# Applied Machine Learning Course Schedule

| DATE | MODULE | CHAPTER | TOPIC |
|---|---|---|---|
| 2021-09-01 | Module 1:Fundamentals of Programming | Python for DataScience | Keywords and identifiers, comments, indentation and statements, Variables and data types in Python, Standard Input and Output, Operators, Control flow: if else, Control flow: while loop |
| 2021-09-02 | Module 1:Fundamentals of Programming | Python for DataScience | Control flow: for loop, Control flow: break and continue,Revision Python for DataScience |
| 2021-09-03 | Module 1:Fundamentals of Programming | Python for DataScience:Data Structures | Lists, Tuples part 1, Tuples part-2, Sets, Dictionary, Strings |
| 2021-09-04 | Module 1:Fundamentals of Programming | Python for DataScience:Data Structures | Revision Python for DataScience:Data Structures |
| 2021-09-05 | Module 1:Fundamentals of Programming | Python for DataScience:Functions | Introduction, Types of functions, Function arguments, Recursive functions, Lambda functions, Modules, Packages, File Handling, Exception Handling, Debugging Python,Revision Python for DataScience:Functions |
| 2021-09-06 | Module 1:Fundamentals of Programming | Python for DataScience:Numpy | Numpy Introduction, Numerical operations on Numpy,Revision Python for DataScience:Numpy |
| 2021-09-07 | Module 1:Fundamentals of Programming | Python for DataScience:Matplotlib | Getting started with Matplotlib,Revision Python for DataScience:Matplotlib |
| 2021-09-08 | Module 1:Fundamentals of Programming | Python for DataScience:Pandas | Getting started with pandas, Data Frame Basics, Key Operations on Data Frames,Revision Python for DataScience:Pandas |

| Date | Module | Topic | Details |
|---|---|---|---|
| 2021-09-09 | Module 1:Fundamentals of Programming | Python for DataScience:Compputational Complexity | Space and Time Complexity: Find largest number in a list , Binary search, Find elements common in two lists, Find elements common in two lists using a Hashtable/Dict,Revision Python for DataScience:Compputational Complexity |
| 2021-09-10 | Module 1:Fundamentals of Programming | SQL | Introduction to Databases, Why SQL?, Execution of an SQL statement., IMDB dataset, Installing MySQL, Load IMDB data., USE, DESCRIBE, SHOW TABLES |
| 2021-09-11 | Module 1:Fundamentals of Programming | SQL | SELECT , LIMIT, OFFSET, ORDER BY, DISTINCT , WHERE, Comparison operators, NULL, Logical Operators, Aggregate Functions: COUNT, MIN, MAX, AVG, SUM, GROUP BY, HAVING, Order of keywords., Join and Natural Join, Inner, Left, Right and Outer joins., Sub Queries/ Nested Queries/Inner Queries |
| 2021-09-12 | Module 1:Fundamentals of Programming | SQL | DML:INSERT, DML:UPDATE , DELETE, DDL:CREATE TABLE, DDL:ALTER: ADD, MODIFY, DROP, DDL:DROP TABLE, TRUNCATE, DELETE, Data Control Language: GRANT, REVOKE, Learning resources,Revision SQL |
| 2021-09-13 | Module 2: Datascience: Exploratory Data Analysis and Data Visualization | Plotting for exploratory data analysis (EDA) | Introduction to IRIS dataset and 2D scatter plot, 3D scatter plot, Pair plots, Limitations of Pair Plots, Histogram and Introduction to PDF(Probability Density Function), Univariate Analysis using PDF, CDF(Cumulative Distribution Function), Mean, Variance and Standard Deviation |

| | | | |
|---|---|---|---|
| **2021-09-14** | Module 2: Datascience: Exploratory Data Analysis and Data Visualization | Plotting for exploratory data analysis (EDA) | Median, Percentiles and Quantiles, IQR(Inter Quartile Range) and MAD(Median Absolute Deviation), Box-plot with Whiskers, Violin Plots, Summarizing Plots, Univariate, Bivariate and Multivariate analysis, Multivariate Probability Density, Contour Plot, Exercise: Perform EDA on Haberman dataset |
| **2021-09-15** | Module 2: Datascience: Exploratory Data Analysis and Data Visualization | Plotting for exploratory data analysis (EDA) | Revision Plotting for exploratory data analysis (EDA) |
| **2021-09-16** | Module 2: Datascience: Exploratory Data Analysis and Data Visualization | Linear Algebra | Why learn it ?, Introduction to Vectors(2-D, 3-D, n-D) , Row Vector and Column Vector, Dot Product and Angle between 2 Vectors, Projection and Unit Vector, Equation of a line (2-D), Plane(3-D) and Hyperplane (n-D), Plane Passing through origin, Normal to a Plane, Distance of a point from a Plane/Hyperplane, Half-Spaces, Equation of a Circle (2-D), Sphere (3-D) and Hypersphere (n-D), Equation of an Ellipse (2-D), Ellipsoid (3-D) and Hyperellipsoid (n-D), Square ,Rectangle, Hyper Cube,Hyper Cuboid |
| **2021-09-17** | Module 2: Datascience: Exploratory Data Analysis and Data Visualization | Linear Algebra | Revision Questions,Revision Linear Algebra |

