



GLOBAL CERTIFICATE IN DATA SCIENCE AND ARTIFICIAL INTELLIGENCE®

Online | 11 months | 11 Terms



PROGRAM SNAPSHOT

- TERM 01
 - Data Analysis basics with Python
- TERM 07
 - ► Tensorflow & Keras for Al

- TERM 02
 - Data Visualization & EDA
- * TERM 08
- Deep learning foundation

- TERM 03
 - Machine Learning Foundation
- TERM 09
- ▶ Deep Learning for NLP-I

- **TERM 04**
- Machine Learning Intermediate
- **TERM 10**
- ► Computer Vision-I

- TERM 05
 - ML Basics for AI

- TERM 11
- SpecialisationDeep Learning for NLP-II
- SpecialisationComputer vision-II
 - Al Capstone project

TERM 06

Capstone Project & IndustryImmersion







Module 1 : Data Science Fundamentals

- Thought Experiment: Data Science from a layman's perspective
- Brief intro to Data Science
- How companies use Data Science
- Overview of Data Science project lifecycle
- Walkthrough of data types and data challenges

Module 2 : Recap of Python for Data Science

- In class quiz for Python Basics
- Common Python concepts and sample questions
- Variable, Inbuilt datatyes, functions, modules and Packages
- File operations and error handling

Module 3 : Recap of Statistics for Data Science

- In class quiz for Descriptive Statistics
- Common charts used
- In class guiz for Inferential Statistics
- Probability, Central Limit theorem, Normal Distribution & Hypothesis testing

Module 4: Mathematical operations using Numpy

- Introduction to Numpy Arrays
- How to apply mathematical operations in Numpy
- Array manipulation using Numpy
- Broadcast values across Arrays using Numpy

Module 5 : Data manipulation with Pandas

- Types of Data Structures in Pandas
- Clean data using Pandas
- Manipulating data in Pandas
- How to deal with missing values
- Hands-on: Implement Numpy arrays and Pandas Dataframes



Mini Project : Data manipulation Projects

- Project Expectation setting
- Project Timelines
- Github for Project Submission
- Presentation guidelines



Module 1 : Data Visualization in Python - 1

- Plotting basic statistical charts in Python
- Data visualization with Matplotlib
- Case study: Analysis of Wine dataset through visualizations

Module 2 : Data Visualization in Python - 2

- Statistical data visualization with Seaborn
- Interactive data visualization with Bokeh
- Case study: Analysis of Fifa data using Seaborn & Bokeh packages

Module 3 : Exploratory Data Analysis - 1

- Introduction to Exploratory Data Analysis (EDA) steps
- Plots to explore relationship between two variables
- Histograms, Box plots to explore a single variable
- Heat maps, Pair plots to explore correlations
- Case study: Perform EDA to explore survival using Titanic dataset

Module 4: Industry immersion & project discussion

- Interaction with Industry experts
- QnA

Module 5 : Exploratory Data Analysis - 2

• Case study: Analyse mental health of IT professionals



Module 6: Project presentation

•Top 3 teams present their projects details to the entire cohort





Term 3: MACHINE LEARNING FOUNDATION

Module 1 : Introduction to Machine Learning (ML)

- What is Machine Learning?
- Use Cases of Machine Learning
- Types of Machine Learning Supervised to Unsupervised methods
- Machine Learning workflow

Module 2: Linear Regression

- Introduction to Linear Regression
- Use cases of Linear Regression
- How to fit a Linear Regression model?
- Evaluating and interpreting results from Linear Regression models
- Case study: How linear regression helps determine demand

Module 3 : Logistic Regression

- Introduction to Logistic Regression
- Logistic Regression use cases
- Understand use of odds & Logit function to perform logistic regression
- Case study:Predicting default cases in the Banking Industry

Module 4 : Decision trees & Random Forests

- Introduction to Decision Trees & Random Forest
- Understanding criterion(Entropy & Information Gain) used in Decision Trees
- Using Ensemble methods in Decision Trees
- Applications of Random Forest
- Case study:Predict passengers survival in a Ship mishap

