Fusemachines
AI School

Mathematics for AI
Foundations in Artificial Intelligence Program

Syllabus

Version	Significant Changes (Marked with a Symbol)	Modified by	Modification Date
1.0	First Version of the course	Miran, Bipin	Dec 2019
2.0	Second Version	Rojesh	Dec 2020

Introduction to the Course

Course Objectives

The objectives of this course on the Mathematics for AI are:

- to develop computational thinking required before diving into AI
- to acquire the skills necessary to apply this understanding to develop computer-based solutions to problems.

Guided Hours

12 weeks course maximum of 4 hours per week.

Modules	**Tentative Time(hrs)**
Linear Algebra	12
Calculus	14.5
Probability and Statistics	14
Information Theory	4
Numerical Computation	3.5

Pre-requisites

Students must have completed CS for AI and have a strong foundation in Python programming.

Distinct Features Used in the Syllabus

Bold Outcomes refers to **Must Have** learning outcomes,

Normal Text refers to Should Have learning outcomes,

Italic Outcomes refers to *Good to Have or Could Have learning outcomes*.

Course Contents

Module 1. Introduction to the Course

Unit 1.1. Introduction to the Course

Students should be able to:

1.1.1. Introduction to the Course	• Understand what course covers and what it does not cover • Understand the course structure and flow of the course through the course overview • Appreciate how the math topics covered in this course are related to AI engineer
1.1.2. Course Logistics	• Understand how blended course will work and be clear about their expectations from different components • Understand the assessment and evaluation criteria • Appreciate the honor code and violation issues

Module 2. Linear Algebra

Main Reference:

1. **MIT 18.06SC Linear Algebra:** https://www.youtube.com/playlist?list=PL221E2BBF13BECF6C

2. **Essence of Linear Algebra:** https://www.youtube.com/watch?v=fNk_zzaMoSs&list=PLZHQObOWTQDPD3MizzM2xVFitgF8hE_ab

Unit 2.1. Introduction to the Module

Students should be able to :

2.1.1. Introduction to the Module	• Understand what they are expected to learn out of the module.

Unit 2.2. Scalars, Vectors and their operations [1 hr]

Update proposition:Vector Outer Product, Vector Cross Product, P-norm, norms and distances, *change of basis and its application.

Students should be able to :

Scalar and Vector	• Differentiate scalar and vector • Appreciate the difference in viewpoint of vectors applied to CS and physics. • Understand magnitude as L2 norm and direction concept in vector in 2d and 3d space • Visualize vectors [2D, 3D] and generalize in multi-dimension with examples in data science.
Vector Norms, Dot product, Orthogonal Vectors	• Compare the formula of L2 norm with L1, L0 and L infinity norm. • Define dot products and interpret geometrically in terms of projection and angle between vectors, including special cases such as orthogonality.
Linear Combination	• Understand addition and scaling operation in vector and visualize geometrically • Understand and visualise geometrically, the linear combination of vectors. [Combine Add and Scaling operator]
Subspace, span	• Understand the vector space. • Distinguish between vector space and subspace. • Understand span in relation to the linear combination.
Linear dependence,	• Understand and visualise geometrically the linear

independence	dependence of vectors ● Understand and visualise geometrically the linear independence of vectors
Basis and Dimension	● Relate basis and dimension of vector space with span. ● visualize basis in R^2 and R^3 other than standard orthonormal basis: i, j, k.

Unit 2.3 Linear Transformations and Matrix [0.5 hr]

Update Proposition: Properties of matrix multiplication and other operations

Students should be able to :

Linear transformation	● Understand the concept of linear functions and linear transformations ● Understand matrix as a linear transformation. ● Define matrix based on the understanding of the above two points.
Matrix Multiplication	● Understand and be able to perform product between matrix and vector, and matrix and matrix. ● Understand matrix-vector multiplication as a linear combination of columns, and vector-matrix multiplication as linear combination of rows.
Other matrix operations	● Carry out sum, transpose operation, trace and Frobenius norm of a matrix ● Define special type of matrix: Identity, Scalar, Diagonal, Symmetric, Anti-Symmetric.