| | | | |
|---|---|---|---|
| **2021-09-18** | Module 2: Datascience: Exploratory Data Analysis and Data Visualization | Probability And Statistics | Introduction to Probability and Statistics, Population and Sample, Gaussian/Normal Distribution and its PDF(Probability Density Function), CDF(Cumulative Distribution function) of Gaussian/Normal distribution, Symmetric distribution, Skewness and Kurtosis, Standard normal variate (Z) and standardization, Kernel density estimation, Sampling distribution & Central Limit theorem, Q-Q plot:How to test if a random variable is normally distributed or not?, How distributions are used?, Chebyshev's inequality, Discrete and Continuous Uniform distributions, How to randomly sample data points (Uniform Distribution) |
| **2021-09-19** | Module 2: Datascience: Exploratory Data Analysis and Data Visualization | Probability And Statistics | Bernoulli and Binomial Distribution, Log Normal Distribution, Power law distribution, Box cox transform, Applications of non-gaussian distributions?, Co-variance, Pearson Correlation Coefficient, Spearman Rank Correlation Coefficient, Correlation vs Causation, How to use correlations? , Confidence interval (C.I) Introduction, Computing confidence interval given the underlying distribution, C.I for mean of a normal random variable, Confidence interval using bootstrapping, Hypothesis testing methodology, Null-hypothesis, p-value |

| | | | |
|---|---|---|---|
| **2021-09-20** | Module 2: Datascience: Exploratory Data Analysis and Data Visualization | Probability And Statistics | Hypothesis Testing Intution with coin toss example, Resampling and permutation test, K-S Test for similarity of two distributions, Code Snippet K-S Test, Hypothesis testing: another example, Resampling and Permutation test: another example |
| **2021-09-21** | Module 2: Datascience: Exploratory Data Analysis and Data Visualization | Probability And Statistics | How to use hypothesis testing?, Propotional sampling, Revision Questions, Assignment :Python (without Numpy) |
| **2021-09-22** | Module 2: Datascience: Exploratory Data Analysis and Data Visualization | Probability And Statistics | Revision Probability And Statistics |
| **2021-09-23** | Module 2: Datascience: Exploratory Data Analysis and Data Visualization | Dimensionality Reduction And Visualization | What is Dimensionality reduction?, Row Vector and Column Vector, How to represent a data set?, How to represent a dataset as a Matrix., Data Preprocessing: Feature Normalisation, Mean of a data matrix, Data Preprocessing: Column Standardization, Co-variance of a Data Matrix, MNIST dataset (784 dimensional) |
| **2021-09-24** | Module 2: Datascience: Exploratory Data Analysis and Data Visualization | Dimensionality Reduction And Visualization | Code to Load MNIST Data Set,Revision Dimensionality Reduction And Visualization |

| | | | |
|---|---|---|---|
| **2021-09-25** | Module 2: Datascience: Exploratory Data Analysis and Data Visualization | Principal Component Analysis | Why learn PCA?, Geometric intuition of PCA, Mathematical objective function of PCA, Alternative formulation of PCA: Distance minimization, Eigen values and Eigen vectors (PCA): Dimensionality reduction, PCA for Dimensionality Reduction and Visualization, Visualize MNIST dataset, Limitations of PCA, PCA Code example, PCA for dimensionality reduction (not-visualization),Revision Principal Component Analysis |
| **2021-09-26** | Module 2: Datascience: Exploratory Data Analysis and Data Visualization | T-Sne | What is t-SNE?, Neighborhood of a point, Embedding, Geometric intuition of t-SNE, Crowding Problem, How to apply t-SNE and interpret its output, t-SNE on MNIST, Code example of t-SNE, Revision Questions,Revision T-Sne |
| **2021-09-27** | Module 3: Foundations of Natural Language Processing and Machine Learning | Predict rating given product reviews on amazon | Dataset overview: Amazon Fine Food reviews(EDA), Data Cleaning: Deduplication, Why convert text to a vector?, Bag of Words (BoW), Text Preprocessing: Stemming, Stop-word removal, Tokenization, Lemmatization., uni-gram, bi-gram, n-grams. |
| **2021-09-28** | Module 3: Foundations of Natural Language Processing and Machine Learning | Predict rating given product reviews on amazon | tf-idf (term frequency-inverse document frequency), Why use log in IDF?, Word2Vec., Avg-Word2Vec, tf-idf weighted Word2Vec, Bag of Words( Code Sample), Text Preprocessing( Code Sample) |

| | | | |
|---|---|---|---|
| **2021-09-29** | Module 3: Foundations of Natural Language Processing and Machine Learning | Predict rating given product reviews on amazon | Bi-Grams and n-grams (Code Sample), TF-IDF (Code Sample), Assignment :Implementing TFIDF vectorizer |
| **2021-09-30** | Module 3: Foundations of Natural Language Processing and Machine Learning | Predict rating given product reviews on amazon | Word2Vec (Code Sample), Avg-Word2Vec and TFIDF-Word2Vec (Code Sample),Revision Predict rating given product reviews on amazon |
| **2021-10-01** | Module 3: Foundations of Natural Language Processing and Machine Learning | Classification And Regression Models: K-Nearest Neighbors | How "Classification" works?, Data matrix notation, Classification vs Regression (examples), K-Nearest Neighbours Geometric intuition with a toy example, Failure cases of KNN, Distance measures: Euclidean(L2) , Manhattan(L1), Minkowski, Hamming, Cosine Distance & Cosine Similarity, How to measure the effectiveness of k-NN? |
| **2021-10-02** | Module 3: Foundations of Natural Language Processing and Machine Learning | Classification And Regression Models: K-Nearest Neighbors | Test/Evaluation time and space complexity, KNN Limitations, Decision surface for K-NN as K changes, Overfitting and Underfitting, Need for Cross validation, K-fold cross validation, Visualizing train, validation and test datasets, How to determine overfitting and underfitting?, Time based splitting, k-NN for regression, Weighted k-NN, Voronoi diagram, Binary search tree, How to build a kd-tree |

