Cognate/Professional Electives

Introduction to Data Science

Renato R. Maaliw III, *DIT*
Professor 1, College of Engineering
Southern Luzon State University
Lucban, Quezon, Philippines

Cognate/Professional Electives

Data Science is a multidisciplinary field that combines statistical analysis, computer science, and domain-specific knowledge to extract meaningful insights from data.

Cognate/Professional Electives

It involves various techniques to process, analyze, and interpret large volumes of structured and unstructured data.

By leveraging algorithms, machine learning, and statistical models, you can uncover patterns and trends that inform decision-making.

Cognate/Professional Electives

In today's data-driven world, your ability to harness data effectively will be your competitive advantage

CPE15 – Programming for Data Science

CPE21 – Machine Learning 1

CPE28 – Machine Learning 2

Week 1 (Preliminaries)

- Data science overview
- IDE (Setup, offline and online)
- Jupyter Notebook overview
- Python Basics – Part 1
  a. data types, variables, list, dictionary, boolean, tuple, sets
  b. comparison operators
  c. logical operators
  d. if constructs
  e. loops

Week 2 (Numpy)

- Array manipulations: reshaping, flattening, concatenation, splitting, etc.
- Universal functions and aggregation
- Working with multidimensional arrays
- Linear Algebra operations with Numpy

Week 3 (Introduction to Pandas)

- Series & DataFrames
- Data Loading, Storage, and File Formats
- Indexing and Selecting Data
- Handling Missing Data

Week 4 (Data Manipulation with Pandas)

- DataFrame Operations: Merging, Joining, Concatenating, and Resphaping
- GroupBy Operations and Aggregations
- Functions: Mapping, Applying, and Vectorizing
- Handling Time Series Data with Pandas

Week 5 (Data Cleaning & Preprocessing)

- Data Cleaning: Duplicates, Inconsistencies, and Outliers
- Data Transformation: Scaling, Normalization, and Encoding
- Working with Text Data in Pandas
- Handling Missing Data for DataFrames

Week 6 (Introduction to Data Wrangling)

- Data Wrangling Fundamentals
- Data Inspection and Profiling
- Common Data Wrangling Tasks and Challenges
- Introduction to Regular Expressions for Data Cleaning

Week 7 (Data Wrangling Techniques)

- Imputation Techniques
- Combining and Reshaping Datasets
- Outlier Treatment
- Feature Engineering for Data Science

Week 8 (Data Extraction)

- Extracting Data using Native Python
- Extraction using SQL

Week 9 (Web Scraping)

- Introduction to Web Scraping
- Tools: BeautifulSoup, etc.
- Parsing HTML and XML data
- Ethical Considerations and Legal Issues in Web Scraping

Week 10 (Exploratory Data Analysis)

- Statistics and Distributions
- Patterns, Anomalies, and Relationships in Data
- Correlational Analysis

Week 11 (Data Visualizations – Part 1)

- Overview of Matplotlib
- Plots: Line plots, Bar plots, Histograms, etc.
- Plot Cusomizations

Week 12 (Data Visualizations – Part 2)

- Subplots & Grids
- More Customizations
- 3D Plotting

Week 13 (Statistical Data Visualization)

- Overview of Seaborn
- Distribution Plots: KDE Plots, Box Plots
- Relational Plots: Scatter Plots, Line Plots, Pair Plots
- Categorical Plots: Bar Plots, Count Plots, Violin Plots

Week 14 (GeoSpatial Data Visualization)

- Overview of GeoSpatial Data and GeoPandas
- Shapefiles and GeoJSON
- Plotting Geospatial with GeoPandas and Folium

Week 15 (Interactive Data Visualizations)

- Dashboards and Reports
- Best Practices for Effective Data Visualization
- End-to-End Data Science Pipeline Project

Thank you very much for listening.