Unit 2.4 Solving Linear Equations [3 hr]

Update proposition: Remove Four Fundamental spaces

Students should be able to :

Linear equation as a matrix	● Understand and visualize row picture, and column picture of a linear equation. ● Represent Linear equations in the form Ax = b
Gauss Elimination, LU factorization	● solve linear equations using gauss elimination method. ● decompose matrix into LU, LDU, PLU, PLDU to solve linear combination
Systems with unique solutions	● Relate system of equations with linear dependence and independence. ● relate rank of a matrix based on linear dependence and independence. ● recognize when a system of equations has unique solutions
Systems with multiple or no solutions	● recognize when a system of equations has multiple or no solution. ● Recognize low-rank matrices and their relation to such systems ● describe null space

Four fundamental spaces	- Understand the core concept of four fundamental spaces and link it with the solution to the system of equations. - Understand the relation between subspaces.

Unit 2.5 Determinant and Inverses

Update proposition: Properties of determinants, Inverse concept without fundamental subspace concept(last chapter)

Students should be able to :

Inverse	- Realize the importance of Ax=b with real world example - Understand inverse from the view point of solving Ax=b - Understand a condition where inverse is defined and undefined (from the viewpoint of geometry and rank) - Compute inverse by Gauss-Jordan method
Determinant	- Understand determinant and interpret geometrically as a scale of space - Define singular matrix in terms of - Rank, - solution to Ax=b and - determinants
Left, right and pseudo-inverse	Understand the condition for left, right and pseudo inverses with respect to four fundamental subspaces and rank

Unit 2.6 Orthogonality

Students should be able to :

Orthogonal matrix	- Understand rotating objects with rotation matrix and lead this concept to orthogonal matrix - Define orthogonal matrix - Understand it’s properties: determinants, inverse
Projection	- Understand the importance of projection - formulate a projection matrix
Least Squares	- understand the application of least squares - formulate the problem in terms of least squares - compute the least-squares solution using matrix form with ideas from projection matrix and left inverse

Unit 2.7 Eigen and Singular Value Decomposition

Students should be able to :

Eigenvalues and Eigenvectors	- Understand and visualize geometrically the eigenvalues and eigenvectors - Find the eigenvalues and eigenvectors with the concept of

	○ determinant and ○ null space
Eigen-decomposition	● Compute the eigen decomposition for ○ square matrices and ○ Symmetric matrices ● Relate it with PCA and Eigen-Faces
Definitiveness	● Realize the use case of positive definite and positive semi-definite matrices ● define in terms of energy
SVD	● Realize the use case of SVD through real-world example ● Understand the formulation of SVD and relate $U \Sigma V^T$ with respect to eigenvectors and eigenvalues of AA^T and $A^T A$ ● Relate SVD with PCA

Unit 2.8 Module Summary

Students should be able to :

Module Summary	A. Summarise the overall concepts covered in the module

Additional Resources

- Books
  http://joshua.smcvt.edu/linearalgebra/book.pdf

  http://148.206.53.84/tesiuami/S_pdfs/Linear%20Algebra%20Done%20Right.pdf

  https://github.com/CompPhysics/ComputationalPhysicsMSU/blob/master/doc/Lectures/Golub%2C%20Van%20Loan%20-%20Matrix%20Computations.pdf

  https://twiki.cern.ch/twiki/pub/Main/AVFedotovHowToRootTDecompQRH/Golub_VanLoan.Matr_comp_3ed.pdf

  http://vmls-book.stanford.edu/vmls.pdf

  http://www.deeplearningbook.org/contents/linear_algebra.html

  https://www.math.ucdavis.edu/~linear/linear-guest.pdf

- Video Lectures from university
  3 blue 1 brown: https://www.youtube.com/watch?v=fNk_zzaMoSs&list=PLZHQObOWTQDPD3MizzM2xVFitgF8hE_ab

  https://www.edx.org/course/applications-linear-algebra-part-1-davidsonx-d003x-1

  https://www.khanacademy.org/math/linear-algebra http://cs.brown.edu/courses/cs053/current/lectures.htm

https://github.com/fastai/numerical-linear-algebra/(seems to be focused mainly on implementation aspects)

https://www.edx.org/course/linear-algebra-foundations-to-frontiers-2

- Lecture Notes
  https://stanford.edu/~shervine/teaching/cs-229/refresher-algebra-calculus