| | | | |
|---|---|---|---|
| **2021-10-03** | Module 3: Foundations of Natural Language Processing and Machine Learning | Classification And Regression Models: K-Nearest Neighbors | Find nearest neighbours using kd-tree, Limitations of Kd tree, Extensions, Hashing vs LSH, LSH for cosine similarity, LSH for euclidean distance, Probabilistic class label, Code Sample:Decision boundary ., Code Sample:Cross Validation, Assignment : Implement RandomSearchCV with k fold cross validation on KNN |
| **2021-10-04** | Module 3: Foundations of Natural Language Processing and Machine Learning | Classification And Regression Models: K-Nearest Neighbors | Question and Answers,Revision Classification And Regression Models: K-Nearest Neighbors |
| **2021-10-05** | Module 3: Foundations of Natural Language Processing and Machine Learning | Classification Algorithms in Various Situations | Introduction, Imbalanced vs balanced dataset, Multi-class classification, k-NN, given a distance or similarity matrix, Train and test set differences, Impact of outliers, Local outlier Factor (Simple solution :Mean distance to Knn) |
| **2021-10-06** | Module 3: Foundations of Natural Language Processing and Machine Learning | Classification Algorithms in Various Situations | K-Distance(A),N(A), Reachability-Distance(A,B), Local reachability-density(A), Local outlier Factor(A), Impact of Scale & Column standardization, Interpretability, Feature Importance and Forward Feature selection, Handling categorical and numerical features |
| **2021-10-07** | Module 3: Foundations of Natural Language Processing and Machine Learning | Classification Algorithms in Various Situations | Handling missing values by imputation, Curse of dimensionality, Bias-Variance tradeoff, Intuitive understanding of bias-variance., Best and worst cases for an algorithm, Question and Answers |

| | | | |
|---|---|---|---|
| **2021-10-08** | Module 3: Foundations of Natural Language Processing and Machine Learning | Classification Algorithms in Various Situations | Revision Classification Algorithms in Various Situations |
| **2021-10-09** | Module 3: Foundations of Natural Language Processing and Machine Learning | Performance Measurement of Models | Accuracy, Confusion matrix, TPR, FPR, FNR, TNR, Precision and recall, F1-score, Receiver Operating Characteristic Curve (ROC) curve and AUC, Log-loss, R-Squared/Coefficient of determination, Median absolute deviation (MAD), Distribution of errors, Assignment:Compute Performance metrics without Sklearn |
| **2021-10-10** | Module 3: Foundations of Natural Language Processing and Machine Learning | Performance Measurement of Models | Revision Performance Measurement of Models |
| **2021-10-11** | Module 3: Foundations of Natural Language Processing and Machine Learning | Naive Bayes | Conditional probability, Independent vs Mutually exclusive events, Bayes Theorem with examples, Exercise problems on Bayes Theorem, Naive Bayes algorithm |
| **2021-10-12** | Module 3: Foundations of Natural Language Processing and Machine Learning | Naive Bayes | Toy example: Train and test stages, Naive Bayes on Text data, Laplace/Additive Smoothing, Log-probabilities for numerical stability, Bias and Variance tradeoff |

| | | | |
|---|---|---|---|
| **2021-10-13** | Module 3: Foundations of Natural Language Processing and Machine Learning | Naive Bayes | Feature importance and interpretability, Imbalanced data, Outliers, Missing values, Handling Numerical features (Gaussian NB), Multiclass classification, Similarity or Distance matrix, Large dimensionality, Best and worst cases, Code example, Assignment: Apply Multinomial NB on Donors Choose Dataset |
| **2021-10-14** | Module 3: Foundations of Natural Language Processing and Machine Learning | Naive Bayes | Revision Naive Bayes |
| **2021-10-15** | Module 3: Foundations of Natural Language Processing and Machine Learning | Logistic Regression | Geometric intuition of Logistic Regression, Sigmoid function: Squashing, Mathematical formulation of Objective function |
| **2021-10-16** | Module 3: Foundations of Natural Language Processing and Machine Learning | Logistic Regression | Weight vector, L2 Regularization: Overfitting and Underfitting, L1 regularization and sparsity, Probabilistic Interpretation: Gaussian Naive Bayes, Loss minimization interpretation, Hyperparameter search: Grid Search and Random Search, Column Standardization, Feature importance and Model interpretability, Collinearity of features, Test/Run time space and time complexity, Real world cases, Non-linearly separable data & feature engineering |

