative lecture hours:

Affine sets, convex sets, cone, examples – hyperplanes, halfspaces, polyhedra, simplexes, positive semidefinite cones. Operations that preserve convexity. Separating and supporting hyperplanes. Dual cones.

(6 lectures)

Convex function, first and second order conditions, epigraph, operations that preserve convexity, conjugate function.

(6 lectures)

Convex optimization – linear, quadratic, geometric, conic, semidefinite programming. Formulation of - unconstrained, equality constrained, inequality constrained and both – problems.

(7 lectures)

Duality – Lagrange dual function, bounds on the optimal value. Lagrange dual problem, weak and strong duality, optimality conditions.

(8 lectures)

Gradient methods – gradient descent, Lipschitz functions, smooth functions, projected gradient descent, Frank-Wolfe algorithm, Chebyhsev iterations, conjugate gradient, Nesterov's accelerated gradient descent. Dual gradient ascent, ADAM.

(10 lectures)

Nonconvex optimization – alternating minimization and expectation maximization algorithms, convex relaxations.

(5 lectures)

Textbook:

1. Boyd, Stephen, and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

Reference books:

1. Yurii, Nesterov, *Introductory lectures on convex optimization: a basic course*. Kluwer Academic Publishers, 2004.
2. Luenberger, D. G., and Y. Ye. *Linear and nonlinear programming*, Springer New York, 2008.
3. Nocedal, Jorge, and Stephen Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

IIT Mandi
Proposal for a New Course

Course Number	:	DS402
Course Name	:	Matrix Computations for Data Science
Credits	:	2-0-2-3
Prerequisites	:	IC110-Engineering Mathematics, IC111-Linear Algebra
Intended for	:	UG
Distribution	:	Discipline core for B. Tech. Data Science and Engineering, Discipline elective for B. Tech. Computer Science and Engineering, Electrical Engineering, Free elective for other B. Tech. disciplines
Semester	:	odd/even

Preamble:

Many of the machine learning algorithms when applied to some practical problems often requires dealing with massive systems of linear equations or matrices. Due to increasing complexity in the real world scenarios and recent advances in the area of data science, understanding of numerical linear algebra and large scale matrix computations has become essential for data science practitioners.

Objective and Learning Outcomes:

After doing courses on engineering mathematics and linear algebra, students get conversant with various concepts related to multivariable calculus, vector spaces, linear transformations and basic matrix properties. To further enhance their expertise to cater the requirements to deal with large scale applications, this course introduces matrix factorization methods and their implementations. Also, direct and iterative methods for solving system of linear equations and eigenvalue problems of large dimensions are discussed.

This course tries to answer the fundamental question of choice of suitable matrix computation method with objective of having adequate speed and suitable accuracy in the computations. Hands-on experience with all the methods covered is the most crucial part of this course. All the topics discussed in this course would be accompanied with parallel practical session to reinforce the learning outcome of the course.

On completion of the module, students will be able to

- do some key matrix factorizations with the help of elementary transformations,
- identify suitable technique to solve linear systems of equations, least squares problems, and the eigenvalue problem,
- equate the efficacy and stability of different algorithms applied to solve linear systems, least squares problems, and the eigenvalue problem,
- understand and develop suitable matrix structures to deal with cases such as sparsity, banded structure, and symmetric positive definiteness,

Course Modules with Quantitative Hours:

Basics : Errors in numerical computations. Review matrices and transformations, Matrix and Vector Norms.

(2 lectures)

Matrix factorizations : Cholesky factorization, QR factorization, Householder reflectors and Givens rotations. LU factorization and Gaussian elimination. Numerical stability Pivoting strategies and solution of triangular and full systems by factorization.

(5 lectures)

Eigenvalue problem: Basic theory, Power method, bisection method, QR algorithm. Similarity reduction.

(4 lectures)

Linear least squares problem : Singular value decomposition, Moore-Penrose Pseudoinverse. Perturbation theory. Least square and normal equations. SVD and rank deficiency, Principal Component Analysis, Linear Discriminant Analysis.

(6 lectures)

Iterative methods for linear systems : Iterative methods for linear systems Iterative methods: Jacobi, Gauss-Seidel and SOR iterations. Kronecker product. Krylov subspace methods, conjugate gradient method, preconditioning.

(5 lectures)

Sparse and banded linear systems :

Storage schemes for banded and sparse matrices, Sparse matrices and sparse solutions: approximate inverses, eigenvalues, incomplete factorization. Matrix regularization: matrix completion.

