



DATA SCIENCE

www.semicoretech.com

Contact us : 9700779739, 8688018839

1st floor, sri sai Arcade, Near Aditya Trade Center,
Ameerpet, hyderabad, Telangana, 500016

PYTHON

- Introduction to Data Science
- Python Basics
- Python Data Types & Utilities.
- Control Structures & Loops
- Strings
- Exception Handling
- Collection Framework
- List
- Set
- Tuple
- Dictionary and Dictionary comprehension
- Functions
- Inbuilt vs User Defined
- User Defined Function
- Function Argument
- Types of Function Arguments
- Actual Argument
- Global variable vs Local variable
- Packages
- Anonymous Function | LAMBDA
- Map Reduce
- OOPS
- Class & Object
- Methods
- Python Decorator
- Python Production Level
- Polymorphism
- Error/Exception Handling
- File Handling
- Docstrings
- Modularization
- Pickling & Unpickling
- Pandas
- Numpy
- Matplotlib
- Sea born
- Spicy
- Stats models

Mathematics

- Set Theory
- Data Representation & Database Operations

Combinatorics

- Feature Selection
- Permutations and Combinations for Sampling
- Hyper parameter Tuning
- Experiment Design
- Data Partitioning and Cross-Validation



Probability

- Basics
- Theoretical Probability
- Empirical Probability □
- Addition Rule
- Conditional Probability □
- Multiplication Rule
- Total Probability
- Probability Decision Tree
- Bayes Theorem □
- Sensitivity & Specificity in Probability □
- Bernouli Naïve Bayes, Gaussian Naïve Bayes, Multinomial NaïveBayes

Linear Algebra

- Linear Equations
- Matrices (Matrix Algebra: Vector Matrix Vector matrix multiplication)
- Matrix matrix multiplication)
- Determinant
- Eigen Value and Eigen Vector
- Euclidean Distance & Manhattan Distance
- Calculus
- Differentiation
- Max & Min
- Partial Differentiation
- Indices & Logarithms

STATISTICS

INTRODUCTION

- Types of Data
- Measures of Central Tendency
- Descriptive statistic Measures of symmetry
- Measurement of Spread
- Measurement of Spread
- Levels of Data Measurement
- Variable Frequency Distribution Table
- Types of Variables
- Frequency Distribution Table
- Correlation, Regression & Collinearity
- Others
- Bias and Variance in ML
- Entropy in ML
- Information Gain
- Surprise in ML
- Loss Function & Cost Function
- Inferential Statistics

SQL

Introduction

- Keys
- Constraints
- CRUD Operations
- SQL Languages
- SQL Commands
- Operators
- Wild Cards
- Aggregate Functions
- SQL Joins



EDA & ML

- EDA
- Data Visualisation
- ML. Introduction
- Important Element of Machine Learning.
- Multiclass Classification
- Data Processing
- Regression
- Evaluation Metrics for Regression
- Classification
- Clustering Introduction
- K-Means Clustering
- Hierarchical Clustering
- Agglomerative clustering
- DBSCAN Clustering
- Association Rules.
- Recommendation Engines Time Series & Forecasting
- Model Selection & Evaluation
- Over Fitting & Under Fitting
- Others
- ML Pipeline
- ML Model Deployment in Flask

Power BI

Introduction



- Basic Report Design
- Visual Sync, Grouping
- Hierarchies, Filters
- Power Query
- Power Query transformation
- Table and Column Transformations
- Text and time transformations
- Power query functions
- Merge and append transformations
- DAX Functions
- DAX Architecture, Entity Sets
- DAX Data types. Syntax Rules
- DAX measures and calculations
- Creating measures
- Creating measures