Module 5 : Model evaluation techniques

- Introduction to evaluation metrics and model selection in Machine Learning
- Importance of Confusion matrix for predictions
- Measures of model evaluation Sensitivity, specificity, precision, recall & f-score
- Use AUC-ROC curve to decide best model
- Case study: Applying model evaluation techniques to prior case study





Term 4:

MACHINE LEARNING INTERMEDIATE

Module 1 : Dimensionality Reduction using PCA

- Unsupervised Learning: Introduction to Curse of Dimensionality
- What is dimensionality reduction?
- Technique used in PCA to reduce dimensions
- Applications of Principle component Analysis (PCA)
- Case study: Optimize model performance using PCA on high dimension dataset

Module 2 : KNN (K- Nearest neighbours)

- Introduction to KNN
- Calculate neighbours using distance measures
- Find optimal value of K in KNN method
- Advantage & disadvantages of KNN
- Case Study:Classify malicious websites using close neighbour technique

Module 3 : Naïve Bayes classifier

- Introduction to Naïve Bayes classification
- Refresher on Probability theory
- Applications of Naive Bayes Algorithm in Machine Learning
- · Case study: Classify Junk emails based on probability

Module 4 : K-means clustering technique

- Introduction to K-means clustering
- Decide clusters by adjusting centroids
- Find optimal 'k value' in kmeans
- Understand applications of clustering in Machine Learning
- Case study: Segment flower species in Iris flower data

Module 5 : Support vector machines (SVM)

- Introduction to SVM
- Figure decision boundaries using support vectors
- Identify hyperplane in SVM
- Applications of SVM in Machine Learning
- Case Study: Predicting wine quality without tasting the wine



Module 6: Time series forecasting

- Introduction to Time Series analysis
- Stationary vs non stationary data
- Components of time series data
- Interpreting autocorrelation & partial autocorrelation functions
- Stationarize data and implement ARIMA model
- Case Study: Forecast demand for Air travel





MACHINE LEARNING BASICS FOR AI

Module 1 : Intro to Machine Learning

- What is Machine Learning?
- Use Cases of Machine Learning
- Types of Machine Learning Supervised to Unsupervised methods
- Machine Learning workflow

Module 2: Linear Regression

- Introduction Supervised learning method: Linear Regression
- Use cases of Linear Regression
- How to fit a Linear Regression model?
- Evaluating and interpreting results from Linear Regression models
- Case study: Predict Bike sharing demand

Module 3: Logistic Regression

- Introduction to Supervised Learning: Logistic Regression
- Logistic Regression use cases
- Applying Logistic Regression using Logit function in Python
- Case study: Predicting credit card default cases

Module 4 : Model Evaluation

- Introduction to evaluation metrics for Machine Learning models
- Interpreting AUC-ROC curve
- Weights and Bias estimation using gradient descent optimization.
- Case study: Comparison of models Credit card default cases

Module 5 : Gradient Descent & Optimization

Learn common Machine Learning optimization "tricks"

- Initialization
- L2 and dropout regularization
- Batch normalization
- Gradient checking





James Telco Bond

In this capstone project, students will be provided with data collected by a major Telecom operator on the demographic behaviour of users using different handsets.

Students are required to do the initial bit of data cleansing, pre-processing and then upload this data to SQL server via a web hosting platform that will be provided to them.

This data from SQL server will be used to create a dashboard for the company using D3.js scripts. D3,js scripts will be provided to students upfront. These dashboards are reflective of how interactive visualizations can help companies make strategies such as what demographies to cater to, how men and women customers behave differently, which geographies are popular and ones that need more investment from the company in terms of finance and marketing?





