List and Syllabi of Courses for B.Tech. (Minor) in Data Science and AI

Mehta Family School of Data Science and Artificial Intelligence

1. List of the courses (Semester-wise) offered to B.Tech. (Minor) in Data Science and AI

S.No.	Course Number and Title	Semester
1.	DA241M & Mathematical Foundations of Data Science	3rd
2.	DA221M & Artificial Intelligence	4th
3.	DA321M & Machine Learning	5th
4.	DA322M & Deep Learning	6th
5.	DA421M & Multi-modal Data Processing & Learning	7th

2. Syllabi of the courses offered to B.Tech. (Minor) in Data Science and AI

DA241M & Mathematical Foundations of Data Science

Course Number & Title: DA241M & Mathematical foundations of data science
L-T-P-C: 3-0-0-6
Type of Letter Grading (Regular Letter Grades / PP or NP Letter Grades): Regular Letter Grades
Kind of Proposal (New Course / Revision of Existing Course): New Course
Offered as (Compulsory / Elective): Minor
Offered to: B.Tech. (Minor) in Data Science and AI
Offered in (Odd/ Even / Any): Odd
Offered by (Name of Department/ Center): Mehta Family School of Data Science and Artificial Intelligence
Pre-Requisite: None
Preamble / Objectives (Optional):
Course Content/ Syllabus Fundamentals: Review of linear algebra and multi-variate calculus from data science perspective. Probability and random variables: Basics of Probability Theory, Conditional Probability, Bayes' Theorem, Random Variables, Discrete and Continuous Distributions, Moments, Law of large numbers and Central Limit Theorem. Statistical inference: Parametric and nonparametric methods, Point estimation, Confidence Intervals, Maximum Likelihood Estimators; Hypothesis testing; Bayesian Inference. Optimization: Unconstrained and Constrained optimization for single and multiple variables: Gradient descent methods, Newton's method, Simplex method; Convexity and duality
Books (In case UG compulsory courses, please give it as "Text books" and "Reference books". Otherwise give it as "References".
Texts: (Format: Authors, *Book Title in Italics font*, Volume/Series, Edition Number, Publisher, Year.)

1.	M.H. DeGroot, and M. J. Schervish, *Probability and statistics*, 4th Ed., Pearson Education, 2010
2.	E. K. P. Chong and S. H. Zak, *An Introduction to Optimization*, 4th Ed., Wiley India Pvt. Ltd., 2017

References: (Format: Authors, *Book Title in Italics font*, Volume/Series, Edition Number, Publisher, Year.)

1.	Wasserman, L., *All of statistics: a concise course in statistical inference*, 1st Ed., Springer, New York, 2004.
2.	Strang, Gilbert. *Linear algebra and learning from data*. Cambridge: Wellesley-Cambridge Press, 2019

Detailed Course Content (Optional)		
It will not be included in the Courses of Study Booklet		
Sl. No.	Broad Title / Topics	Number of Lectures
1	Fundamentals	5
2	Probability and Random variables	12

3	Statistical Inference	11
4	Optimization	12
Total Number of Lectures =		40

DA221M & Artificial Intelligence

Course Number & Title: DA221M & Artificial Intelligence
L-T-P-C: 3-0-0-6
Type of Letter Grading (Regular Letter Grades / PP or NP Letter Grades): Regular Letter Grades
Kind of Proposal (New Course / Revision of Existing Course): New Course
Offered as (Compulsory / Elective): Minor
Offered to: B.Tech. (Minor) in Data Science and AI
Offered in (Odd/ Even / Any): Even
Offered by (Name of Department/ Center): Mehta Family School of Data Science and Artificial Intelligence
Pre-Requisite: None
Preamble / Objectives (Optional):
Course Content/Syllabus Introduction to AI and Intelligent Agents; Problem solving by Searching: Uninformed and informed strategies; Logical Agents: Propositional and first order logic, inference; Knowledge representation and Automated Planning; Uncertain Knowledge and Reasoning: Quantifying uncertainty, probabilistic reasoning. Introduction to Learning: Supervised Learning, Unsupervised Learning, Reinforcement Learning; Markov Process: Discrete-time Markov chain, Stationary Distribution; Markov Decision Process: Dynamic Programming, Finite Horizon MDP, Infinite Horizon MDP; Value Iteration Algorithm, Policy Iteration Algorithm; Multi-armed Bandit, Application & Case Studies; Q-learning; TD learning; Policy Gradient.
Books (In case UG compulsory courses, please give it as "Text books" and "Reference books". Otherwise give it as "References".

Texts: (Format: Authors, *Book Title in Italics font*, Volume/Series, Edition Number, Publisher, Year.)

1.	S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th Ed., Pearson, 2020
2.	R.S. Sutton and A.G. Barto, *Reinforcement Learning: An Introduction*, 2nd Ed., MIT Press, 2018.

References: (Format: Authors, *Book Title in Italics font*, Volume/Series, Edition Number, Publisher, Year.)

