EXPLORE

MACHINE LEARNING
PROSPECTUS
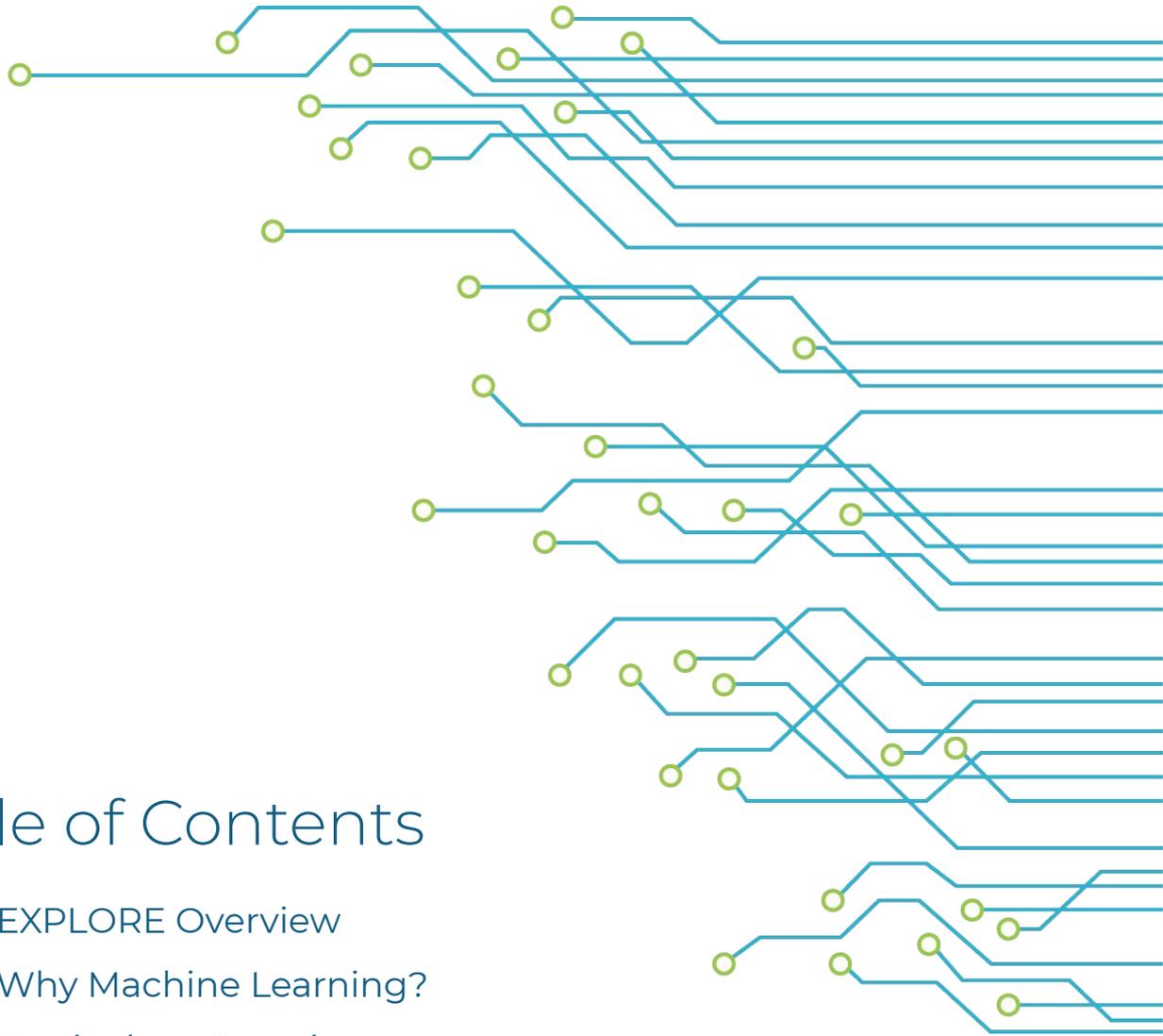
Table of Contents

EXPLORE

EXPLORE Overview

EXPLORE is a next generation Learning Institution that teaches students the skills of the future. From Data Science to Data Engineering to Machine Learning to Deep Learning we deliver cutting edge courses to satisfy your hunger to learn. Our Programmes are built by an amazing Faculty - we employ some of the world's most talented Scientists who have experience solving difficult problems on a global stage.