Module 1 : Introduction to TensorFlow

- What is TensorFlow
- Hello World prgram in TF
- How does a TF program work Computational Graphs

Module 2: TensorFlow Basic Syntax

- Commonly used syntax
- Variables and Placeholders in TF
- 'Hello World' in TensorFlow

Module 3 : Programming in TF - |

- Common TF constructs and syntax
- Building a Regressor in TensorFlow
- ML model visualization using Tensorboard

Module 4: TF Regression Exercise - 1

- Building Linear regression models on TF
- Visualize using Tensorflow graphs

Module 5 : Programming in TF - ||

- Building Classifier in TensorFlow
- Predict class probabilities using Softmax
- Training with large data using Mini-Batching



Module 6: TF Regression Exercise - 2

- Building Logistic regression models on TF
- Visualize using Tensorflow graphs

Module 7 : Intro to Keras - 1

- What is Keras
- Keras vs TensorFlow
- Building Regressor & Classifier in Keras
- Using different hyperparameters in Keras

Module 8: TF Regression Exercise - 1

- Building Linear regression models on Keras
- Building Logistic regression models on Keras



Module 1: Introduction to Deep learning

- Basics of Neural Networks
- How do Neural Networks work?
- How do Deep Neural networks learn?
- Buiding a DNN classifier in TensorFlow

Module 2: Components affecting Deep Learning Models

- Learn what components affect DL models
- Simple implementations of DL models
- How to evaluate the models

Module 3 : Improving Deep Learning Models - |

- Learn how Activation Functions trigger neurons
- Review different Activation functions
- How to make Machines generalize and not memorize?
- Speed of learning Learning Rate for Machines

Module 4 : Overview of NLP

- Brief walk into the world of NLP
- Common problems that NLP solves
- Few solutions by popular companies in the NLP space

Module 5 : Overview of Computer vision

- Brief walk into the world of CV
- Common problems that CV solves
- Few solutions by popular companies in the CV space



Module 6: Improving Deep Learning Models - ||

- Limitations of Gradient Descent
- Overcome local minima problem using Momentum
- Tuning using different Optimizers -Nestrove, AdaGrad,

Module 7 : Deep Learning Models Practicals - |

- Manual Creation of Neural Network Part One
- Manual Creation of Neural Network Part Two Operations

Module 8 : Deep Learning Models Practicals - ||

- Manual Creation of Neural Network Part Three Placeholders and Variables
- Manual Creation of Neural Network Part Four Session





Term 9: **NLP-I**

Module 1 : Intorduction to NLP & Word Vectors

- Introduction to NLP
- Converting text to Numbers
- Using Word2Vec to covert text to numbers
- Using Pre-Built Word2Vec embeddings

Module 2: Decoding a textual data

- Implement text to number operations
- Build Word2Vec using regular NLTK package
- Identify word to number relationship

Module 3 : NLP using Recurrent Neural Networks

- How can we make Neural networks remember the past.
- Introduction to RNN
- LSTM and Gated Recurrent Unit (GRU) theory

Module 4: NLP using Memory alterations

- Using Long Short-term Memory (LSTM) in place of RNN
- RNN Plotting
- Using RNN/LSTM/GRU for classification in Keras

Module 5 : Teach machines to generate new textual data

- Language Modeling generating new Text data
- Building Character Level RNN (Char-RNN)
- Building Word Level RNN



Module 6: Build new dialogues using Keras

- The Republic by Plato introduction
- Preparation of Data
- Train Language Model
- Use the language model

Module 7: Translate language e.g. English to Spanish using Seq2Seq Models - I

- Understanding requirements of Machine Translation
- Introduction to Sequence to Sequence (Seq2Seq) model
- Understanding Keras Functional API

Module 8 : Build a simple language translator

- Using Standford Core NLP package
- Identify grammar
- Understand interlingual approaches
- Translation models in Standford Core NLP package





Module 1 : Intro to Convolutional Neural Networks

- Understand Convolution
- Techniques applied in a CNN Pooling, Padding
- Components of a Convolutional Neural Network (CNN)
- Building a CNN in Keras

Module 2 : Decoding image components

- · Popularly used techniques for image processing
- How to collect, resize and reduce image
- Components of an image

Module 3 : Use of CNN in Computer Vision (CV)

- Role of ImageNet in CV
- Common architectures built using CNN
 - a. VGG Network
 - b. Inception Network
 - c. Resnet Architecture

Module 4 : Identifying MNIST using CNN

- Try identifying numbers using CNN
- Apply few common architectures to identify better

Module 5 : Using pretrained CNN models

- Can machines Learn from each other:
- Understanding Transfer Learning*
- Image classification using Transfer Learning



Module 6: Transfer learning on ImageNet Data

- Train a model using ImageNet data
- Classify objects using CNN Models

Module 7 : Breaking down what CNN learns?