Module 3. Calculus

Main Reference:

1. The essence of Calculus Video Series -Grant Sanderson for 3 blue 1 brown
https://www.youtube.com/watch?v=WUvTyaaNkzM&list=PLZHQObOWTQDMsr9K-rj53DwVRMYO3t5Yr
2. Calculus-1 - Khan Academy
https://www.khanacademy.org/math/calculus-1
3. Multivariable Calculus - Khan Academy
https://www.khanacademy.org/math/multivariable-calculus

Prerequisite / Initial assignment:

Students should be able to :

Review videos: Essence of Calculus	• Refresh high school level calculus. • Develop intuitions and link concepts to high school level calculus.
Limits and Continuity	• Understand the concepts of limits of a function • Understand when a function is continuous and differentiable.

Unit 3.1 Introduction to the module

Students should be able to :

Introduction to the Module	• Understand what they are expected to learn out of the module.

Unit 3.2 Derivatives [4]

Update proposition: Remove Taylor series and numerical methods of calculus

Students should be able to :

Refresher concepts and tools for derivatives	• Understand the need for derivative via problems relating to rate of changes • Understand derivative as a tangent and limit of secant of a function. • Make use of sum rule, product rule and quotient rule. • **Make use of chain rule. (Numerical method: Differential programming)** • Perform derivatives through implicit differentiation. • Compute (higher orders, at least up to second-order) derivatives of polynomial, trigonometric, exponential, logarithmic, hyperbolic, piecewise functions like ReLU.
Taylor's series for single variable	• Understand how derivatives can be used to approximate a function at a point and its neighborhood. • Approximate functions such as exponential,

	trigonometric, logarithmic functions by polynomials for x = 0 and generalize for x = a. ● Understand when Taylor's Series converges (radius of convergence).
Numerical Differentiation: Finite Difference	● Understand how derivatives can be calculated for discrete functions using finite difference methods: ○ First-order derivative by forward, backward, and central difference ○ Second-order derivative by forward, backward, and central difference ● Understand caveats related to approximating derivative by finite difference methods
Optimization	● Link properties of stationary points to Taylor's expansion. ● Locate Stationary Points and visualize them. ● Use the second derivative test to check if a stationary point is a local minima or local maxima. ● Differentiate between local and global minima/maxima. ● Solve simple single variable optimization problems analytically. The selected problems should illustrate the link to learning: finding the best parameters with the target task.
Solving Optimization Numerically	● Understand the need for an iterative method to compute maxima/minima of function with some real-world examples. The selected problems should illustrate a link to learning: finding best parameters with target task. ● Understand how to use gradient descent/ascent to compute minima/maxima[first order iterative optimization algorithm] ● *Understand how to use Newton's method for finding roots to compute minima/maxima[second order iterative optimization algorithm]*
Concave and Convex	● Understand graphically and be able to check analytically if any function is concave upwards or downwards at a certain point. ● Understand the concept of inflection point visually as well as analytically. ● Be able to understand the role of inflection points in the nature of stationary points.

Unit 3.3 Integrals [3]

Update Proposition: Remove "Relate Taylor's Series with fundamental Theorem of Calculus" and Numerical methods for integration

Students should be able to :

Refresher concepts	● Understand the need for integrals via problems relating to the area under a curve. ● Understand the fundamental theorem of calculus

	and relate integrals to derivatives. ● Understand integral as a reverse process of the derivative. ● Differentiate between definite and indefinite integrals. ● Relate Taylor's Series with fundamental Theorem of Calculus
Tools for Integrals	● Understand and deduce basic integral formulas for polynomial, exponential and trigonometric functions. ● Make use of sum rule, change of variable techniques. ● *Make use of techniques such as partial fractions and integration by parts.*
Numerical Integration	● Understand the computation of definite integral with: ○ **Trapezoidal method as a one of the case of Newton Cotes Quadrature formula to compute integral and the error order in the approximation of the integral** ○ *Simpson's ⅓ rule* ○ *Simpson's ⅜ rule*
Application	● Compute the area of a region bounded by a curve and lines parallel to the axes, or between a curve and a line or between two curves. ● Compute the volume of revolution about one of the axes.

