

Full Stack Data Science Feb'21 Batch

Instructors:

Sunny Bhaveen Chandra:

Sr. Data Scientist and lecturer at iNeuron.ai with working experience in computer vision, natural language processing and embedded systems. Hands-on experience leveraging machine learning, deep learning, transfer learning models to solve challenging business problems. Also, he has a vast interest in Robotics.

krish naik:

Having 10+ years of experience in Data Science and Analytics with product architecture design and delivery. Worked in various product and service based Company. Having an experience of 5+ years in educating people and helping them to make a career transition.

Sudhanshu Kumar:

Having 8+ years of experience in Big data, Data Science and Analytics with product architecture design and delivery. Worked in various product and service based Company. Having an experience of 5+ years in educating people and helping them to make a career transition.

Curriculum:

Course introduction

- a. course overview and dashboard description
- b. Introduction of data science and its application in day to day life
- c. Programming language overview
- d. Installation (tools: sublime, vscode, pycharm, anaconda, atom, jupyter notebook, kite)
- e. Virtual environment

- f. Why python

Python basic

- a. Introduction of python and comparison with other programming language
- b. Installation of anaconda distribution and other python ide
- c. Python objects, number & Booleans, strings.
- d. Container objects, mutability of objects
- e. Operators - arithmetic, bitwise, comparison and assignment operators, operator's precedence and associativity
- f. Conditions (if else, if-elif-else), loops (while, for)
- g. Break and continue statement and range function

String objects

- a. basic data structure in python
- b. String object basics
- c. String inbuilt methods
- d. Splitting and joining strings
- e. String format functions

List object basics

- a. List methods
- b. List as stack and queues
- c. List comprehensions

Tuples, set, dictionaries & its function

- Dictionary object methods

- Dictionary comprehensions
- Dictionary view objects
- Functions basics, parameter passing, iterators.
- Generator functions
- Lambda functions
- Map, reduce, filter functions.

Memory management

- Multithreading
- Multiprocessing

Oops concepts

- oops basic concepts.
- Creating classes
- Pillars of oops
- Inheritance
- Polymorphism
- Encapsulation
- Abstraction
- Decorator
- Class methods and static methods
- Special (magic/dunder) methods
- Property decorators - getters, setters, and deletes

Files

- Working with files
- Reading and writing files
- Buffered read and write
- Other file methods.
- Logging, debugger
- Modules and import statements

Exception handling difference between exceptions and error

- Exceptions handling with try-except
- Custom exception handling
- List of general use exception
- Best practice exception handling

Gui framework

- What is desktop and standalone application
- Use of desktop app
- Examples of desktop app
- Tinker
- Kivy

Database

- SQLite
- MySQL
- Mongo dB
- NoSQL - Cassandra

Web API

- What is web API
- Difference b/w API and web API
- Rest and soap architecture
- Restful services

Flask

- Flask introduction
- Flask application
- Open link flask
- App routing flask
- Url building flask
- Http methods flask
- Templates flask
- Flask project: food app
- Postman
- Swagger

Django

- Django introduction
- Django project: weather app
- Django project: memes generator
- Django project: blog app
- Django project in cloud

Stream lit

- Stream lit introduction
- Stream lit project structure
- Stream lit project in cloud

Pandas basic

- Python pandas - series
- Python pandas – data frame
- Python pandas – panel
- Python pandas - basic functionality
- Reading data from different file system

Pandas advance

- Python pandas – re indexing python
- Pandas – iteration
- Python pandas – sorting.
- Working with text data options & customization
- Indexing & selecting
- Data statistical functions
- Python pandas - window functions
- Python pandas - date functionality
- Python pandas –time delta
- Python pandas - categorical data
- Python pandas – visualization

- Python pandas - iotools

Dask

- Dask Array
- Dask Bag
- Dask DataFrame
- Dask Delayed
- Dask Futures
- Dask API
- Dask SCHEDULING
- Dask Understanding Performance
- Dask Visualize task graphs
- Dask Diagnostics (local)
- Dask Diagnostics (distributed)
- Dask Debugging
- Dask Ordering

Python numpy

- Numpy - ND array object.
- Numpy - data types.
- Numpy - array attributes.
- Numpy - array creation routines.
- Numpy - array from existing.
- Data array from numerical ranges.
- Numpy - indexing & slicing.