Our philosophy is to teach our students how to solve problems in the real world. We emphasise team-work, collaboration and working within constraints, under pressure, with deadlines while understanding context, audience and implementation challenges. We are not a theoretical institution (although we cover the theory) - we are a 'practical, hands-on, roll-up-your-sleeves and get stuff done' kind of institution. As real-world Scientists who have delivered impact in the world of work we're well positioned to deliver these skills.

EXPLORE launched during 2013 and since then has taught 1,000's of students and solved many problems for businesses across multiple Industries across the world. We're reinventing education and invite you to join us to change things for the better.

EXPLORE

Why Machine Learning?

Four megatrends are fundamentally changing the shape of our world:

1. Vast amounts of data are being generated every minute.
2. The processing speed of our machines is increasing exponentially.
3. We now have cloud providers who can store insane amounts of data for a few dollars.
4. Powerful open source algorithms that can read, write, translate and see are now available to everyone.

All of these things above mean that it's becoming easier and quicker to produce models that can analyze bigger/more complex data and deliver more accurate results – even on a very large scale. By building prediction models a business has a better chance of identifying profitable opportunities – or avoiding unknown risks.

GROWTH IN DATA SCIENCE JOBS SINCE 2012

650%

2012 2013 2014 2015 2016 2017 2018 2019

EXPLORE
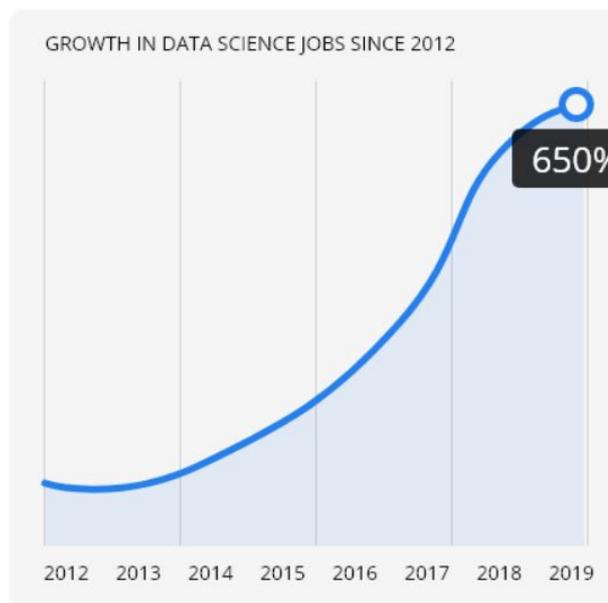
Curriculum Overview

The course will start by covering the basics of programming in python - building functions to automate analysis. After the foundations have been set, participants will master the model building process and the various machine learning algorithms to use as well as exploring techniques such as clustering and natural language processing. In the final phase participants will dive deeper into the more complex machine learning techniques - deep learning, where they will cover all the various neural networks.

Phase	Module	Weeks
Programming	Python Fundamentals	3 weeks
	Python Data Structures	3 weeks
Machine Learning	Classification Techniques	4 weeks
	Regression Techniques	4 weeks
	Unsupervised Learning	4 weeks
	Deep Learning	4 weeks

EXPLORE

Python Fundamentals

Project: Build a financial calculator

What is covered in Module 1:

Python Programming Basics

Basic Concepts
- ➔ What is programming
- ➔ Installing Jupyter environment
- ➔ Pseudo code and debugging concepts
- ➔ Interactive vs scripting mode

Variables, Expressions and Statements
- ➔ Print statements
- ➔ Working with primitive data types - variables, strings, integers, floating points, booleans
- ➔ Math functions and relational operators
- ➔ Importing modules

Logic and Functions

Working with Logic
- ➔ Conditional statements - IF and ELSE IF
- ➔ Working with lists
- ➔ For loops and while loops
- ➔ Break | Continue principles

Writing Functions
- ➔ How functions work - basic syntax
- ➔ Creating and working with functions
- ➔ Parameters in functions
- ➔ Local and global variables

EXPLORE

Module 2

Python Data Structures

①②③❹❺❻⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰⑱⑲⑳㉑㉒

Project: Build functions to select the most optimal football team

What is covered in Module 2:

Data Types

Working with Basic Structures
- ➜ What is a String?
- ➜ What is a Number?
- ➜ What is a Boolean?
- ➜ Working with Strings, Numbers, Booleans

Working with Lists and Tuples
- ➜ Lists and Tuples Semantics
- ➜ Lists and Tuples Operations
- ➜ Working with Comparisons
- ➜ Working with Statements

Dataframes and using Libraries

Working with Sets and Dictionaries
- ➜ Sets and Dictionaries Semantics
- ➜ Sets and Dictionaries Operations
- ➜ Working with Comparisons
- ➜ Working with Statements