Unit 3.4 Multivariable Calculus [4 hrs]

Update proposition: remove Taylor series for multivariate calculus

Students should be able to :

Introduction to multivariable functions	● Understand and differentiate between multivariable functions and single variable functions. ● Understand and differentiate between scalar-valued and vector-valued functions. ● Graph or visualize scalar-valued and vector-valued functions with two independent variables using 3D graphs, contour plots, fluid flow, etc. ● Motivate multivariable functions using real-world examples such as: ○ image as a function, ○ ML datasets.
Partial Differentiation	● Understand visually the concepts of partial differentiation. ● Compute Partial Differentiation of scalar-valued functions with examples of gradient of image to produce edges.
Directional Derivatives and	● Understand and compute (analytically and visually)

Gradient	concept of directional derivatives. ● Understand and compute the Gradient of a scalar-valued function. ● Relate Gradient to slope/derivative of a single-variable function ● Understand and relate Gradient to the direction of steepest ascent.
Jacobian and Hessian	● Introduce the concept of derivative of a vector by a vector ● Define and understand Jacobian ● Define hessian and relate it to gradient and jacobians
Other derivatives	● Differentiate a scalar by a vector (relate with gradient) ● Differentiate a scalar by a matrix ● Differentiate a matrix by a scalar ● Differentiate a vector by a scalar ● Differentiate a vector by a vector (relate with Jacobian)
Taylor Series for multivariable function	● Generalize the single variable Taylor Series for multivariable functions. ● Understand the role of Gradient and Hessian in the Quadratic Approximation of Taylor's series.

Unit 3.5 Optimization of multivariable Functions [2 hrs]

Update proposition: Remove Numerical Optimization

Students should be able to :

Stationary Points	● Relate gradient and Hessian with stationary points: minima, maxima and saddle points. ● Understand and visualize minimas, maximas and saddle points. ● Locate stationary points and analyze their nature (minima or maxima or saddle point). ● Relate definitiveness of the Hessian matrix with the nature of stationary points. ● Compute Simple convex optimization Problems Analytically. ● Examples where analytic solutions are not possible and motivate iterative methods.
Numerical Optimization	● Understand how to use gradient descent/ascent to compute minima/maxima[first order iterative optimization algorithm] ● *Understand how to use Newton's method for finding roots to compute minima/maxima[second order iterative optimization algorithm]*
Constrained Optimization	● Introduce Constraints to optimization problems with the help of Lagrange Multipliers. ● Grab intuition behind how the introduction of Lagrange multipliers changes constrained

	optimization to simple convex optimization. ● Compute simple constrained optimization problems analytically.

Unit 3.6: Integrals of multivariable functions [1 hr]

Update proposition: Add change of integrals in calculating multiple integrals

Students should be able to :

Double Integrals	● Understand the basic intuition of dependence/independence of variables in performing double integrals. ● Compute area under curves using double integrals. ● Examples relating to marginalizing probability distributions
Triple Integration	● Generalize concept of Double Integrals to triple Integrals

Unit 3.7 Module Summary

Students should be able to :

Module Summary	B. Summarise the overall concepts covered in the module

Additional Resources

1. Videos:
3blue 1 brown:
https://www.youtube.com/playlist?list=PL0-GT3co4r2wlh6UHTUeQsrf3mlS2lk6x
Khan Academy:
https://www.khanacademy.org/math/calculus-all-old
Khan Academy - Multivariable Calculus
https://www.youtube.com/playlist?list=PLSQl0a2vh4HC5feHa6Rc5c0wbRTx56nF7
MIT - Single variable Calculus
https://www.youtube.com/watch?v=7K1sB05pE0A
MIT - Multivariable Calculus
https://www.youtube.com/playlist?list=PL4C4C8A7D06566F38