| | | | |
|---|---|---|---|
| **2021-10-17** | Module 3: Foundations of Natural Language Processing and Machine Learning | Logistic Regression | Code sample: Logistic regression, GridSearchCV, RandomSearchCV, Extensions to Logistic Regression: Generalized linear models,Revision Logistic Regression |
| **2021-10-18** | Module 3: Foundations of Natural Language Processing and Machine Learning | Linear Regression | Geometric intuition of Linear Regression, Mathematical formulation, Real world Cases, Code sample for Linear Regression, Question and Answers,Revision Linear Regression |
| **2021-10-19** | Module 3: Foundations of Natural Language Processing and Machine Learning | Solving Optimization Problems | Differentiation, Online differentiation tools, Maxima and Minima, Vector calculus: Grad, Gradient descent: geometric intuition, Learning rate, Gradient descent for linear regression |
| **2021-10-20** | Module 3: Foundations of Natural Language Processing and Machine Learning | Solving Optimization Problems | SGD algorithm, Constrained Optimization & PCA, Logistic regression formulation revisited, Why L1 regularization creates sparsity?, Assignment : Implement SGD Classifier with Log Loss and L2 regularization Using SGD: without using sklearn |
| **2021-10-21** | Module 3: Foundations of Natural Language Processing and Machine Learning | Solving Optimization Problems | Revision Solving Optimization Problems |
| **2021-10-22** | Module 4: Machine Learning-II (Supervised Learning Models) | Support Vector Machines | Geometric Intuition, Why we take values +1 and and -1 for Support vector planes, Mathematical derivation, Loss function (Hinge Loss) based interpretation, Dual form of SVM formulation |

| | | | |
|---|---|---|---|
| **2021-10-23** | Module 4: Machine Learning-II (Supervised Learning Models) | Support Vector Machines | Kernel trick, Polynomial kernel, RBF-Kernel, Domain specific Kernels, Train and run time complexities, nu-SVM: control errors and support vectors, SVM Regression, Cases, Code Sample, Assignment : Behaviour of Linear Models |
| **2021-10-24** | Module 4: Machine Learning-II (Supervised Learning Models) | Support Vector Machines | Revision Support Vector Machines |
| **2021-10-25** | Module 4: Machine Learning-II (Supervised Learning Models) | Decision Trees | Geometric Intuition of decision tree: Axis parallel hyperplanes, Sample Decision tree, Building a decision Tree:Entropy, Building a decision Tree:Information Gain, Building a decision Tree: Gini Impurity, Building a decision Tree: Constructing a DT, Building a decision Tree: Splitting numerical features, Feature standardization, Building a decision Tree:Categorical features with many possible values, Overfitting and Underfitting, Train and Run time complexity, Regression using Decision Trees, Cases, Code Samples, Assignment : Apply Decision Trees on Donors Choose Dataset |
| **2021-10-26** | Module 4: Machine Learning-II (Supervised Learning Models) | Decision Trees | Revision Decision Trees |

| | | | |
|---|---|---|---|
| **2021-10-27** | Module 4: Machine Learning-II (Supervised Learning Models) | Ensemble Models | What are ensembles?, Bootstrapped Aggregation (Bagging) Intuition, Random Forest and their construction, Bias-Variance tradeoff, Bagging :Train and Run-time Complexity., Bagging:Code Sample, Extremely randomized trees, Assignment : Application of Bootstrap samples in Random Forest |
| **2021-10-28** | Module 4: Machine Learning-II (Supervised Learning Models) | Ensemble Models | Random Tree :Cases, Boosting Intuition, Residuals, Loss functions and gradients, Gradient Boosting, Regularization by Shrinkage, Train and Run time complexity, XGBoost: Boosting + Randomization, AdaBoost: geometric intuition, Stacking models |
| **2021-10-29** | Module 4: Machine Learning-II (Supervised Learning Models) | Ensemble Models | Cascading classifiers, Kaggle competitions vs Real world, Assignment : Apply GBDT/XGBOOST/LIGHT-GBM on Donors Choose Dataset |
| **2021-10-30** | Module 4: Machine Learning-II (Supervised Learning Models) | Ensemble Models | Revision Ensemble Models |
| **2021-10-31** | Module 5: Feature Engineering, Productionization and Deployment of ML Models | Featurization And Feature Importance | Introduction, Moving window for Time Series Data, Fourier decomposition, Deep learning features: LSTM, Image histogram, Keypoints: SIFT., Deep learning features: CNN, Relational data, Graph data, Indicator variables, Feature binning, Interaction variables, Mathematical transforms, Model specific featurizations, Feature orthogonality |

| | | | |
|---|---|---|---|
| 2021-11-01 | Module 5: Feature Engineering, Productionization and Deployment of ML Models | Featurization And Feature Importance | Domain specific featurizations, Feature slicing, Kaggle Winners solutions,Revision Featurization And Feature Importance |
| 2021-11-02 | Module 5: Feature Engineering, Productionization and Deployment of ML Models | Miscellaneous Topics | Calibration of Models:Need for calibration, Calibration Plots., Platt's Calibration/ Scaling., Isotonic Regression, Code Samples, Modeling in the presence of outliers: RANSAC, Productionizing models, Retraining models periodically., A/B testing. |
| 2021-11-03 | Module 5: Feature Engineering, Productionization and Deployment of ML Models | Miscellaneous Topics | Data Science Life cycle, Productionization and deployment of Machine Learning Models, Productionization and deployment + Spark |
| 2021-11-04 | Module 5: Feature Engineering, Productionization and Deployment of ML Models | Miscellaneous Topics | Hands on Live Session: Deploy an ML model using APIs on AWS |
| 2021-11-05 | Module 5: Feature Engineering, Productionization and Deployment of ML Models | Miscellaneous Topics | Building web apps for ML/AI using StreamLit |
| 2021-11-06 | Module 5: Feature Engineering, Productionization and Deployment of ML Models | Miscellaneous Topics | Building web apps for ML/AI using StreamLit-ii, VC dimension,Revision Miscellaneous Topics |