- Numpy – advanced indexing.
- Numpy – broadcasting.
- Numpy - iterating over array.
- Numpy - array manipulation.
- Numpy - binary operators.
- Numpy - string functions.
- Numpy - mathematical functions.
- Numpy - arithmetic operations.
- Numpy - statistical functions.
- Sort, search & counting functions.
- Numpy - byte swapping.
- Numpy - copies & views.;
- Numpy - matrix library.
- Numpy - linear algebra

Visualization

- Matplotlib
- Seaborn
- Cufflinks
- Plotly
- Bokeh

Statistics basic

- Introduction to basic statistics terms
- Types of statistics

- Types of data
- Levels of measurement
- Measures of central tendency
- Measures of dispersion
- Random variables
- Set
- Skewness
- Covariance and correlation

Probability distribution function

- Probability density/distribution function
- Types of the probability distribution
- Binomial distribution
- Poisson distribution
- Normal distribution (Gaussian distribution)
- Probability density function and mass function
- Cumulative density function
- Examples of normal distribution
- Bernoulli distribution
- Uniform distribution
- Z stats
- Central limit theorem
- Estimation

Statistics advance

- a Hypothesis
- Hypothesis testing's mechanism
- P-value
- T-stats
- Student t distribution
- T-stats vs. Z-stats: overview
- When to use a t-tests vs. Z-tests
- Type 1 & type 2 error
- Bayes statistics (Bayes theorem)
- Confidence interval(ci)
- Confidence intervals and the margin of error
- Interpreting confidence levels and confidence intervals
- Chi-square test
- Chi-square distribution using python
- Chi-square for goodness of fit test
- When to use which statistical distribution?
- Analysis of variance (anova)
- Assumptions to use anova
- Anova three type
- Partitioning of variance in the anova
- Calculating using python
- F-distribution
- F-test (variance ratio test)
- Determining the values of f

- F distribution using python

Linear algebra

- linear algebra
- Vector
- Scaler
- Matrix
- Matrix operations and manipulations
- Dot product of two vectors
- Transpose of a matrix
- Linear independence of vectors
- Rank of a matrix
- Identity matrix or operator
- Determinant of a matrix
- Inverse of a matrix
- Norm of a vector
- Eigenvalues and eigenvectors
- Calculus

Solving stats problem with python

Stats problem implementation with spicy

Introduction to machine learning

- Ai vs ml vs dl vs ds

- Supervised, unsupervised, semi-supervised, reinforcement learning
- Train, test, validation split
- Performance
- Overfitting, under fitting
- Bias vs variance

Feature engineering

- Handling missing data
- Handling imbalanced data
- Up-sampling
- Down-sampling
- Smote
- Data interpolation
- Handling outliers
- Filter method
- Wrapper method
- Embedded methods
- Feature scaling
- Standardization
- Mean normalization
- Min-max scaling
- Unit vector
- Feature extraction
- Pca (principle component analysis)
- Data encoding

- Nominal encoding
- One hot encoding
- One hot encoding with multiple categories
- Mean encoding
- Ordinal encoding
- Label encoding
- Target guided ordinal encoding
- Covariance
- Correlation check
- Pearson correlation coefficient
- Spearman's rank correlation
- Vif

Feature selection

- Feature selection
- Recursive feature elimination
- Backward elimination
- Forward elimination

Exploratory data analysis

- Feature engineering and selection.
- Analyzing bike sharing trends.
- Analyzing movie reviews sentiment.
- Customer segmentation and effective cross selling.
- Analyzing wine types and quality.

- Analyzing music trends and recommendations.
- Forecasting stock and commodity prices

Regression

- Linear regression
- Gradient descent
- Multiple linear regression
- Polynomial regression
- R square and adjusted r square
- Rmse , mse, mae comparison
- Regularized linear models
- Ridge regression
- Lasso regression
- Elastic net
- Complete end-to-end project with deployment on cloud and ui

Logistics regression

- Logistics regression in-depth intuition
- In-depth mathematical intuition
- In-depth geometrical intuition
- Hyper parameter tuning
- Grid search cv
- Randomize search cv
- Data leakage
- Confusion matrix

- Precision, recall, f1 score, roc, auc
- Best metric selection
- Multiclass classification in lr
- Complete end-to-end project with deployment in multi cloud platform