- Understand what CNN is learning
- Visualize CNN filters
- Build some fun applications of CNN
 - a. Neural Style transfer
 - b Daydream

Module 8 : Preprocessing Image data to apply CNN

Augment Image data for better learning

- Scaling
- Translation
- Rotation
- Flipping





Term 11: [Specialisation]

NLP-II

Module 1:
Deep dive in
language translation
- English to Spanish
using Seq2Seq
Models - II

- Building Seq2Seq Model in Keras
- Working with different outputs of LSTM and GRU in Keras
- Stacked LSTM
- Bi-Directional LSTM
- Examples where Seq2Seq model can be used

Module 2: Hands-on: Optimize language translator Eng-Spanish using Seq2Seq Models - II

- Find how KSTM & GRU affect the language translator
- Evaluate methods that are helpful for language trnslation

Module 3:
Techniques to enhance
Seq2Seq Models
(Attention Mechanism)

- What is Attention Mechanism
- Applying Attention to Seq2Seq models

Module 4: NLP using CNN

- Making a case for CNNs to use in NLP
- Using CNN for Sentiment Analysis
- Using Attention with CNN for NLP

Module 5 : Help Neural Networks to memorize

- What is Question-Answering (Q&A) System
- Why they are difficult to build.
- What are Memory networks



Module 6: Memory based Neural Networks

- End to End Memory Network (MemN2N) for Q&A
- Dynamic Memory Networks (DMN)

Module 7 : Building a Chatbot - Part 1

- Steps to move a model to Production
- Deploying Model as Web Service

Module 8 : Building a Chatbot - Part 2

- Creating an application in Slack
- Integration Slack with ML Model





Module 1:
Object Detection in an image using
CNN based algorithms -I

- What is required to detect multiple objects in an Image
- Using Region based CNN (R-CNN) for Object Detection
- Improving R-CNN using Fast R-CNN network
- Going a step further Faster R-CNN network

Module 2 : Cat and Dog classification

- Study the cat and dog dataset
- Build a simple classifier that identifies the type

Module 3:
Object Detection in an image using CNN based algorithms -II

- How do we detect object in Real time.
- Understanding Single Shot Detection (SSD) approach
- You Only Look Once (YOLO) for Object Detection

Module 4:
Object detection at pixel Level (Instance Segmentation)

- How do identify objects at pixel level
- Mask R-CNN model for Instance Segmentation

Module 5 : How to trick Neural networks (Adversarial ML) and ways to safegaurd

- Adversarial ML Can we trick neural networks?
- Building an Adversarial use case
- Techniques to defend against Adversarial examples



Module 6: Teaching machines to mimic real world (Generative Models)

- Can machines build new things by observing the world around them?*
- Understanding Generative Adversarial Network (GAN)
- Building DCGAN in Keras

Module 7 : DCGAN in action

• Use Deep Convolutional GAN to generate color images of house numbers.





Demand Planners

This capstone project will focus more on applying machine learning concepts rather than data gathering and storing aspects. Students will be provided with data collected by a major Taxi Aggregator of taxi bookings done in a leading city. As budding data science consultants, students are required to do exploratory data analysis & present an initial report.

After that the students are required to create an UI that displays the observations regarding taxi usage across the city from the analysis and the website should also have a provision for the company to forecast demand for taxis at a specific time in the day.

The taxi bookings data provided will be in csv format and dashboards for the company need to be created using D3.js scripts. The D3.js scripts will be provided to the students beforehand.



Need to know

Program Start

> First week of every month

Duration

> 11 months (Incl. Capstone Projects)

Prerequisite

- > Background in Programming (Mandatory)
- > Laptop with 4 GB RAM

Program Fee

INR 3 Lakh + GST

Scholarships

50 scholarships (each with 70% tuition waiver) for professionals passionate about making a career in Data Science & furthering INSAID's mission of putting India on the global AI map.

Talk to our Admissions team today or attend the next Data Science MasterClass to know more

For Further details, write to us info@insaid.co