(6 lectures)

Lab Exercises: Lab would be conducted on a 2-hour weekly slot. Lab sessions would be conducted in tandem with the theory course so the topics for problems given in the lab are already initiated in the theory class.

Textbook

1. Trefethen, Lloyd N., and David Bau III. *Numerical linear algebra*. Vol. 50. SIAM, 1997.

2. Eldén, Lars. *Matrix methods in data mining and pattern recognition*. Vol. 4. SIAM, 2007.

Reference

1. Watkins, David S. *Fundamentals of matrix computations*. Vol. 64. John Wiley & Sons, 2004.

2. Demmel, James W. *Applied numerical linear algebra*. Vol. 56. SIAM, 1997.

3. Golub, Gene H., and Charles F. Van Loan. *Matrix computations*. Vol. 3. JHU press, 2012.

4. Cullen, Charles G. *An introduction to numerical linear algebra*. PWS Publishing Company, 1994.

IIT Mandi

Proposal for a New Course

Course Number : DS403

Course Name : Introduction to Statistical Learning

Credits : 2-0-2-3 (L-T-P-C)

Prerequisites : DS201 - Data Handling and visualization or equivalent,
DS303 – Statistical Foundations of Data Science or equivalent,
DS402 – Matrix Computations for Data Science

Intended for : UG

Distribution : Discipline core for B. Tech. Data Science and Engineering,
Discipline elective for B. Tech. Computer Science and Engineering,
Electrical Engineering, Free elective for other B. Tech. disciplines

Semester : Even/Odd

Preamble: Data science involves using the scientific methods to process the data to extract knowledge and insights from data in various forms, both structured and unstructured. Vast amounts of data are being generated in many fields, and the data analyst's job is to make sense of it all: to extract important patterns and trends. This involves learning from data. This course intends to provide fundamental ideas and techniques in learning from data in the statistical framework.

Objective: After a course on data handling and visualization, and course(s) on statistical and mathematical foundations of data science, the students acquire familiarity with data in various forms, their representations, statistical and mathematical properties for analysing the data. This introductory course aims to provide foundations to different techniques for learning from useful information/patterns and trends from data. This course provide rigorous theoretical foundations to those techniques and integrate those with practical applications in various fields. However, the course shall strive to introduce relevant applications as and when needed to motivate/illustrate various topics. This course also aims at providing practical implementation of different techniques used for learning from data and making predictions on the real world data.
On completion of this course, the student should be able to

- Understand the different learning mechanisms like supervised and unsupervised learnings
- Learn to analyse data to gain insights using an appropriate learning algorithm under a given task and context
- Learn various learning algorithms for tasks such as regression, classification and clustering
- Evaluate the learning algorithms on the data from different areas and disciplines and make sense from them

Course modules with Quantitative lecture hours:

5. **Introduction to learning from data:** **(2 lecture)**
   Introduction to supervised learning and unsupervised learning.

6. **Supervised learning: Regression** **(6 lectures)**
   Linear regression models and least squares, Shrinkage methods: ridge regression and the LASSO.

7. **Supervised learning: Classification** **(8 lectures)**
   Logistic regression, nearest neighbour's method, Bayes classifier with unimodal and multimodal density - maximum likelihood estimation, expectation-maximization (EM) algorithm; decision trees, support vector machines (SVMs), basics of neural networks

8. **Model Assessment and Selection:** **(6 lectures)**
   Bias, variance and model complexity, The Bayesian approach, AIC and BIC, cross-validation, bootstrap methods, hypothesis testing, confidence intervals, significance testing.

9. **Unsupervised learning:** **(6 lectures)**
   Introduction to association rules, clustering, and dimension reduction.

Lab Exercises:

- Lab to be conducted on a 2-hour slot. It will be conducted in tandem with the theory course so the topics for problems given in the lab are already initiated in the theory class. The topics taught in the theory course should be appropriately be sequenced for synchronization with the laboratory.

Textbooks:

1. Hastie, T., Tibshirani, R. and Friedman, J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition, Springer, 12th Reprint, 2017.

Reference books:

1. Duda, R. O., Hart, P. E. and Stork, D. G., Pattern Classification, John Wiley, 2001.
2. Bishop, C. M., Pattern Recognition and Machine Learning, Springer, 2006.
3. Theodoridis, S. and Koutroumbas, K., Pattern Recognition, Academic Press, 2009.

IIT Mandi

Proposal for a New Course

Course Number : DS404

Course Name : Information Security and Privacy

Credits : 3-0-0-3

Prerequisites : DS203-Mathematical Foundations of Data Science I

Intended for : B.Tech.