Decision tree

- Decision tree classifier
- In-depth mathematical intuition
- In-depth geometrical intuition
- Confusion matrix
- Precision, recall, f1 score, roc, auc
- Best metric selection
- Decision tree regressor
- In-depth mathematical intuition
- In-depth geometrical intuition
- Performance metrics
- Complete end-to-end project with deployment in multi cloud platform

Support vector machines

- Linear svm classification
- In-depth mathematical intuition
- In-depth geometrical intuition
- Soft margin classification
- Nonlinear svm classification
- Polynomial kernel

- Gaussian, rbf kernel
- Data leakage
- Confusion matrix
- precision, recall, f1 score, roc, auc
- Best metric selection
- Svm regression
- In-depth mathematical intuition
- In-depth geometrical intuition
- Complete end-to-end project with deployment

Naïve Bayes

- Bayes theorem
- Multinomial naïve Bayes
- Gaussian naïve Bayes
- Various type of Bayes theorem and its intuition
- Confusion matrix
- precision, recall, f1 score, roc, auc
- Best metric selection
- Complete end-to-end project with deployment

Ensemble techniques and its types

- Definition of ensemble techniques
- Bagging technique
- Bootstrap aggregation
- Random forest (bagging technique)

- Random forest regressor
- Random forest classifier
- Complete end-to-end project with deployment

Boosting

- Boosting technique
- Ada boost
- Gradient boost
- Xgboost
- Complete end-to-end project with deployment

Stacking

- Stacking technique
- Complete end-to-end project with deployment

Knn

- Knn classifier
- Knn regressor
- Variants of knn
- Brute force knn
- K-dimension tree
- Ball tree
- Complete end-to-end project with deployment

Dimensionality reduction

- The curse of dimensionality
- Dimensionality reduction technique
- Pca (principle component analysis)
- Mathematics behind pca
- Scree plots
- Eigen-decomposition approach

Clustering

- Clustering and their types
- K-means clustering
- K-means++
- Batch k-means
- Hierarchical clustering
- Dbscan
- Evaluation of clustering
- Homogeneity, completeness and v-measure
- Silhouette coefficient
- Davies-bouldin index
- Contingency matrix
- Pair confusion matrix
- Extrinsic measure
- Intrinsic measure
- Complete end-to-end project with deployment

Anomaly detection

- Anomaly detection types
- Anomaly detection applications
- Isolation forest anomaly detection algorithm
- Density-based anomaly detection (local outlier factor) algorithm
- Support vector machine anomaly detection algorithm
- DbSCAN algorithm for anomaly detection
- Complete end-to-end project with deployment

Time-series

- What is a time series?
- Old techniques
- Arima
- ACF and PACF
- Time-dependent seasonal components.
- Autoregressive (AR),
- Moving average (MA) and mixed ARMA-modeler.
- The random walk model.
- Box-Jenkins methodology.
- Forecasts with ARIMA and VAR models.
- Dynamic models with time-shifted explanatory variables.
- The Koyck transformation.
- Partial adjustment and adaptive expectation models.
- Granger's causality tests.
- Stationarity, unit roots and integration
- Time series model performance

- Various approach to solve time series problem
- Complete end-to-end project with deployment
- Prediction of nifty stock price and deployment

NLP basic

- Tokenization
- Pos tags and chunking
- Stop words
- Stemming and lemmatization
- Named entity recognition (ner)
- Word vectorization (word embedding)
- Tfidf
- Complete end-to-end project with deployment

Machine learning pipeline

- Aws segmaker
- Aure ml studio
- Ml flow
- Kube flow

Model retraining approach

Auto ML

- H2o
- Pycaret

- Auto sklearn
- Auto time series
- Auto viml
- Auto gluon
- Auto viz
- Tpot
- Auto neuro

Neural network a simple perception

- Detail mathematical explanation
- Neural network overview and its use case.
- Various neural network architect overview.
- Use case of neural network in nlp and computer vision.
- Activation function -all name
- Multilayer network.
- Loss functions. - all 10
- The learning mechanism.
- Optimizers. - all 10
- Forward and backward propagation.
- Weight initialization technique
- Vanishing gradient problem
- Exploding gradient problem
- Visualization of nn

Hardware setup – GPU

- Gpu introduction.
- Various type of gpu configuration.
- Gpu provider and its pricing.
- Paper space gpu setup.
- Running model in gpu

Tensor flow installation environment setup for deep learning

- Colab pro setup
- Tensor flow installation 2.0 .
- Tensor flow installation 1.6 with virtual environment.
- Tensor flow 2.0 function.
- Tensor flow 2.0 neural network creation.
- Tensor flow 1.6 functions.
- Tensor flow 1.6 neural network and its functions.
- Keras introduction.
- Keras in-depth with neural network creation.
- Mini project in tensorflow.
- Tensorspace
- Tensorboard integration
- Tensorflow playground
- Netron

Pytorch

- pytorch installation.
- Pytorch functional overview.