| | | | |
|---|---|---|---|
| **2021-11-07** | Module 6: Machine Learning Real World Case studies | Quora Question Pair Similarity | Business/Real world problem : Problem definition , Business objectives and constraints., Mapping to an ML problem : Data overview , Mapping to an ML problem : ML problem and performance metric., Mapping to an ML problem : Train-test split, EDA: Basic Statistics., EDA: Basic Feature Extraction, EDA: Text Preprocessing, EDA: Advanced Feature Extraction, EDA: Feature analysis., EDA: Data Visualization: T-SNE., EDA: TF-IDF weighted Word2Vec featurization., ML Models :Loading Data, ML Models: Random Model, ML Models : Logistic Regression and Linear SVM, ML Models : XGBoost,Revision Quora Question Pair Similarity |
| **2021-11-08** | Module 6: Machine Learning Real World Case studies | Personalized Cancer Diagnosis | Business/Real world problem : Overview, Business objectives and constraints., ML problem formulation :Data, ML problem formulation: Mapping real world to ML problem., ML problem formulation :Train, CV and Test data construction, Exploratory Data Analysis:Reading data & preprocessing, Exploratory Data Analysis:Distribution of Class-labels, Exploratory Data Analysis: "Random" Model, Univariate Analysis:Gene feature |

| | | | |
|---|---|---|---|
| **2021-11-09** | Module 6: Machine Learning Real World Case studies | Personalized Cancer Diagnosis | Univariate Analysis:Variation Feature, Univariate Analysis:Text feature, Machine Learning Models:Data preparation, Baseline Model: Naive Bayes, K-Nearest Neighbors Classification, Logistic Regression with class balancing, Logistic Regression without class balancing, Linear-SVM. |
| **2021-11-10** | Module 6: Machine Learning Real World Case studies | Personalized Cancer Diagnosis | Random-Forest with one-hot encoded features, Random-Forest with response-coded features, Stacking Classifier, Majority Voting classifier,Revision Personalized Cancer Diagnosis |
| **2021-11-11** | Module 6: Machine Learning Real World Case studies | Facebook Friend Recommendation Using Graph Mining | Problem definition. , Overview of Graphs: node/ vertex, edge/link, directed-edge, path. , Data format & Limitations. , Mapping to a supervised classification problem. , Business constraints & Metrics. , EDA:Basic Stats |
| **2021-11-12** | Module 6: Machine Learning Real World Case studies | Facebook Friend Recommendation Using Graph Mining | EDA:Follower and following stats., EDA:Binary Classification Task, EDA:Train and test split., Feature engineering on Graphs:Jaccard & Cosine Similarities, PageRank |
| **2021-11-13** | Module 6: Machine Learning Real World Case studies | Facebook Friend Recommendation Using Graph Mining | Shortest Path, Connected-components, Adar Index, Kartz Centrality, HITS Score, SVD, Weight features, Modeling, Assignment : Facebook Friend Recommendation |
| **2021-11-14** | Module 6: Machine Learning Real World Case studies | Facebook Friend Recommendation Using Graph Mining | Assignment: SQL |

| Date | Module | Topic | Details |
|---|---|---|---|
| **2021-11-15** | Module 6: Machine Learning Real World Case studies | Facebook Friend Recommendation Using Graph Mining | Revision Facebook Friend Recommendation Using Graph Mining |
| **2021-11-16** | Module 6: Machine Learning Real World Case studies | Taxi Demand Prediction in New York City | Business/Real world problem Overview, Objectives and Constraints, Mapping to ML problem :Data, Mapping to ML problem :dask dataframes, Mapping to ML problem :Fields/Features., Mapping to ML problem :Time series forecasting/ Regression, Mapping to ML problem :Performance metrics, Data Cleaning :Latitude and Longitude data, Data Cleaning :Trip Duration., Data Cleaning :Speed., Data Cleaning :Distance., Data Cleaning :Fare, Data Cleaning :Remove all outliers/ erroneous points, Data Preparation:Clustering/ Segmentation |
| **2021-11-17** | Module 6: Machine Learning Real World Case studies | Taxi Demand Prediction in New York City | Data Preparation:Time binning, Data Preparation:Smoothing time-series data., Data Preparation:Smoothing time-series data cont.., Data Preparation: Time series and Fourier transforms., Ratios and previous-time-bin values, Simple moving average, Weighted Moving average., Exponential weighted moving average, Results., Regression models :Train-Test split & Features, Linear regression., Random Forest regression, Xgboost Regression, Model comparison,Revision Taxi Demand Prediction in New York City |

| | | | |
|---|---|---|---|
| **2021-11-18** | Module 6: Machine Learning Real World Case studies | Stack Overflow Tag Predictor | Business/Real world problem, Business objectives and constraints, Mapping to an ML problem: Data overview, Mapping to an ML problem:ML problem formulation., Mapping to an ML problem:Performance metrics., Hamming loss, EDA:Data Loading, EDA:Analysis of tags, EDA:Data Preprocessing, Data Modeling : Multi label Classification |
| **2021-11-19** | Module 6: Machine Learning Real World Case studies | Stack Overflow Tag Predictor | Data preparation., Train-Test Split, Featurization, Logistic regression: One VS Rest, Sampling data and tags+Weighted models., Logistic regression revisited, Why not use advanced techniques,Revision Stack Overflow Tag Predictor |
| **2021-11-20** | Module 6: Machine Learning Real World Case studies | Microsoft Malware Detection | Problem Definition, Objectives and Constraints, Data Overview, ML Problem, Train and Test Splitting, Exploratory Data Analysis:Class Distribution, Exploratory Data Analysis:Feature Extraction from Byte Files, Exploratory Data Analysis:Multivariate analysis of features from byte files, Train-Test class Distribution, ML models – using byte files only :Random Model, K-NN, Logistic regression, Random Forest and XGBoost, Feature Extraction and Multi Threading, File Size Feature, Univariate Analysis, T-SNE Analysis, ML Models on ASM File features, Models on all features: t-SNE, Models on all features: RandomForest and XGBoost, Assignment : Microsoft Malware Detection |