2. Books:
Thomas' Calculus
http://www.maths.sci.ku.ac.th/suchai/417167/thomas.pdf

Module 4.   Probability and Statistics

Reference list:
MIT

https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/class-slides/
Khan Academy
https://www.khanacademy.org/math/statistics-probability
Stanford
http://cs229.stanford.edu/section/cs229-prob.pdf
Goodfellow
https://www.deeplearningbook.org/contents/prob.html
Ross
http://julio.staff.ipb.ac.id/files/2015/02/Ross_8th_ed_English.pdf
Radke videos:
https://www.youtube.com/watch?v=sa_ibR7Cqug&list=PLuh62Q4Sv7BU1dN2G6ncyiMbML7OXh_Jx
Crash Course
https://www.youtube.com/watch?v=JBlm4wnjNMY&list=PL8dPuuaLjXtNM_Y-bUAhblSAdWRnmBUcr&index=42
Statical inference:
https://fsalamri.files.wordpress.com/2015/02/casella_berger_statistical_inference1.pdf
Bishop:
http://users.isr.ist.utl.pt/~wurmd/Livros/school/Bishop%20-%20Pattern%20Recognition%20And%20Machine%20Learning%20-%20Springer%20%202006.pdf

Prerequisites
- Measure of central tendency: mean, median, mode
- Measure of dispersion: variance, standard deviation
- Other statistical measures: quartiles, range, percentile

* While providing the examples to the problems/concept, examples should be motivated with real world usage; which could also be reused to build up/relate the different concept

Unit 4.1. Introduction to Probability and Statistics

Students should be able to :

4.1.1. Introduction to the Module	A. Distinguish between different views/interpretation of probability: frequentism and subjectivism interpretation B. Describe relation and distinction between probability and statistics C. Discuss the importance and application of probability and statistics D. Describe a probabilistic model to involve: the sample space and probability law	- Life is uncertain, and we need to make decisions under uncertainty. - Example of uncertainty and how probability/statistics can be useful in relation with real-life, big-data and other disciplines - Describe what do we mean by a probabilistic model and what it consist of - Describe the relation between probability and statistics - Prerequisites needed - What students will learn with certainty Eg: https://youtu.be/MkUqWE-wUY0

Unit 4.2. Review and Introductory Concepts

Students should be able to :

4.2.1. Sample space and the Basic principle of counting	A. Define sample space as mutually exclusive and collectively exhaustive set of all possible outcomes of an experiment, defined at the right granularity
	B. Describe the basic principles of counting
	    a. addition principle of counting
	    b. multiplication principle of counting
	C. Define permutation and combination and solve simple problems
	D. Generalize combination to multinomial coefficients
	E. Compute simple problems(preferably real world example that can be reused while discussing the distribution and rules of probability) relating to multinomial coefficients
		- Define Sample space and each term in the definition with examples
		- Work out on sample space in both matrix and sequential (Tree) form

4.2.2. Set Theory and Real Numbers	A. Define set B. Define and perform different set operations using Venn diagram viz. Union, Intersection, Complement, Difference C. Use De Morgan's Laws to work on set algebraic problems D.	

4.2.3. Probability	A. Recall and use the axioms of probability to solve numerical problems
	    a. Non-Negativity
	    b. Normalization
	    c. Additivity and its corollaries
	B. Distinguish between discrete and continuous random variables with relation to sample space and events and how it is used to model real world experiments
	C. Recall and use discrete and continuous uniform law of probability
	D. Compute probability problems related to sample space and probability axioms	

4.2.4. Conditioning and independence	A. Understand the conditional probability and its use cases
	B. Derive Bayes Theorem by applying chain rule on conditional probability to solve conditional probability problems
	C. Describe each term (Posterior, Prior, Likelihood, and Evidence) in the Bayes rule and their significance
	D. Set up a model based on conditional probability and calculate
	    a. Composite probabilities $P(A \cap B)$
	    b. Total probability $P(B)$ from conditional using theorem of Total Probability
	    c. Inference $P(B\|A)$ using Bayes' Rule
	E. Define independence of two or **more events** and its properties
	    a. Symmetric properties
	    b. multiplication rule of probability for dependent and independent events
	F. Show distinction between independence and disjoint jets
	G. justify that independence in one model doesn't imply independence in a conditioned model and vice versa
	H. Justify that pairwise independence does not imply independence	

4.2.5. Statistics	A. Distinguish between population and sample	

	B. Distinguish between parameter and statistic C. Understand the concept of sufficient statistics and minimal sufficient statistics D. Define and understand the concept of estimators E. Understand the concept of law of large numbers F. Understand the concept of bias and distinguish between biased and unbiased estimators	

Unit 4.3. Discrete Random Variables

Update proposition: Add simple introduction and use cases of other discrete probability distributions.