- Pytorch neural network creation.

Mxnet

- Mxnet installation
- Mxnet in depth function overview
- Mxnet model creation and training

Keras tuner

- Keras tuner installation and overview
- Finding best parameter from keras tuner
- Keras tuner application across various neural network

Cnn overview

- Cnn definition
- Various cnn based architecture
- Explanation end to end cnn network
- Cnn explainer
- Training cnn
- Deployment in azure cloud
- Performance tuning of cnn network

Advance computer vision – part 1

- Various cnn architecture with research paper and mathematics
- Lenet-5 variants with research paper and practical
- Alexnet variants with research paper and practical

- Googlenet variants with research paper and practical
- Transfer learning
- Vggnet variants with research paper and practical
- Resnet variants with research paper and practical
- Inception net variants with research paper and practical
- Darknet variants with research paper and practical

Advance computer vision – part 2

- Object detection in-depth
- Transfer learning
- Rcn with research paper and practical
- Fast rcnn with research paper and practical
- Faster r cnn with research paper and practical
- Ssd with research paper and practical
- Ssd lite with research paper and practical

Training of custom object detection

- Tfod introduction
- Environment setup with tfod
- Gpu vs tpu vs cpu
- Various gpu comparison

Advance computer vision – part 3

- Yolo v1 with research paper and practical
- Yolo v2 with research paper and practical

- Yolo v3 with research paper and practical
- Yolo v4 with research paper and practical
- Yolo v5 with research paper and practical
- Retina net
- Face net
- Detectron2 with practical and live testing

Object segmentation

- Semantic segmentation
- Panoptic segmentation
- Masked rcnn
- Practical with detectron
- Practical with tfod

Object tracking

- Detail of object tracking
- Kalman filtering
- Sort
- Deep sort
- Object tracking live project with live camera testing

OCR

- Introduction to ocr
- Various framework and api for ocr
- Practical implementation of ocr

Advance NLP with deep-learning

- Overview computational linguistic.
- History of nlp.
- Why nlp
- Use of nlp

Text processing importing text

- Web scrapping.
- Text processing
- Understanding regex.
- Text normalization
- Word count.
- Frequency distribution.
- Text annotation.
- Use of annotator.
- String tokenization
- Annotator creation.
- Sentence processing.
- Lemmatization in text processing
- Pos.
- Named entity recognition
- Dependency parsing in text.
- Sentimental analysis

Spacy

- Spacy overview.
- Spacy function
- Spacy function implementation in text processing.
- Pos tagging, challenges and accuracy.
- Entities and named entry recognition
- Interpolation, language models
- Nltk
- Text blob
- Stanford nlp

RNN

- Recurrent neural networks.
- Long short term memory (lstm)
- Bi lstm.
- Stacked lstm
- Gru implementation.
- Building a story writer using character level rnn.

Word embedding

- Word embedding
- Co-occurrence vectors
- Word2vec
- Doc2vec

Attention based model

- Seq 2 seq.
- Encoders and decoders.
- Attention mechanism.
- Attention neural networks
- Self-attention

Transfer learning in nlp

- Introduction to transformers.
- Bert model.
- Elmo model.
- Gpt1 model
- Gpt2 model.
- Albert model.
- Distilbert model

Deployment of model and performance tuning

- Deep learning model deployment strategies.
- Deep learning project architecture
- Deep learning model deployment phase.
- Deep learning model retraining phase.
- Deep learning model deployment in aws.
- Deep learning model deployment in azure.
- Deep learning model deployment in gcloud.