| | | | |
|---|---|---|---|
| **2021-11-21** | Module 6: Machine Learning Real World Case studies | Microsoft Malware Detection | Revision Microsoft Malware Detection |
| **2021-11-22** | Module 7: Data Mining (Unsupervised Learning) and Recommender systems+Real World Case studies | Clustering | What is Clustering?, Unsupervised learning, Applications, Metrics for Clustering, K-Means: Geometric intuition, Centroids, K-Means: Mathematical formulation: Objective function, K-Means Algorithm., How to initialize: K-Means++ |
| **2021-11-23** | Module 7: Data Mining (Unsupervised Learning) and Recommender systems+Real World Case studies | Clustering | Failure cases/Limitations, K-Medoids, Determining the right K, Code Samples, Time and space complexity, Assignment :Clustering on Graph Dataset |
| **2021-11-24** | Module 7: Data Mining (Unsupervised Learning) and Recommender systems+Real World Case studies | Clustering | Revision Clustering |
| **2021-11-25** | Module 7: Data Mining (Unsupervised Learning) and Recommender systems+Real World Case studies | Hierarchical Clustering | Agglomerative & Divisive, Dendrograms, Agglomerative Clustering, Proximity methods: Advantages and Limitations., Time and Space Complexity, Limitations of Hierarchical Clustering, Code sample,Revision Hierarchical Clustering |

| | | | |
|---|---|---|---|
| 2021-11-26 | Module 7: Data Mining (Unsupervised Learning) and Recommender systems+Real World Case studies | DBSCAN Technique | Density based clustering, MinPts and Eps: Density, Core, Border and Noise points, Density edge and Density connected points., DBSCAN Algorithm, Hyper Parameters: MinPts and Eps, Advantages and Limitations of DBSCAN, Time and Space Complexity, Code samples., Question and Answers,Revision DBSCAN Technique |
| 2021-11-27 | Module 7: Data Mining (Unsupervised Learning) and Recommender systems+Real World Case studies | Recommender Systems and Matrix Factorization | Problem formulation: IMDB Movie reviews, Content based vs Collaborative Filtering, Similarity based Algorithms, Matrix Factorization: PCA, SVD, Matrix Factorization: NMF, Matrix Factorization for Collaborative filtering, Matrix Factorization for feature engineering, Clustering as MF, Hyperparameter tuning, Matrix Factorization for recommender systems: Netflix Prize Solution, Cold Start problem, Word vectors as MF |
| 2021-11-28 | Module 7: Data Mining (Unsupervised Learning) and Recommender systems+Real World Case studies | Recommender Systems and Matrix Factorization | Eigen-Faces, Code example., Assignment : Recommendation Systems and Truncated SVD: Implement SGD algorithm to predict the ratings |
| 2021-11-29 | Module 7: Data Mining (Unsupervised Learning) and Recommender systems+Real World Case studies | Recommender Systems and Matrix Factorization | Revision Recommender Systems and Matrix Factorization |

| 2021-11-30 | Module 7: Data Mining (Unsupervised Learning) and Recommender systems+Real World Case studies | Amazon Fashion Discovery Engine | Problem Statement: Recommend similar apparel products in e-commerce using product descriptions and Images, Plan of action, Amazon product advertising API, Data folders and paths, Overview of the data and Terminology, Data cleaning and understanding:Missing data in various features, Understand duplicate rows, Remove duplicates : Part 1 , Remove duplicates: Part 2, Text Pre-Processing: Tokenization and Stop-word removal, Stemming, Text based product similarity :Converting text to an n-D vector: bag of words, Code for bag of words based product similarity, TF-IDF: featurizing text based on word-importance, Code for TF-IDF based product similarity |

| | | | |
|---|---|---|---|
| **2021-12-01** | Module 7: Data Mining (Unsupervised Learning) and Recommender systems+Real World Case studies | Amazon Fashion Discovery Engine | Code for IDF based product similarity, Text Semantics based product similarity: Word2Vec(featurizing text based on semantic similarity), Code for Average Word2Vec product similarity, TF-IDF weighted Word2Vec, Code for IDF weighted Word2Vec product similarity, Weighted similarity using brand and color, Code for weighted similarity, Building a real world solution, Deep learning based visual product similarity:ConvNets: How to featurize an image: edges, shapes, parts, Using Keras + Tensorflow to extract features, Visual similarity based product similarity, Measuring goodness of our solution :A/B testing, Exercise :Build a weighted Nearest neighbor model using Visual, Text, Brand and Color,Revision Amazon Fashion Discovery Engine |
| **2021-12-02** | Module 7: Data Mining (Unsupervised Learning) and Recommender systems+Real World Case studies | Netflix Movie Recommendation system | Business/Real World Problem:Problem Definition, Objectives and Constraints, Mapping to ML problem : Data Overview, Mapping to ML problem : ML problem formulation, Exploratory Data Analysis: Data preprocessing, Exploratory Data Analysis: Temporal Train-Test split, Exploratory Data Analysis: Preliminary Data Analysis, Exploratory Data Analysis: Sparse matrix representation, Exploratory Data Analysis:Average ratings for various slices , Exploratory Data Analysis:Cold start problem, Computing Similarity matrices:User-User similarity matrix |