Distribution : Discipline core for B. Tech. Data Science and Engineering, Discipline elective for B. Tech. Computer Science and Engineering, Electrical Engineering, Free elective for other B. Tech. disciplines

Semester : Even/Odd.

Objective: During the last two decades, there has been a major paradigm shift in generation, processing, communication, and storage of information from predominantly analog domain to digital domain for the reasons such as ease of implementation, better efficiency, greater robustness against noise, and enhanced performance and security. This shift has not only resulted in opening up exciting new possibilities in numerous areas of sciences, engineering, and humanities, but have also created plethora of possibilities for the misuse of this data, raising serious security, privacy and ethical challenges. To address these challenges, DSE curriculum must not only introduce the students to spectrum of such challenges, but also equip them with various concepts, tools, techniques. This will enable the students to better understand information security, privacy, and ethical challenges and appreciate the current developments at the frontiers of these areas, both in theory and application, and prepare them to contribute to further advancement of such frontiers.

On completion of this course, students should be able to demonstrate their understanding of and apply methods of information security, privacy, and ethics to subsequent courses in data mining, pattern recognition, knowledge representation, information systems, to name a few. In particular, students should be able to

- demonstrate understanding and awareness of information security, privacy, and ethical issues in a given application scenario.

- propose practical solutions to address such issues using relevant concepts, tools, and techniques

Course modules with Quantitative lecture hours:

1. **Introduction to information security:** **(8 lectures)**
   Information security models; attacks, threats, vulnerabilities, and risks. Operations security: Haas' Laws.
   Identification and authentication: identity verification, falsifying identification, multifactor and mutual authentication, passwords, biometrics, hardware tokens, performance evaluation.

Authorization and access control: principle of least privilege, access control lists, and access control methodologies, physical security and access controls.
Auditing and accountability: non-repudiation, deterrence, intrusion detection and prevention, logging, monitoring, assessments.

2. **Cryptography:** **(8 lectures)**
   Protocols (key exchange, public key cryptography, secret sharing), techniques (key length, key management, etc), cryptographic algorithms (mathematical background, data encryption standard, block and stream ciphers, public-key, digital signatures).

3. **Network security:** **(6 lectures)**
   Protecting networks and network traffic, mobile device security, network security tools.

4. **Operating system security:** **(4 lectures)**
   OS hardening, protecting against malware, firewalls and host intrusion detection, OS security tools.

5. **Application security:** **(6 lectures)**
   Software development vulnerabilities, web security, database security, and application security tools.

6. **Information privacy:** **(6 lectures)**
   Static and dynamic data anonymization and threats to anonymization, privacy in synthetic and test data, privacy regulations.

7. **Information Ethics:** **(4 lectures)**
   Ownership, privacy, anonymity, validity, algorithmic fairness, societal consequences, code of ethics, attributions.

Textbook Books:

1. Andress, J. and Winterfeld, S., The Basics of Information Security, 2/e, Syngress, 2014.
2. Venkataraman, N. and Shriram, A., Data Privacy: Principles and Practice, Chapman and Hall/CRC, 2016.

Reference Books:

1. Guise, P. D., *Data Protection*, Routledge, 2017.
2. Katz, J. and Lindell, Y., *Introduction to Modern Cryptography,* Chapman and Hall/CRC, 2015.
3. Torra, V., *Data Privacy: Foundations and the Big data Challenge*, Springer, 2017.
4. Weigand, A., *Data for the People*, Basic Books, 2017.

IIT Mandi

Proposal for a New Course

Course Number : DS411

Course Name : Optimization for Data Science

Credits : 3-1-0-4 (L-T-P-C)

Prerequisites : IC111-Linear Algebra, IC110-Engineering Mathematics, DS301-Mathematical Foundations of Data Science I

Intended for : UG

Distribution : Discipline core for B. Tech. Data Science and Engineering, Discipline elective for B. Tech. Computer Science and Engineering, Electrical Engineering, Free elective for other B. Tech. disciplines

Semester : Even

Preamble

Optimization lies at the heart of data science. Most of the problems in data science and machine learning are formulated as optimization problems. Techniques such as regression, estimation, learning etc. used in data science when applied to large scale data sets need efficient optimization techniques to find the solution. This course is meant as a first course in optimization for the undergraduate data science and engineering program. The first part of this course gives a strong foundation on convex optimization. An exposure to typical algorithms and acceleration methods used in data science as well as machine learning are provided in the second part. An expertise in these topics will prepare the student for dealing with advanced optimization methods used in data science and machine learning.

On completion of this course, students should be able to demonstrate their understanding of and apply various concepts of optimization in DSE and subsequent courses in statistics, Pattern recognition, and Deep learning.