1.	E. Rich, K. Knight, and S. B. Nai, *Artificial Intelligence*, McGraw Hill, 3rd Ed., 2017.
2.	D. Khemani, *A First Course in Artificial Intelligence*, 1st Ed., McGraw-Hill Education, 2017.

Detailed Course Content (Optional) It will not be included in the Courses of Study Booklet		
Sl. No.	Broad Title / Topics	Number of Lectures
1	Introduction to AI and Intelligent agents	1
2	Problem solving by Searching	5
3	Logical Agents	5
4	Knowledge representation and Automated planning	4
5	Uncertain Knowledge and Reasoning	5
6	Introduction to Learning	1
7	Markov Process	4
8	Markov Decision Process	2
9	Finite Horizon MDP	2
10	Infinite Horizon MDP	1
11	Value and Policy Iteration Algorithm	3
12	Multi-armed Bandit, Application and Case Studies	3
13	Q-Learning	2
14	TD Learning	1
15	Policy Gradient	1
Total Number of Lectures =		40

DA321M & Machine Learning

Course Number & Title: DA321M & Machine Learning
L-T-P-C: 3-0-0-6

Type of Letter Grading (Regular Letter Grades / PP or NP Letter Grades): Regular Letter Grades
Kind of Proposal (New Course / Revision of Existing Course): New Course
Offered as (Compulsory / Elective): Minor
Offered to: B.Tech. (Minor) in Data Science and AI
Offered in (Odd/ Even / Any): Odd
Offered by (Name of Department/ Center): Mehta Family School of Data Science and Artificial Intelligence.
Pre-Requisite: None
Preamble / Objectives (Optional):
Course Content/ Syllabus:

Introduction to learning: supervised and unsupervised, generative and discriminative models, classification and regression problems, performance measures, design of experiments; Feature space and dimensionality reduction: distance measures, PCA, LDA; Unsupervised learning: K-means clustering, hierarchical agglomerative clustering, EM algorithm, Mixture model; Supervised learning: Bayesian classification, linear and logistic regression, simple perceptron and multi-layer perceptron, Parzen windows, k-nearest neighbor, decision trees, support vector machines; Hidden Markov models; Applications and case studies.

Books (In case UG compulsory courses, please give it as "Text books" and "Reference books". Otherwise give it as "References".

Texts: (Format: Authors, *Book Title in Italics font*, Volume/Series, Edition Number, Publisher, Year.)

1.	E. Alpaydin, *Introduction to Machine Learning*, 4th Ed., Prentice Hall (India) 2020
2.	R. O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification*, 2nd Ed., Wiley India, 2007

References: (Format: Authors, *Book Title in Italics font*, Volume/Series, Edition Number, Publisher, Year.)

1.	C. M. Bishop, *Pattern Recognition and Machine Learning*, Information Science and Statistics, Springer, 2016
2.	S. O. Haykin, *Neural Networks and Learning Machines*, 3rd Ed., Pearson Education (India), 2016

Detailed Course Content (Optional)
It will not be included in the Courses of Study Booklet

Sl. No.	Broad Title / Topics	Number of Lectures
1	Introduction to learning	1
2	Performance measures, design of experiments	3
3	Feature space and dimensionality reduction	4
4	Unsupervised learning	8
5	Supervised learning	15
6	Hidden Markov Model	3
7	Applications and case studies	6
Total Number of Lectures =		40

DA322M & Deep Learning

Course Number & Title: DA322M & Deep Learning
L-T-P-C: 3-0-0-6
Type of Letter Grading (Regular Letter Grades / PP or NP Letter Grades): Regular Letter Grades
Kind of Proposal (New Course / Revision of Existing Course): New Course
Offered as (Compulsory / Elective): Minor
Offered to: B.Tech. (Minor) in Data Science & AI
Offered in (Odd/ Even / Any): Even
Offered by (Name of Department/ Center): Mehta Family School of Data Science and Artificial Intelligence
Pre-Requisite: None
Preamble / Objectives (Optional):
Course Content/ Syllabus:

Introduction: Engineered and learned features, discriminative models, decision surfaces, shallow and deep learning; Feature extraction: Correlation, cross-correlation, auto-correlation, convolution; Revisiting MLP: Multilayer perceptron, back-propagation learning; Activation functions; Loss functions; Optimization techniques:

Stochastic gradient descent, batch optimization, momentum optimizer, RMSProp, Adam; Autoencoders; Convolutional Neural Network: Building blocks of CNN, vanishing and exploding gradient problems; Popular CNN architectures: LeNet, AlexNet, VGGNet, ResNet skip connections, inception blocks; Training issues: Early stopping, dropout, batch normalization, instance normalization, group normalization; Recurrent Neural Networks and variants; Applications of Deep Networks.

Books (In case UG compulsory courses, please give it as "Text books" and "Reference books". Otherwise give it as "References".