| | | | |
|---|---|---|---|
| **2021-12-03** | Module 7: Data Mining (Unsupervised Learning) and Recommender systems+Real World Case studies | Netflix Movie Recommendation system | Computing Similarity matrices:Movie-Movie similarity , Computing Similarity matrices:Does movie-movie similarity work?, ML Models:Surprise library , Overview of the modelling strategy. , Data Sampling. , Google drive with intermediate files , Featurizations for regression. , Data transformation for Surprise. , Xgboost with 13 features , Surprise Baseline model. , Xgboost + 13 features +Surprise baseline model , Surprise KNN predictors , Matrix Factorization models using Surprise , SVD ++ with implicit feedback |
| **2021-12-04** | Module 7: Data Mining (Unsupervised Learning) and Recommender systems+Real World Case studies | Netflix Movie Recommendation system | Final models with all features and predictors., High Level + End-End Design of a Music Recommendation system - I, High Level + End-End Design of a Music Recommendation system - II |
| **2021-12-05** | Module 7: Data Mining (Unsupervised Learning) and Recommender systems+Real World Case studies | Netflix Movie Recommendation system | Building a simple Youtube recommendation using basic Math,Revision Netflix Movie Recommendation system |
| **2021-12-06** | Module 8: Neural Networks, Computer Vision and Deep Learning | Neural Networks | History of Neural networks and Deep Learning., How Biological Neurons work?, Growth of biological neural networks, Diagrammatic representation: Logistic Regression and Perceptron, Multi-Layered Perceptron (MLP). |

| Date | Module | Topic | Subtopics |
|---|---|---|---|
| **2021-12-07** | Module 8: Neural Networks, Computer Vision and Deep Learning | Neural Networks | Notation, Training a single-neuron model., Training an MLP: Chain Rule, Training an MLP:Memoization |
| **2021-12-08** | Module 8: Neural Networks, Computer Vision and Deep Learning | Neural Networks | Backpropagation., Activation functions, Vanishing Gradient problem., Bias-Variance tradeoff., Decision surfaces: Playground |
| **2021-12-09** | Module 8: Neural Networks, Computer Vision and Deep Learning | Neural Networks | Revision Neural Networks |
| **2021-12-10** | Module 8: Neural Networks, Computer Vision and Deep Learning | Deep Multi Layer Perceptrons | Deep Multi-layer perceptrons:1980s to 2010s, Dropout layers & Regularization., Rectified Linear Units (ReLU)., Weight initialization., Batch Normalization. |
| **2021-12-11** | Module 8: Neural Networks, Computer Vision and Deep Learning | Deep Multi Layer Perceptrons | Optimizers:Hill-descent analogy in 2D, Optimizers:Hill descent in 3D and contours., SGD Recap, Batch SGD with momentum., Nesterov Accelerated Gradient (NAG), Optimizers:AdaGrad, Optimizers : Adadelta andRMSProp, Adam, Which algorithm to choose when?, Gradient Checking and clipping, Softmax and Cross-entropy for multi-class classification., How to train a Deep MLP?, Auto Encoders. |
| **2021-12-12** | Module 8: Neural Networks, Computer Vision and Deep Learning | Deep Multi Layer Perceptrons | Word2Vec :CBOW, Word2Vec: Skip-gram, Word2Vec :Algorithmic Optimizations., Assignment : Backpropagation and Gradient Checking |

| | | | |
|---|---|---|---|
| **2021-12-13** | Module 8: Neural Networks, Computer Vision and Deep Learning | Deep Multi Layer Perceptrons | Revision Deep Multi Layer Perceptrons |
| **2021-12-14** | Module 8: Neural Networks, Computer Vision and Deep Learning | Tensorflow And Keras | Tensorflow and Keras overview, GPU vs CPU for Deep Learning., Google Colaboratory., Install TensorFlow, Online documentation and tutorials, Softmax Classifier on MNIST dataset. |
| **2021-12-15** | Module 8: Neural Networks, Computer Vision and Deep Learning | Tensorflow And Keras | MLP: Initialization, Model 1: Sigmoid activation., Model 2: ReLU activation., Model 3: Batch Normalization., Model 4 : Dropout., MNIST classification in Keras., Hyperparameter tuning in Keras., Assignment : Working with Callbacks |
| **2021-12-16** | Module 8: Neural Networks, Computer Vision and Deep Learning | Tensorflow And Keras | Exercise: Try different MLP architectures on MNIST dataset.,Revision Tensorflow And Keras |
| **2021-12-17** | Module 8: Neural Networks, Computer Vision and Deep Learning | Convolutional Neural Nets | Biological inspiration: Visual Cortex, Convolution:Edge Detection on images., Convolution:Padding and strides, Convolution over RGB images., Convolutional layer. |
| **2021-12-18** | Module 8: Neural Networks, Computer Vision and Deep Learning | Convolutional Neural Nets | Max-pooling., CNN Training: Optimization, Example CNN: LeNet [1998], ImageNet dataset., Data Augmentation., Convolution Layers in Keras, AlexNet, VGGNet, Residual Network., Inception Network., What is Transfer learning., Code example: Cats vs Dogs., Code Example: MNIST dataset., Assignment : Transfer Learning - (Given an rvl-cdip dataset, classify the given document using transfer learning) |

