Mathematical Foundations Course Outline

04. Mathematical Foundations for Data Science and Machine Learning Course Description:

Mathematics plays a very important role in Data Science and Machine Learning as it is helpful in understanding the nuances of data and the relationships existing in the data. It helps in modelling the data science problems in a more formal way so that patterns in data are understood and extracted. This course focuses on the topics essential to build concepts in Data Science and Machine Learning. The math behind data representation, transformation, analysis and useful functions for optimization of the same are discussed.

Learning Outcomes

After completing this course, you will be able to:

- Understand the mathematical foundations essential to data science
- Represent data in vector and matrix forms
- Perform operations on vectors and matrices
- Solve a linear system of equations
- Calculate the Eigen Values for Eigen Vectors
- Perform matrix factorization and dimensionality reduction using SVD and PCA
- Calculate the derivatives of a function
- Use different function optimizations in ML

Pedagogy

The course is a mixture of classroom lectures, quizzes, assignments and mini-projects. Jupyter notebook will be the medium of coding in python.

Course Content

Session – 1 Introduction to Linear Algebra

- Linear Algebra Overview
- Applications of Linear Algebra
- Vector properties and operations
- Linear Independence
- Change of basis (Vector)
- Matrix Introduction
- Types of Matrix

Session – 2 Linear Transformations

- Basic Matrix Transformations
- Determinant
- Matrix Inverse
- Determinant and Inverse for special matrices
- Orthogonal matrix & Gram-Schmidt Process
- Eigen Values and Vectors
- Eigen basis and transformations

Session – 3 Linear Algebra for ML

- Mathematical modelling of Machine Learning problem -
    - Linear Regression Overview

- ML problem formulation
- Explanation through an ex● Mathematical modelling of Machine Learning problem - ○ Linear Regression Overview ○ ML problem formulation ○ Explanation through an example ● Image Pre-processing ○ Image basics - Pixel, representation ○ Linear Transformations ○ Filtering Session – 4 Linear Algebra Application ex ample

- Image Pre-processing
  - Image basics - Pixel, representation
  - Linear Transformations
  - Filtering

Session – 4 Linear Algebra Applications

- Covariance Matrix
- Principal Component Analysis - Dimensionality Reduction
- Matrix Factorization - SVD
- Image compression

Session – 5 Functions and its Derivatives

- Function Definition and Representation
- Linear and Non-Linear functions
- Single and Multivariate functions
- Function Differentiation and Integration
- Rules of Derivatives

Session - 6 Function Optimizations

- Introduction to Optimization - Maxima & Minima
- First order optimization algorithms
- Gradient Descent and Stochastic Gradient Descent

Some frequently asked Questions

1. Why do we need to learn the Mathematical foundation in the mtech course ?

Ans: This course focuses on the topics essential to build concepts in Data Science and Machine Learning. The math behind data representation, transformation, analysis and useful functions for optimization of the same are discussed. Along with the theoretical aspect the course also covers a set of applications using the core concepts and hands on for the same.

This course structure helps in modelling the data science problems in a more formal way so that patterns in data are understood and extracted.

2. What are the key topics available in the course?

Ans:
- Represent data in vector and matrix forms
- Perform operations on vectors and matrices

- Solve a linear system of equations
- Calculate the Eigen Values for Eigen Vectors
- Perform matrix factorization and dimensionality reduction using SVD and PCA
- Calculate the derivatives of a function
- Use different function optimizations in ML

3. Why are we teaching the basic concepts like vectors and matrices in this module?

Ans: Vectors and matrices are fundamental concepts in linear algebra, and many advanced topics in data science and machine learning build upon these foundations. Providing a strong understanding of these basics sets the stage for more complex mathematical concepts.

Also we have learners with diverse experience levels and from various disciplines as well. Teaching a course with students from diverse experience levels, especially in the context of data science and machine learning, requires careful consideration and strategies to ensure that all students can benefit from the learning experience.

4. What are the topics for which Flip videos are not available in olympus?
Ans: There are a total 30 flip videos available and they are session wise divided in the portal.

However for students reference, some extra materials are shared available here
https://olympus.mygreatlearning.com/courses/107859/files/9703623?module_item_id=5247700

5. Can we have extra questions to solve for practice?
Ans: We have already shared the take homes which you can practise. Also 3 ESA qps are shared, which can be used for practice purposes.
https://olympus.mygreatlearning.com/courses/100312/files/9684289?module_item_id=5231102

6. What are the ISA (Internal Semester Assessment) components for this course?
Ans: Quiz – 30 marks, Graded Assessment – 40 marks (python), mini project – 20 marks, Attendance – 10 marks (Total 100).

7. What is the pattern of End Semester Assessment (ESA) and Graded assessment (GA)?
Ans:ESA is both handwritten (20 marks) and coded python based (80) marks. Please refer to the model QP provided in the sample.

8. Why is the exam pattern python (coding) based?
Ans: Python-based coding exams internally may be used to assess the application of mathematical concepts in solving real-world problems. This aligns with the idea of using programming as a tool for mathematical modeling and problem-solving. In mathematics, especially in areas like computational mathematics, coding skills can be valuable. Internal assessments in Python focus on numerical methods, algorithms, and simulations that involve coding.

9. Is there a group assignment in the course?
Ans: There is no group assignment for this course.

10. Students' performance in this course in the past?
Ans: Students in the past have performed well and failure percentage is <5% and have benefited from this course during interviews.

11. Are there any textbooks to supplement the material shared?
Ans: Yes. Please find the list of reference books below:

Books:

1. Introduction to Linear Algebra, Sixth Edition (2023) by Gilbert Strang
2. Interactive Linear Algebra by Dan Margalit, Joseph Rabinoff
3. Linear Algebra Done Right by Sheldon Axler