



# Big Data & Data Science Program Diploma Courses

Date: March 2019 – v 3.0

## Diploma Structure

The Big Data & Data Science Diploma requires the attendance of 4 courses and 1 hands-on group project according to the following structure:

### Semester #1 (2 Courses)

- 1- Introduction to Big Data, Developing with Spark and Hadoop (42 Hours, 14 Lectures)
- 2- Introduction to Machine Learning and Statistical Analysis (42 Hours, 14 Lectures)

### Semester #2 (2 Courses)

- 3- Advanced Big Data Analytics Technologies and Applications (42 Hours, 14 Lectures)
- 4- Only 1 of the 3 following courses:
  - Practical Data Mining (42 Hours, 14 Lecture)

OR

- Practical Data Science Using Machine Learning Technique (42 Hours, 14 Lectures)

OR

- Selected topics in Deep Learning (42 Hours, 14 Lectures)

### Semester #3 (Final Project)

- Hands-on group project based on real life use case (14 Weeks of Mentoring)

Please refer to **Appendix A** for the description of each of those courses.

## Important Notes

- All enrollments are subject to the admission rules and acceptance criteria of Nile University and the Big Data and Data Science Program.
- The default training location in Nile University premises and any change will be decided upon case by case by the program management team.
- Timing, lecture distribution, assigned instructors and schedules will be assigned and announced to students upon registration completion subject to Nile University and the program administrative decisions.
- The courses details and outlines might get changed due to continuous development and enhancements to cope with trending theories, technologies, methods and applications in this domain.

For more details and pricing, please contact us: [bigdata@nu.edu.eg](mailto:bigdata@nu.edu.eg)

## Appendix A: Course Descriptions

### CIT-653: Practical Data Mining (42 Hours, 14 Lectures)

#### Description

Data Mining generally refers to the process of exploring patterns and regularities in data for enabling forecasting, predication, and providing a better understanding of the data for guiding decision making. This course provides an introduction to data mining concepts over structured and un-structured data with special emphasis on practical applications of this important research area. Throughout the course, students will be provided with practical examples related to the topics being covered and which will serve to illustrate concepts being introduced. Students will also learn how to solve real-life problems using state-of-the-art technologies for data analysis through provided hands on assignments, and course projects.

#### Pre-requisites

- CIT-651: Introduction to Machine Learning and Statistical Analysis  
(Or equivalent knowledge subject to NU evaluation)

#### Reference Textbooks

“Mining of Massive Dataset,” by Anand Rajaraman and Jeffrey D. Ullman; Cambridge University Press, 2011.

“Data Mining: Practical Machine Learning Tools and Techniques (Second Edition)” by Ian H. Witten and Eibe Frank, Morgan Kaufmann, 2005.

#### Course Outlines

- Data mining concepts
- Association rule mining
- Introduction to text mining and classification
- Using Python for efficient classification and clustering (implementing pipeline and using grid search for parameter tuning)
- PageRank as an example of social network analysis
- Recommendation systems
- Mining data streams