4.3.1. Introduction to discrete random variables	A. discuss how **real world experiments** are modeled using parametric distributions of discrete random variables B. Understand the probability mass function(PMF) C. Understand cumulative distribution function(CDF) D. Understand the concept of expectation, variance and covariance E. Understand the LOTUS theorem and its use cases F. Formally define expectation, variance and covariance G. Compute numerical problems related to PMF, CDF, expectation, variance and covariance	
4.3.2. Bernoulli Random Variable	A. Understand the need for parametric model with some examples B. Understand the Bernoulli and Binomial model with real life examples C. Define formally Bernoulli and Binomial PMFs, CDFs D. Compute expectation and variance for Bernoulli and Binomial random variables	
4.3.3. Multinomial Random Variable	A. Understand the multinomial model with real life examples B. Generalize the binomial random variable to multinomial random variable C. Generalize the concept of PMFs, CDFs of binomial to multinomial random variable D. Understand expectation and variance of multinomial random variable	

Unit 4.4. Continuous Random Variable

4.4.1. Introduction to continuous	A. Recall how real world experiments are modeled using continuous random variables B. Understand the probability density function(PDF)	

random variables	and distinguish between PMF and PDF C. Understand cumulative distribution function(CDF) for continuous random variables D. Understand the concept of expectation, variance and covariance for continuous random variables E. Formally define expectation, variance and covariance continuous random variables F. Compute numerical problems related to PDF, CDF, expectation, variance and covariance	
4.4.2. Gaussian Distribution	A. Understand the gaussian model with real life examples B. Define formally PDFs, CDFs of gaussian C. Understand the expectation and variance for gaussian distribution D. Compute the expectation and variance for gaussian distribution *E. Understand the limitation of gaussian model to modeling data and relate how gaussian mixture model could be used instead*	
4.4.3. Other distributions	A. Understand the PMFs and CDFs of uniform, beta and gamma distribution and list other distribution B. Use formulae to calculate expectation and variance of beta and gamma distribution	

Unit 4.5. Multiple Random Variables

4.5.1. Marginal Probability, Joint Probability	A. Understand the concept of marginal probability and its use cases B. Compute marginal probability C. Understand the real life use case of a jointly distributed random variable D. Relate joint probability with conditional and marginal probability reinforcing the concept of sum and product rule E. Understand law of total probability and its use to compute marginal probability F. Relate conditional, marginal and joint probability to conditional, marginal and joint distribution G. Compute numerical problems related to marginal and joint probability	
4.5.2. Bayes' Rule	A. Understand and derive Bayes' rule B. Interpret Bayes rule in terms of likelihood, prior, posterior, and model evidence C. Compute numerical problems related to Bayes' rule	
4.5.3. Sum of random variables and Central Limit Theorem	A. Understand the properties of a sum of random variables B. Understand and formally define central limit theorem C. Use central limit theorem in problems concerning sampling distribution of sample	

	mean D. Use central limit theorem in problems concerning sampling distribution of sampling proportion	

Unit 4.6. Basics of hypothesis testing

4.6.1. Introduction to hypothesis testing	A. Understand the need for hypothesis testing with some real world examples B. Understand null hypothesis, alternative hypothesis, one-tailed tests and two-tailed tests C. Formulate the null and alternative hypotheses for both one-tailed and two-tailed test	
4.6.2. P-values	A. Understand p-values, significance level, acceptance region, rejection region B. Compare p-values for different significance levels C. Make conclusions on the hypothesis using p-values D. Carry out hypothesis testing on cases where p-values are already known	
4.6.3. Errors	A. Understand Type-I and Type-II errors in relation to hypothesis testing B. Distinguish between Type-I and Type-II errors C. Understand the consequences of errors and significance	
4.6.4. Hypothesis testing on population proportion	A. Formulate null and alternative hypothesis in cases about population proportion B. Understand z-statistics and compute a z-statistic in a test about a proportion C. Compute a p-value given z-statistic D. Make conclusion using p-value and significance level for a test	

