TECH TALENT SOUTH

DATA SCIENCE TRACK

Syllabus - Data Science - Learning Python and other topics in Data and Machine Learning

Data Science Topics Overview
- Introduction to the Python programming language
- Python data science packages like NumPy and Scikit-Learn
- Data analysis of and visualization of large datasets

This course will give you the opportunity to learn Python, the first-class tool for working with scientific data and large datasets. You'll also be introduced to more Machine Learning topics and the enterprise-level tools used to create business solutions around Machine Learning.

Pre-requisites
Prior knowledge of MS Excel and a desire to step into the bigger world of data science.

Topics Outline
- Intro to Python 3
- How to create and share documents with live code and visualizations using Jupyter Notebook
- Storing and manipulating data using NumPy and Pandas
- Create data visualizations in Python using Matplotlib
- Machine Learning algorithms and Scikit-Learn
- Solving real problems with Microsoft AzureML

Course Schedule

Week 1: Introduction to Python
- how to run Python code
- Python's basic syntax
- Variables and Operators
- Data Structures

Week 2: Python Basics Continued
- Control Flow statements
- Defining functions
- Error handling

Week 3: Working with Data Basics
- Iterators
- List Comprehensions
- Generators and Generator Expressions
- Regular Expressions

Week 4: Object-Oriented Programming in Python
- basic OOP principles
- creating an Object in Python
- Inheritance
- Defining and working with Methods

Week 5: IPython and NumPy
- IPython as our "control panel"
- Jupyter Notebook
- NumPy Arrays and Computation

Week 6: Data Manipulation and Visualization
- Pandas Objects
- Operating on Data
- Visualizing Data using Matplotlib

Week 7: Introducing Machine Learning
- Introducing Scikit-Learn
- Model validation
- Feature Engineering
- Using AzureML

Week 8: Machine Learning Continued
- Classification and Regression
- Decision Trees
- Clustering