Big data introduction

- What is big data?
- Big data application
- Big data pipeline

Hadoop

- Hadoop introduction
- Hadoop setup and installation

Spark

- Spark
- Spark overview.
- Spark installation.
- Spark rdd.
- Spark data frame.
- Spark architecture.
- Spark ml lib
- Spark NLP
- Spark linear regression
- Spark logistic regression
- Spark decision tree
- Spark naive bayes
- Spark xg boost.
- Spark time series
- Spark deployment in local server

- Spark job automation with
- Scheduler

Kafka

- Kafka introduction
- Kafka installation
- Spark streaming
- Spark with Kafka

Tableau

- Talking about Business Intelligence
- Tools and Methodologies used in BI
- Why Visualization is getting more popular
- Why Tableau?
- Gartner Magic Quadrant of Market Leaders
- Future business impact of BI
- Tableau Products
- Tableau Architecture
- BI Project Execution
- Tableau Installation in local system
- Introduction to Tableau Prep
- Tableau Prep Builder User Interface
- Data Preparation techniques using Tableau Prep Builder tool
- How to connect Tableau with different data source
- Visual Segments

- Visual Analytics in depth
- Filters, Parameters & Sets
- Tableau Calculations using functions
- Tableau Joins
- Working with multiple data source (Data Blending)
- Building Predictive Models
- Dynamic Dashboards and Stories
- Sharing your Reports
- Tableau Server
- User Security
- Scheduling

Power BI

- Power BI introduction and overview
- Key Benefits of Power BI
- Power BI Architecture
- Power BI Process
- Components of Power BI
- Power BI - Building Blocks
- Power BI vs other BI tools
- Power Installation
- Overview of Power BI Desktop
- Data Sources in Power BI Desktop
- Connecting to a data Sources
- Query Editor in Power BI

- Views in Power BI
- Field Pane
- Visual Pane
- Custom Visual Option
- Filters
- Introduction to using Excel data in Power BI
- Exploring live connections to data with Power BI
- Connecting directly to SQL Azure, HD Spark, SQL Server Analysis Services/ My SQL
- Import Power View and Power Pivot to Power BI
- Power BI Publisher for Excel
- Content packs
- Introducing Power BI Mobile
- Power Query Introduction
- Query Editor Interface
- Clean and Transform your data with Query Editor
- Data Type
- Column Transformations vs Adding Columns
- Text Transformations
- Cleaning irregularly formatted data -Transpose
- Date and Time Calculations
- Advance editor: Use Case
- Query Level Parameters
- Combining Data – Merging and Appending
- Data Modelling

- Calculated Columns
- Measures/New Quick Measures
- Calculated Tables
- Optimizing Data Models
- Row Context vs Set Context
- Cross Filter Direction
- Manage Data Relationship
- Why is DAX important?
- Advanced calculations using Calculate functions
- DAX queries

Reinforcement Learning

Python Project

- Weeding script
- Image resizing
- Jupyter notebook merging, reading etc.
- Sending emails
- Weather app
- Memes generator
- Food log app
- Web scrapping
- Web crawlers for image data sentiment analysis and product review sentiment analysis.
- Integration with web portal.
- Integration with rest api, web portal and mongo db. on azure

- Deployment on web portal on azure.
- Text mining
- Social media data churn
- Mass copy, paste

Chatbot projects

- Chatbot using Microsoft Luis
- Chatbot using google dialog flow
- Chatbot using amazon lex
- Chatbot using rasa nlu
- Deployment of Chabot with web , telegram , WhatsApp, skype

Major projects

- Healthcare analytics prediction of medicines based on Fitbit band.
- Revenue forecasting for startups.
- Prediction of order cancellation at the time of ordering inventories.
- anomaly detection in inventory packaged material.
- Fault detection in wafers based on sensor data.
- Demand forecasting for fmcg product.
- Threat identification in security system.
- Defect detection in vehicle engine.
- Food price forecasting with zomato dataset.
- Fault detection in wafers based on sensor data.
- Cement strength reg.
- Credit card fraud.

- Forest cover classification.
- Fraud detection.
- Income prediction.
- Mushroom classifier.
- phishing classifier
- Thyroid detection.
- Visibility climate

Computer vision project

- Traffic surveillance system.
- Object identification.
- Object tracking.
- Object classification.
- Tensorflow object detection.
- Image to text processing.
- Speech to speech analysis.
- Vision based attendance system

Mini NLP project

- Machine translation.
- Abstractive text summarization.
- Keyword spotting.
- Language modelling.
- Document summarization