In particular, students should be able to

- understand concepts related with convex optimization and optimality conditions.
- understand some of the modern powerful and efficient algorithms to solve problems originating in data science.
- have familiarity with algorithms dealing with problems in data science having non convex optimization formulation

Course modules with Quantitative lecture hours:

Affine sets, convex sets, cone, examples – hyperplanes, halfspaces, polyhedra, simplexes, positive semidefinite cones. Operations that preserve convexity. Separating and supporting hyperplanes. Dual cones.

(6 lectures)

Convex function, first and second order conditions, epigraph, operations that preserve convexity, conjugate function.

(6 lectures)

Convex optimization – linear, quadratic, geometric, conic, semidefinite programming. Formulation of - unconstrained, equality constrained, inequality constrained and both – problems.

(7 lectures)

Duality – Lagrange dual function, bounds on the optimal value. Lagrange dual problem, weak and strong duality, optimality conditions.

(8 lectures)

Gradient methods – gradient descent, Lipschitz functions, smooth functions, projected gradient descent, Frank-Wolfe algorithm, Chebyhsev iterations, conjugate gradient, Nesterov's accelerated gradient descent. Dual gradient ascent, ADAM.

(10 lectures)

Nonconvex optimization – alternating minimization and expectation maximization algorithms, convex relaxations.

(5 lectures)

Textbook:

1. Boyd, Stephen, and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

Reference books:

1. Yurii, Nesterov, *Introductory lectures on convex optimization: a basic course*. Kluwer Academic Publishers, 2004.
2. Luenberger, D. G., and Y. Ye. *Linear and nonlinear programming*, Springer New York, 2008.
3. Nocedal, Jorge, and Stephen Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

IIT Mandi

Proposal for a New Course

Course Number	:	DS412
Course Name	:	Matrix Computations for Data Science
Credits	:	3-0-2-4
Prerequisites	:	IC110-Engineering Mathematics, IC111-Linear Algebra
Intended for	:	UG
Distribution	:	Discipline core for B. Tech. Data Science and Engineering, Discipline elective for B. Tech. Computer Science and Engineering, Electrical Engineering, Free elective for other B. Tech. disciplines
Semester	:	odd/even

Preamble:

Many of the machine learning algorithms when applied to some practical problems often requires dealing with massive systems of linear equations or matrices. Due to increasing complexity in the real world scenarios and recent advances in the area of data science, understanding of numerical linear algebra and large scale matrix computations has become essential for data science practitioners.

Objective and Learning Outcomes:

After doing courses on engineering mathematics and linear algebra, students get conversant with various concepts related to multivariable calculus, vector spaces, linear transformations and basic matrix properties. To further enhance their expertise to cater the requirements to deal with large scale applications, this course introduces matrix factorization methods and their implementations. Also, direct and iterative methods for solving system of linear equations and eigenvalue problems of large dimensions are discussed.

This course tries to answer the fundamental question of choice of suitable matrix computation method with objective of having adequate speed and suitable accuracy in the computations. Hands-on experience with all the methods covered is the most crucial part of this course. All the topics discussed in this course would be accompanied with parallel practical session to reinforce the learning outcome of the course.

On completion of the module, students will be able to

- do some key matrix factorizations with the help of elementary transformations,
- identify suitable technique to solve linear systems of equations, least squares problems, and the eigenvalue problem,
- equate the efficacy and stability of different algorithms applied to solve linear systems, least squares problems, and the eigenvalue problem,
- understand and develop suitable matrix structures to deal with cases such as sparsity, banded structure, and symmetric positive definiteness,

Course Modules with Quantitative Hours:

Basics : Errors in numerical computations. Review matrices and transformations, Matrix and Vector Norms.

(2 lectures)

Matrix factorizations : Cholesky factorization, QR factorization, Householder reflectors and Givens rotations. LU factorization and Gaussian elimination. Numerical stability Pivoting strategies and solution of triangular and full systems by factorization.

(5 lectures)

Eigenvalue problem: Basic theory, Power method, bisection method, QR algorithm. Similarity reduction.

(4 lectures)

Linear least squares problem : Singular value decomposition, Moore-Penrose Pseudoinverse. Perturbation theory. Least square and normal equations. SVD and rank deficiency, Principal Component Analysis, Linear Discriminant Analysis.

(6 lectures)

Iterative methods for linear systems : Iterative methods for linear systems Iterative methods: Jacobi, Gauss-Seidel and SOR iterations. Kronecker product. Krylov subspace methods, conjugate gradient method, preconditioning.