Texts: (Format: Authors, *Book Title in Italics font*, Volume/Series, Edition Number, Publisher, Year.)

| 1. | I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, MIT Press, 2016 |
| 2. | M. A. Nielsen, *Neural Networks and Deep Learning*, Determination Press, 2015 |

References: (Format: Authors, *Book Title in Italics font*, Volume/Series, Edition Number, Publisher, Year.)

| 1. | A. Zhang, Z. C. Lipton, M. Li, A. J. Smola, *Dive into Deep Learning*, 2021 (Available online at https://d2l.ai/index.html) |
| 2. | Y. Bengio, *Learning Deep Architectures for AI*, Now Publishers Inc., 2009 |

Detailed Course Content (Optional)
It will not be included in the Courses of Study Booklet

Sl. No.	Broad Title / Topics	Number of Lectures
1	Introduction	1
2	Feature extraction	2
3	Revisiting MLP	6
4	Optimization techniques	3
5	Autoencoders	3
6	Convolutional Neural Network	3
7	Popular CNN architectures	8
8	Training issues	5
9	Recurrent neural network and variants	4
11	Applications of Deep Networks	5
Total Number of Lectures =		40

DA421M & Multi-Modal Data Processing & Learning

Course Number & Title: DA421M & Multi-Modal Data Processing & Learning
L-T-P-C: 3-0-0-6
Type of Letter Grading (Regular Letter Grades / PP or NP Letter Grades): Regular Letter Grades
Kind of Proposal (New Course / Revision of Existing Course): New Course
Offered as (Compulsory / Elective): Minor
Offered to: B.Tech. (Minors) in Data Science and AI
Offered in (Odd/ Even / Any): Odd
Offered by (Name of Department/ Center): Mehta Family School of Data Science and Artificial Intelligence
Pre-Requisite: None
Preamble / Objectives (Optional): Primary aim of the course is to teach essential processing required when dealing with data involving text, speech, image and video. In particular, this course focuses on *text processing for information retrieval perspective, text processing for general natural language processing perspective, digital image and video processing, and speech processing.*
Course Content/ Syllabus Text: Natural Language Processing – Text normalization: subword tokenization, lemmatization, morphology; Language models and smoothing techniques; Vector space models. Text: Information Retrieval- Introduction: Text processing and statistics, Document parsing, Inverted index; Retrieval and Ranking: TFIDF, BM-25, Binary independent model, Page rank, HITS, Query Expansion; Evaluation methods.

Speech Processing - Speech production and perception, Acoustic and articulatory phonetics; Audio and Speech signal processing.

Digital Image and Video Processing - Image/video acquisition and perception; Basic image processing operations; Image and Video features; Motion estimation; Applications of image and video processing.

Learning with multi-modal data: VQA, Emotion Recognition etc.

Books (In case UG compulsory courses, please give it as "Text books" and "Reference books". Otherwise give it as "References".

Texts: (Format: Authors, *Book Title in Italics font*, Volume/Series, Edition Number, Publisher, Year.)

1.	W. B. Croft, D. Metzler, T. Strohman, *Search Engines Information Retrieval in Practice*, Pearson, 2015 (Online available at https://ciir.cs.umass.edu/irbook/)
2.	C. J. Chen, *Elements of Human Voice*, World Scientific Publishing, 2016
3.	M. Sonka, V. Hlavac, R. Boyle, *Image Processing, Analysis and Machine Vision*, 4th Ed., Cengage, 2017

References: (Format: Authors, *Book Title in Italics font*, Volume/Series, Edition Number, Publisher, Year.)

1.	D. Jurafsky, J.H. Martin, *Speech and Language Processing*, 3rd Ed. 2022 (online available at https://web.stanford.edu/~jurafsky/slp3/)
2.	C. D. Manning, P Raghavan, H Schutz, *Introduction to Information Retrieval*, Cambridge University Press, 2008 (Online available at https://nlp.stanford.edu/IR-book/information-retrieval-book.html)
3.	T. F. Quatieri, *Discrete-Time Processing of Speech Signals*, Pearson Education, 2005
4.	L. R. Rabiner, R. W. Schafer, *Digital Processing of Speech Signals*, Pearson Education, 2004
5.	D. O'Shaughnessy, *Speech Communications: Human and Machine*, 2nd Ed. University Press, 2005
6.	R Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2022
7.	M. K. Bhuyan, *Computer Vision and Image Processing – Fundamentals and Applications*, CRC Press, USA, 2019
8.	Y. Wang, J. Ostermann, Y. Q. Zhang, *Video Processing and Communications*, Prentice Hall, 2002
9.	R. C. Gonzalez, R. E. Woods, *Digital Image Processing*, Prentice Hall, 3rd Ed. 2008

Detailed Course Content (Optional)		
It will not be included in the Courses of Study Booklet		
Sl. No.	Broad Title / Topics	Number of Lectures
1	Text: Natural Language Processing	8
2	Text: Information Retrieval	8
3	Speech Processing	9
4	Digital Image and Video Processing	12
5	Learning with Multi-modal Data	3
Total Number of Lectures =		40