Working with Data Sets
- ➜ Importing Data - using Numpy and Pandas libraries
- ➜ Working with Data Frames
- ➜ Retrieving and Storing Data
- ➜ Transforming Data

EXPLORE

Classification Techniques

Project: Predict if someone will take out a personal loan or not

What is covered in Module 3:

Pre-processing and Model Building

Preparing data for modelling
- ➔ Basics of Scikit Learn library in Python
- ➔ Define features and labels
- ➔ Train/test split
- ➔ Encoding categorical features
- ➔ Standardizing features

Steps to build a model
- ➔ Overview of classification
- ➔ Logistic regression models
- ➔ Build and train models in python
- ➔ Make predictions using trained model
- ➔ Test model output - confusion matrix, classification report

Classification Algorithms

Algorithms for classification models
- ➔ K nearest neighbors
- ➔ Decision trees
- ➔ Random forests
- ➔ Support vector machines

Optimizing model performance
- ➔ Optimizing model performance by tuning hyperparameters
- ➔ Gridsearch techniques
- ➔ Which models to use when

EXPLORE

Regression Techniques

Project: Using features of a house alone, predict the sale price

What is covered in Module 4:

Regression Algorithms

Steps to build a model
➔ Overview of regression
➔ Linear regression models
➔ Train models and make predictions
➔ Regularization techniques - Ridge, Lasso
➔ Test model output - MSE, MAE, RMSE

Algorithms for regression models
➔ K nearest neighbors
➔ Decision trees
➔ Random forests
➔ Support vector machines
➔ Optimizing model performance by tuning hyperparameters

Dimensionality Reduction

Feature selection
➔ Determining significant features - linear and non-linear methods
➔ Variance Thresholds to drop features
➔ Correlations Thresholds

Features extraction
➔ Principal component analysis
➔ Linear discriminant analysis
➔ Interpreting extracted features and transforming them back

EXPLORE

Unsupervised Learning

Project: Cluster customer reviews into themes

What is covered in Module 5:

Natural Language Processing

Overview of NLP
- ➔ Overview of NLTK library in python - for NLP
- ➔ Removing punctuation and symbols
- ➔ Tokenizing text
- ➔ Stopwords and removing them

Analysing text
- ➔ Lemmatization of words
- ➔ Bag of words
- ➔ Sentiment analysis

Clustering

Hard and Hierarchical Clustering
- ➔ What is clustering?
- ➔ Various forms of clustering - hard, hierarchical, soft
- ➔ K-Means Clustering
- ➔ Hierarchical Clustering

Soft Clustering
- ➔ Gaussian Mixture Models
- ➔ LDE clustering of text
- ➔ Labelling data using cluster output

EXPLORE

Deep Learning

① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩ ⑪ ⑫ ⑬ ⑭ ⑮ ⑯ ⑰ ⑱ **⑲ ⑳ ㉑ ㉒**

Project: Predict the characters present in a handwritten image

What is covered in Module 6:

Artificial neural Network

How a Neural Network works

➔ Overview of a neuron
➔ Structure of a neural network
➔ Backpropagation and gradient descent
➔ Learning rate and batch sizes when training a neural network

Building and using ANN's

➔ Building classification / regression ANN models in keras
➔ Make predictions using ANN and test model output
➔ Tuning hyperparameters (learning rate, batch size, nodes in each layer)

Other Neural Networks

Building and using CNN's

➔ Building classification / regression CNN models in keras
➔ Make predictions using CNN and test model output
➔ Tuning hyperparameters

Building and using RNN's

➔ Building classification / regression RNN models in keras
➔ Make predictions using RNN and test model output
➔ Tuning hyperparameters

EXPLORE

EXPLORE Philosophy:
Solving problems in the real world

At EXPLORE we focus on building our student's ability to solve problems in the real world. Building things that work and make a difference is hard - that's what we teach.

We're not a traditional learning institution that spends weeks teaching matrix multiplication on a whiteboard (although understanding that is useful) - we're a practical, solution-orientated institution that teaches our students to work in teams, under pressure, with deadlines while understanding context, constraints and the audience.

Our courses are typically broken into Sprints where we teach a core set of concepts within the framework of solving a problem in a team with a tight deadline.

Students cycle from Sprint to Sprint solving different problems in different teams as they build this core muscle over the course.

EXPLORE

Contact Information

For any **admissions related enquiries** mail us on:

admission@explore-ai.net

For any **general enquiries** mail us on:

general@explore-ai.net

Visit: www.explore-datascience.net