Unit 4.7. Module Summary

Students should be able to :

4.7.1. Module Summary	A. Summarise the overall concepts covered in the module	
Module Project		

5. Information Theory

Reference list:
https://www.khanacademy.org/computing/computer-science/informationtheory
http://www.deeplearningbook.org/contents/prob.html

Chain Rule for relative entropy
https://web.stanford.edu/class/ee376a/files/2017-18/lecture_20.pdf

Course Outline:

Unit 5.1 Introduction to the module

Students should be able to :

Introduction to the Module	• Understand what they are expected to learn out of the module.

Unit 5.2 Measure of information

Students should be able to :

Introduction to Measure of information	• Understand information as a degree of surprise of occurence of an event • Describe the criteria to be satisfied by measure of information and formally define it(self-information) and understand the real world application via examples • Understand the concept of coding as a mechanism to encode information with a real world example

Entropy	• Understand the concept of entropy with real world example
	• Understand the relation between entropy and shortest coding length(Noiseless Coding Theorem)
	    ○ Relate shortest coding length with compression
	    ○ Relate compression with dimension reduction conceptually
	• Formally define entropy for discrete random variable and describe the condition for maximality
	• Formally define entropy(differential entropy) for continuous random variable and describe the condition for maximality
	• Understand how to compute joint entropy for jointly distributed random variable
	• Understand how to compute conditional entropy (equivocation) for conditional random variable
	• Understand the relation between joint entropy and conditional entropy

Cross Entropy, Relative Entropy and Mutual Information	● Understand the need of cross-entropy with some real-world example preferably as a loss function in ML ● Formally define the cross-entropy ● Describe the binary cross entropy as a special case of cross entropy and its use case ● Understand the need of relative entropy(Kullback–Leibler[KL] divergence)(preferably to measure the similarity of a distribution) and define it formally ● Understand the properties of KL divergence ● Understand the relation between KL divergence and cross entropy ● Understand the need for mutual information with how close are two random variables are to being independent ● Formally define the mutual information and understand its properties ● Understand the relation between conditional entropy and mutual information ● Understand Jensen-Shanon divergence and its properties ● Understand the relation between mutual information and Jensen-Shanon divergence

6. Numerical Computation

Course Outline:

Unit 6.1 Introduction to the module

Students should be able to :

Introduction to the Module	• Understand what they are expected to learn out of the module.

Unit 6.2 Solving nonlinear equations

Students should be able to :

Root findings	• Understand the need for root finding problem of an equation with real world use case and the necessity of iterative methods for the task
	• Understand how to compute the roots with
	  ○ *False Positive Method*
	  ○ **Bisection Method and it's convergence**
	  ○ **Newton Raphson Method and it's convergence**
	  ○ *Secant Method*

Unit 6.3 Numerical Linear Algebra

Students should be able to :

Solution to the system of linear equation	• Understand Gauss Elimination and Gauss Jordan as a solution to the system of linear equation along with its pitfalls(division by zero, round-off errors, ill-conditioned systems) and potential solution to the pitfalls • Understand the LU/LDU decomposition as the matrix used for gauss elimination • **Understand the limitations of the above methods and the need for iterative methods for finding solutions to systems of linear equations** • Understand jacobi and gauss-seidel method to solve a system of linear equations and their convergence
Inverse of a matrix	• Understand how to compute the inverse of a matrix with Gauss-Jordan Elimination and LU decomposition
Orthogonalization	• Recall the importance of orthogonal matrices • Understand Gram-Schmidt process to

	orthogonalize the matrices

Unit 6.4 Module Summary

Students should be able to :

Module Summary	B. Summarise the overall concepts covered in the module