(5 lectures)

Sparse and banded linear systems :

Storage schemes for banded and sparse matrices, Sparse matrices and sparse solutions: approximate inverses, eigenvalues, incomplete factorization. Matrix regularization: matrix completion.

(6 lectures)

Lab Exercises: Lab would be conducted on a 2-hour weekly slot. Lab sessions would be conducted in tandem with the theory course so the topics for problems given in the lab are already initiated in the theory class.

Textbook

1. Trefethen, Lloyd N., and David Bau III. *Numerical linear algebra*. Vol. 50. SIAM, 1997.

2. Eldén, Lars. *Matrix methods in data mining and pattern recognition*. Vol. 4. SIAM, 2007.

Reference

5. Watkins, David S. *Fundamentals of matrix computations*. Vol. 64. John Wiley & Sons, 2004.

6. Demmel, James W. *Applied numerical linear algebra*. Vol. 56. SIAM, 1997.

7. Golub, Gene H., and Charles F. Van Loan. *Matrix computations*. Vol. 3. JHU press, 2012.

8. Cullen, Charles G. *An introduction to numerical linear algebra*. PWS Publishing Company, 1994.

IIT Mandi

Proposal for a New Course

Course Number : DS413

Course Name : Introduction to Statistical Learning

Credits : 3-1-0-4 (L-T-P-C)

Prerequisites : DS201 - Data Handling and visualization or equivalent,
DS303 – Statistical Foundations of Data Science or equivalent,
DS402 – Matrix Computations for Data Science

Intended for : UG

Distribution : Discipline core for B. Tech. Data Science and Engineering,
Discipline elective for B. Tech. Computer Science and Engineering,
Electrical Engineering, Free elective for other B. Tech. disciplines

Semester : Even/Odd

Preamble: Data science involves using the scientific methods to process the data to extract knowledge and insights from data in various forms, both structured and unstructured. Vast amounts of data are being generated in many fields, and the data analyst's job is to make sense of it all: to extract important patterns and trends. This involves learning from data. This course intends to provide fundamental ideas and techniques in learning from data in the statistical framework.

Objective: After a course on data handling and visualization, and course(s) on statistical and mathematical foundations of data science, the students acquire familiarity with data in various forms, their representations, statistical and mathematical properties for analysing the data. This introductory course aims to provide foundations to different techniques for learning from useful information/patterns and trends from data. This course provide rigorous theoretical foundations to those techniques and integrate those with practical applications in various fields. However, the course shall strive to introduce relevant applications as and when needed to motivate/illustrate various topics. This course also aims at providing practical implementation of different techniques used for learning from data and making predictions on the real world data.
On completion of this course, the student should be able to

- Understand the different learning mechanisms like supervised and unsupervised learnings
- Learn to analyse data to gain insights using an appropriate learning algorithm under a given task and context
- Learn various learning algorithms for tasks such as regression, classification and clustering
- Evaluate the learning algorithms on the data from different areas and disciplines and make sense from them

Course modules with Quantitative lecture hours:

1. **Introduction to learning from data:** **(2 lecture)**

   Introduction to supervised learning and unsupervised learning.

2. **Supervised learning: Regression** **(6 lectures)**

   Linear regression models and least squares, Shrinkage methods: ridge regression and the LASSO.

3. **Supervised learning: Classification** **(8 lectures)**

   Logistic regression, nearest neighbour's method, Bayes classifier with unimodal and multimodal density - maximum likelihood estimation, expectation-maximization (EM) algorithm; decision trees, support vector machines (SVMs), basics of neural networks

4. **Model Assessment and Selection:** **(6 lectures)**

   Bias, variance and model complexity, The Bayesian approach, AIC and BIC, cross-validation, bootstrap methods, hypothesis testing, confidence intervals, significance testing.

5. **Unsupervised learning:** **(6 lectures)**

   Introduction to association rules, clustering, and dimension reduction.

Lab Exercises:

- Lab to be conducted on a 2-hour slot. It will be conducted in tandem with the theory course so the topics for problems given in the lab are already initiated in the theory class. The topics taught in the theory course should be appropriately be sequenced for synchronization with the laboratory.

Textbooks:

2. Hastie, T., Tibshirani, R. and Friedman, J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition, Springer, 12th Reprint, 2017.

Reference books:

4. Duda, R. O., Hart, P. E. and Stork, D. G., Pattern Classification, John Wiley, 2001.
5. Bishop, C. M., Pattern Recognition and Machine Learning, Springer, 2006.
6. Theodoridis, S. and Koutroumbas, K., Pattern Recognition, Academic Press, 2009.