DEEP LEARNING

- Deep learning at Glance
 - Introduction to Neural Network
 - Biological and Artificial Neuron
 - Introduction to perceptron
 - Perceptron and its learning rule and drawbacks
 - Multilayer Perceptron, loss function
 - Neural Network Activation function
- Training MLP: Backpropagation
- Gradient Descent Backpropagation Vanishing and Exploding Gradient
- Problem
- Cost Function
- Introduce to Py-Torch
- Regularization
- Hyper parameters and tuning of the same
- Optimizers
- Tensor Flow Framework
 - Introduction to tensor Flow
 - tensor Flow Basic Syntax
 - tensor Flow Graphs
 - Variables and Placeholders
 - tensor Flow Playground

- **ANN (Artificial Neural Network)**
 - ANN Architecture
 - Forward & Backward Propagation, Epoch
 - Forward & Backward Propagation, Epoch
 - Vanishing Gradient Descend
 - Fine-tuning neural network hyper parameter
 - Number of hidden layers, Number of neurons per hidden layer
 - Activation function
 - INSTALLATION OF YOLO V8, KERAS, THEANO
- **Py-Torch Library**
- **RNN (Recurrent Neural Network)**
 - Introduction to RNN
 - Back Propagation through time
 - Input and output sequences
 - LSTM (Long Short-Term Memory)
 - RNN vs ANN
 - Different types of RNN: LSTM, GRU
 - Biirectional RNN
 - Sequential-to-sequential architecture (Encoder Decoder)
 - BERT Transformers
 - Text generation and classification using Deep Learning
 - Generative-AI (Chat-GPT) Basics of Image Processing
 - Convolutional Neural Networks (CNN)
 - ImageNet Dataset
 - Project: Image Classification
 - Different types of CNN architectures
 - Recurrent Neural Network (RNN)
 - Using pre-trained model: Transfer Learning

NLP

- **Introduction to Natural Language Processing (NLP)**
- **Document Vectorization**
 - Bag of Words
 - TF-IDF Vectorise
 - on-gram: Unigram, Bigram
 - Word vectorizer basics, One Hot Encoding
 - Count Vectorizer
 - Word cloud and genism
 - Word2Vec and Glove
 - Text classification using Word2Vec and Glove
 - Parts of Speech Tagging (PoS Tagging or POST)
 - Topic Modelling using LDA
 - Sentiment Analysis
- **Twitter Sentiment Analysis Using Textblob**
 - TextBlob
 - Installing textblob library
 - Simple TextBlob Sentiment Analysis Example
 - Using NLTK's Twitter Corpus

- **Spacy Library**

- Introduction, what is a Token, Tokenization
- Stop words in spacy library
- Stemming
- Lemmatization
- Lemmatization through NLTK
- Lemmatization using spacy
- Word Frequency Analysis
- Counter
- Part of Speech, Part of Speech Tagging
- Pose by using spacy and nitk
- Dependency Parsing
- Named Entity Recognition (NER)
- NER with NLTK
- NER with spacy

Computer Vision



- Human vision vs Computer vision
- Image Processing with Open CV Video Processing with Open CV
- Reinforcement Learning.
- Generative AI
 - Introduction to Generative AI
 - Generative AI
 - Chat Gpt (3.5)
 - LLM (Large Language Model) Classification Tasks with Generative AI
 - Content Generation and Summarization with Generative AI
 - Information Retrieval and Synthesis workflow with Gen AI
 - Time Series Forecasting using Deep Learning
 - Time Series and Forecasting
 - Seasonal-Trend decomposition using LOESS (STL) models.
 - Bayesian time series analysis.
 - MakerSuite Google
 - PaLM API
 - MUM models
 - Azure ML

GEN-AI MASTERS PROGRAM — CURRICULUM



Getting started with Gen AI

- Introduction to GenAI and LLMs
- Pre training and Fine Tuning
- Real-time Support Chatbot use-case
- RAG, RLHF, Longchain, Few Shot Learning
- Llama model

Internals of Transformers

- Transformers Architecture
- Transformers Architecture
- Token Embeddings
- Tokenization, Embeddings, Context



