

Machine Learning and Generative AI Bootcamp

Duration: 15 days (120 hours)

Labs: Open Source platform and Koenig DC will be provided

Pre-requisite: Fundamental knowledge of Python/similar programming is a must.

Module 01: Python for Machine Learning

A. Python for Data Analytics (4 days)

- Introduction to Python and Basics
- Data types and Control Flows
- Lists, Tuples, Sets, and Dictionaries
- [Lab: Understanding Data Types and Control Flows](#)
- File Handling and Strings
- [Lab: File Handling and String Manipulation in Python](#)
- Iterators & Generators
- Regular Expressions
- [Lab: Regular Expressions for Text Processing](#)
- OOPS Concept
- [Lab: Exploring Object-Oriented Programming Concepts in Python](#)
- Introduction to NumPy, Pandas, SciPy, Seaborn & Matplotlib
- [Lab: Getting Started with Data Science: Introduction to NumPy, Pandas, and matplotlib](#)
- Introduction to Data Preprocessing
- [Lab: Introduction to Data Preprocessing](#)
- Exploratory Data Analysis (EDA)
- [Lab: Data Unveiled- Exploring Insights through EDA](#)
- Hands-on: Project Work- Exploring Data Science Techniques

B. Computer Vision with OpenCV (3 days)

- Introduction to OpenCV - Introduction to OpenCV, Image basics with Numpy, Open an Image with Matplotlib, Get familiar with RGB channels, Differences between Matplotlib and OpenCV
- [Lab: Reading, writing and displaying images](#)
- [Lab: Color Spaces](#)
- [Lab: OpenCV HSV, RGB and Channels](#)
- Image Processing - Introduction to Image Processing, Transform, Resize & Flip an Image, Draw Shapes on an Image, Event Choices for the Mouse, Transformations, Affine and Non-Affine, Image Translations and Rotations, Scaling, Resizing and Cropping
- [Lab: Translations, Scaling, re-sizing and interpolations, Cropping, Different Operations and Sharpening](#)
- Image Segmentation - Segmentation and Contours, Sorting and Matching

- Contour Shapes, Line Detection Game, Circle Detection, Blob Detection
- [Lab: Sorting Contours, Matching Contours Shape](#)
- Object Detection - Object Detection Overview, Feature Description Theory, Finding Corners, SIFT, SURF, FAST, BRIEF & ORB
- [Lab: Finding Corners, Face & Eye Detection](#)
- Object Tracking - Filtering by Color, Background Subtraction and Foreground Subtraction, Using Mean Shift
- [Lab: Background Subtraction](#)

C. Data Mining for Machine Learning (1 day)

- Understanding Data Sources and Acquisition
- Web Scraping Techniques - Overview of Web Scraping, Legal and Ethical Considerations, Tools and Libraries for Web Scraping
- [Lab: Web Scraping for Data Collection](#)

D. Database Operations for Machine Learning (1 day)

- Introduction to Python Database Operations - Overview of Database Management Systems (DBMS), Introduction to SQL and Python Database APIs
- Basic Python Database Operations - Connecting to Databases, Executing SQL Queries, Fetching Data from Databases
- [Lab: Implement CRUD operations](#)

Module 02: Document OCR using Deep Learning (1 day)

- Introduction to Document OCR with Deep Learning
- Data Acquisition and Pre-processing
- Building Deep Learning Models for OCR
- Training and Evaluation
- [Lab: Implementing Document OCR with Deep Learning](#)

Module 03: Predictive Modelling (1 day)

- Introduction to Predictive Modelling
- Data Preparation for Predictive Modelling
- Building Predictive Models
- Evaluating and Tuning Predictive Models
- [Lab: Predictive Modelling Experiment - Regression](#)
- [Lab: Predictive Modelling Experiment - Classification](#)
- [Lab: Predictive Modelling Experiment - Clustering](#)

Module 04: Natural Language Processing (2 days)

- Introduction to NLP and its applications
- Overview of NLTK, SpaCy, and other NLP libraries
- Setting up NLTK environment and configuring for text processing
- Understanding corpora in NLP and their role in training models
- Techniques for importing and reading text files into Python
- Text mining for extracting valuable insights from textual data
- Text pre-processing techniques like tokenization and stemming

- POS-tagging to categorize words by their grammatical properties
- Named Entity Recognition (NER) using SpaCy or NLTK libraries
- One Hot Encoding in NLP and its challenges
- Importance of Vectorization Techniques
- [Lab: Analyzing Text Data with NLP Techniques](#)

Module 05: Generative AI for Image Captioning (2 days)

- Transformer Models and BERT Model
- [Lab: T5 transformer for text related tasks](#)
- Introduction to Generative AI
- Introduction to Large Language Models (LLMs)
- Encoder Decoder Architecture
- Attention Mechanism
- [Lab: Text generation using LLM model](#)
- Introduction to Image Generation
- [Lab: Creating an Image Captioning Model](#)