| Date | Module | Topic | Content |
|---|---|---|---|
| 2021-12-19 | Module 8: Neural Networks, Computer Vision and Deep Learning | Convolutional Neural Nets | Assignment : Document Classification with CNN |
| 2021-12-20 | Module 8: Neural Networks, Computer Vision and Deep Learning | Convolutional Neural Nets | Revision Convolutional Neural Nets |
| 2021-12-21 | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | Why RNNs? , Recurrent Neural Network., Training RNNs: Backprop., Types of RNNs., Need for LSTM/GRU. |
| 2021-12-22 | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | LSTM., GRUs., Deep RNN., Bidirectional RNN., Code example : IMDB Sentiment classification |
| 2021-12-23 | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | Assignment : LSTM on Donors Choose - (LSTM with Text and categorical data) |
| 2021-12-24 | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | Assignment : CNN on CIFR - (Classifying CIFAR-10 dataset images with DenseNet and work with optimization) |
| 2021-12-25 | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | Exercise: Amazon Fine Food reviews LSTM model., Deep Learning: Generative Adversarial Networks (GANs):Live session on Generative Adversarial Networks (GAN), Encoder-Decoder Models:LIVE: Encoder-Decoder Models |
| 2021-12-26 | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | Attention Models in Deep Learning:Attention Models in Deep Learning, Assignment : NLP Attention Mechanism |
| 2021-12-27 | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | Deep Learning: Transformers and BERT:Transformers and BERT |

| | | | |
|---|---|---|---|
| **2021-12-28** | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | Assignment : NLP with Transfer Learning - (Classification of reviews using BERT embeddings) |
| **2021-12-29** | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | Deep Learning: Image Segmentation:Live session on Image Segmentation |
| **2021-12-30** | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | Assignment : Computer Vision: Segmentation - (Self Driving Cars: Detect the Objects on the road using Semantic Segmentation) |
| **2021-12-31** | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | Deep Learning: Object Detection:Object Detection |
| **2022-01-01** | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | Object Detection YOLO V3, Generative Adversarial Networks (GANs) |
| **2022-01-02** | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | Encoder Decoder Models, Attention Models in Deep Learning, Transformers and BERT |
| **2022-01-03** | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | Image Segmentation |
| **2022-01-04** | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | Object Detection |
| **2022-01-05** | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | Object Detection YOLO V3 |

| Date | Module | Topic | Details |
| --- | --- | --- | --- |
| 2022-01-06 | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | GPT-1, 2 and GPT-3 Models |
| 2022-01-07 | Module 8: Neural Networks, Computer Vision and Deep Learning | Long Short-Term Memory(LSTMS) | Design and Build a smart store,Revision Long Short-Term Memory(LSTMS) |
| 2022-01-08 | Module 9: Deep Learning Real World Case Studies | Human Activity Recognition | Human Activity Recognition Problem definition, Dataset understanding, Data cleaning & preprocessing, EDA:Univariate analysis., EDA:Data visualization using t-SNE, Classical ML models., Deep-learning Model., Exercise: Build deeper LSTM models and hyper-param tune them,Revision Human Activity Recognition |
| 2022-01-09 | Module 9: Deep Learning Real World Case Studies | Self Driving Car | Problem Definition, Datasets., Data understanding & Analysis :Files and folders., Dash-cam images and steering angles., Split the dataset: Train vs Test, EDA: Steering angles, Mean Baseline model: simple, Deep-learning model:Deep Learning for regression: CNN, CNN+RNN, Batch load the dataset., NVIDIA's end to end CNN model., Train the model., Test and visualize the output., Extensions.,Revision Self Driving Car |
| 2022-01-10 | Module 9: Deep Learning Real World Case Studies | Music Generation Using Deep Learning | Real-world problem, Music representation, Char-RNN with abc-notation :Char-RNN model, Char-RNN with abc-notation :Data preparation. |

| Date | Module | Topic | Details |
|---|---|---|---|
| 2022-01-11 | Module 9: Deep Learning Real World Case Studies | Music Generation Using Deep Learning | Char-RNN with abc-notation:Many to Many RNN ,TimeDistributed-Dense layer, Char-RNN with abc-notation : State full RNN, Char-RNN with abc-notation :Model architecture,Model training., Char-RNN with abc-notation :Music generation., Char-RNN with abc-notation :Generate tabla music, MIDI music generation., Case Study 13: Semantic Search Engine for Q&A [Design + Code]:Semantic Search for Q&A [Design + Code] --- Part 1 |
| 2022-01-12 | Module 9: Deep Learning Real World Case Studies | Music Generation Using Deep Learning | Case Study 13: Semantic Search Engine for Q&A [Design + Code]:Semantic Search for Q&A [Design + Code] --- Part 2 , Case Study 13: Semantic Search Engine for Q&A [Design + Code]:Semantic Search for Q&A [Design + Code] --- Part 3, Case Study 13: Semantic Search Engine for Q&A [Design + Code]:Semantic Search for Q&A [Design + Code] --- Part 4 , Survey blog, Assignment : Spoken Digit Recognition - (Working with Audio Dataset: Detect the sounds using spectrograms and Deep Learning) |
| 2022-01-13 | Module 9: Deep Learning Real World Case Studies | Music Generation Using Deep Learning | Revision Music Generation Using Deep Learning |

**Applied AI Course Wishes You All The Best**

Please mail us to team@appliedaicourse.com if you have any queries