Gen-AI Practicals

- Tasks (Sentiment Analysis)
- Text Summarization | Translation | Question - Answer Tasks
- Token Classification | Fill Mask | Text Generation
- Feature Extraction | Zero Shot Classification
- Selecting the Right Model for the Task | Pre-Processing
- Model Inference | Working of Softmax
- More about Tokenizer - Auto Tokenizer | Auto Model
- Introduction to huggingface - Downloading a Model from HuggingFace | Execute NLP

Transformers Internals | Attention Blocks

- Model selection
- Preprocessing - tokenizer
- Postprocessing - SoftMax
- Attention Block and Multilayer Perceptron : Decoding Attention Pattern
- Attention Pattern : Vectors and Matrices
- Embedding matrix & Unembedding matrix
- Query matrix & Query vector | Key vector & Key matrix
- Attention Mechanism : Embedding the Contextual Knowledge
- Multi Layer Perceptron : Feed Forward Layer
- How MLP Works

Training a model

- Introduction to the Bigram Model | Understanding tensors
- Bigram Model
- Tensors
- Pretraining with Matrices
- Word Generation
- Precomputed Probabilities
- Model Quality
- One hot Encoding and Forward Pass
- Backward Pass and Adjusting Weights

Model Pre-Training

- Implementation of Multilayer Perceptron
- Probability Calculation
- Tuning Parameters
- Optimized Industrial Approach - Mini Batches

Getting Started with RAG

- Introduction to RAG
- RAG Practicals (Using Google Colab)
- RAG Practicals Using Databricks CE & Local
- Accessing Gated Models
- RAG Internals Explained
- RAG Pipeline : Text to Embeddings
- RAG Pipeline : Text to Embeddings

RAG End-to-End Production Pipeline

- End-to-End Production Grade Pipeline
- Generating Embeddings
- Building Retriever and RAG LangChain Creation
- Registering Lang-chain and Creating Serving Endpoint

LangChain Essentials

- Understanding LangChain Framework
- Inferring Large Models on Cloud
- Working with OpenAI
- Message Structure : System | Human | AI Message
- Usecase : Few Shot Learning
- Prompt Template
- Task & Chain : Runnable Lambda & Runnable Sequence
- Runnable Parallel

RAG Application using LangChain

- Advanced RAG Application Use-case with Langchain
- Document Loaders
- Chunking Strategies & Embedding Model Selection
- Retriever Configs | Search Types
- Conversational RAG Solution
- RAG Challenges

Query Optimization

- Query Transformation
- Query Routing
- Indexing Strategies
- Tracing and debugging with LangSmith
- More on Query Transformation Techniques - Query Rewrite
- Query Optimization Techniques - Multi Query
- Rag Fusion | Reciprocal Ranking
- Query Decomposition | Sub-Query | Chain of Thoughts(COT)
- Query Decomposition | Multiple Independent Sub-Queries
- Step Back Questions | Few Shot Prompting

AGENTIC AI

- Agentic Behavior in AI
- Introduction to LangGraph
- LangGraph Use Case: Natural Language to PySpark DataFrame
- LangGraph Use Case: PySpark DataFrame to Spark SQL
- Binding Tools to LLMs Using LangGraph

AGENTIC AI TOOLS

- Registering Tools with LLMs in LangGraph
- Tool Binding with ReAct Architecture
- State Persistence in LangGraph Using Checkpoints
- Structured Outputs with Pydantic Models | Reducers to Retain Full Conversation History
- Handling Long Conversations with Message Filtering
- Dynamic Summarization in LangGraph Agents
- Persisting LangGraph State with SQLite Checkpointers

CRAG | AGENTIC AI PROJECTS

- Corrective RAG (CRAG)
- Building a LangGraph based CRAG Application
- Agentic AI Project - SQL Querying Agent with Tools
- Designing a Multi-Role LLM System with Agents

LLM Fine-Tuning

- Introduction to LLM Fine-tuning
- RAG Vs Fine-Tuning
- Types of Fine-Tuning
- Full Fine-Tuning & Parameter Efficient Fine-Tuning (PEFT)
- LoRA | QLoRA (Reducing the Memory & Compute Requirement)
- Practicals Fine-Tuning Demo
- Hyperparameters