Name of the Subject

Emphasis

NAME OF THE SUBJECT:
Big Data
- Obligatory ( ): Basic ( ) Complementary ( )
- Elective (X): Intrinsic ( ) Extrinsic ( )

NUMBER OF ACADEMIC CREDITS: Four (4).

COURSE TYPE: THEORETICAL: ___ PRACTICAL: ___ THEORETICAL-PRACTICAL: _X_
Methodological alternatives:
Master Class (X), Seminar ( ), Seminar - Workshop (X), Workshop ( ), Practice (X),
Tutored projects (X), Other: _____

Justification

In a world flooded with data (Big Data), we can make better business decisions, based on the strategic use of data.

- How does data allow me to improve products for my customers?
- How does data improve operation and reduce risks?

What do the world's most successful digital organizations like Amazon and Alphabet (Ggogle) have in common?

> The answer: a focus on the data. Donna Burbank, CEO of Global Data Strategy, Ltd., says the connection between excellent data utilization and business success is not a coincidence. The World Economic Forum has stated that, in the current and future market environment, data is more valuable than physical assets.
> - Data is humanity's new crude

For all the above, it is essential to develop two major academic aspects in our professionals: Big Data Architectures, NO-SQL Databases and the strategic use of data thanks to Artificial Intelligence, Machine Learning and Data Analysis.

PREREQUISITE: Basic Statistics and Computer Sciences

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS
FACULTAD DE INGENIERIA

SYLLABUS

Page 3 de 8

FACULTAD DE INGENIERÍA
Maestría en Ciencias de la
Información y las Comunicaciones

UNIVERSIDAD DISTRITAL
FRANCISCO JOSE DE CALDAS

Content

OBJECTIVE:

The objective of this course is to study and apply in a practical way all the concepts of Big Data Architecture, Machine Learning and Data Analysis, so that the student can create big data and data analysis projects for a company and achieve the digital transformation of the company.

SPECIFIC OBJECTIVES

- At the end of the course the student will be able to:
- Know the Big Data architectures and the different NO-SQL Database models
- Know what data analysis is and its application in business
- Know the different stages of data analysis
- Know different techniques and tools for data analysis
- Know the machine learning and deeplearning techniques for predictive analysis.
- Know the techniques of NLP (Natural Language Processing) for the analysis of unstructured data text.
- Apply data analysis in a business domain using some data analysis tool.

SYNTHETIC PROGRAM:

- INTRODUCTION TO BIG DATA AND THE IMPORTANCE IN THE WORLD TODAY
- BIG DATA
    - DEFINITIONS
    - ARCHITECTURES
        - MAP-REDUCE
        - HADOOP
        - SPARK
        - NO-SQL DATABASES
- DATA ANALYSIS AND SCIENCE
    - DATA AND DECISION-MAKING IN YOUR LIFE AND IN ORGANIZATIONS
- STAGES IN DATA ANALYSIS
- ANALYSIS TYPES
    - DESCRIPTIVE
    - PREDICTIVE
    - PRESCRIPTIVE
    - MACHINE LEARNING
- APPLICATION OF THE DATA ANALYSIS PROCESS

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS
FACULTAD DE INGENIERIA

SYLLABUS

Page 5 de 8

FACULTAD DE INGENIERÍA
Maestría en Ciencias de la
Información y las Comunicaciones

UNIVERSIDAD DISTRITAL
FRANCISCO JOSE DE CALDAS

- ○ DESCRIPTIVE ANALYSIS AND BASIC STATISTICS
- ○ ANALYSIS OF A VARIABLE
- ○ ANALYSIS OF 2 VARIABLES
- ○ ANALYSIS AND GRAPHICAL PRESENTATION OF DATA
- ○ PREDICTIVE ANALYSIS
  - ■ LINEAR REGRESSION MODELING
  - ■ CLASSIFICATION MODELING AND TEXT MINING
  - ■ CLUSTERING AND SIMILARITY: RETRIEVING DOCUMENTS
  - ■ RECOMMENDER SYSTEMS
  - ■ OTHER MACHINE LEARNING TECHNIQUES
  - ■ TREES
  - ■ ETC.
  - ■ DEEP LEARNING.
- ○ NLP NATURAL LANGUAGE PROCESSING

Strategies

METHODOLOGY:
- ● Master presentations by the teacher of all the theoretical concepts of the course.
- ● Development of individual practical cases week to week.
- ● Collaborative activity is essential.

The interaction within the academic space defines a cultural mediation scenario, in which the theoretical contrast, the practical verification and the collective dialogue are strategies for the better assimilation of concepts, the detection of failures in the implementations and the generation of innovation points. .

The presentation of contents is carried out from problem nuclei and the application of the same in their resolution from practical exercises and a class project. The master class is focused on the presentation of example practices, from which previous concepts taken from readings, videos and podcasts will be appropriated as learning objects.

Type of course	Hours			Teacher hours / week	Student hours / week	Total Hours Student / semester	Academic credits
	DW	CW	AW	(DW + CW)	(DW + CW +AW)	X 18 weeks	

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS
FACULTAD DE INGENIERIA

SYLLABUS

Page 6 de 8

UNIVERSIDAD DISTRITAL
FRANCISCO JOSE DE CALDAS

FACULTAD DE INGENIERÍA
Maestría en Ciencias de la
Información y las Comunicaciones

	4	2	3	6	9	144	3

Direct Presential Work (DW): classroom work in plenary session with all students.
Mediated-Cooperative Work (CW): Teacher tutoring work to small groups or individually to students.
Autonomous Work (AW): Student work without the presence of the teacher, which can be done in different instances: in work groups or individually, at home or in a library, laboratory, etc.)

Resources

PHYSICAL RESOURCES REQUIRED:

- Master classes to introduce the basic concepts using different resources: presentations, videos, consultation of links of interest.
- Individual workshops to apply the basic concepts
- Bibliographic review, readings and exhibitions by the students.
- Group development of a course project to implement a case of data analysis and big data applied in a public or private company

BIBLIOGRAPHY:

What is Your Digital Business Model? Six Questions to Help You Build the Next-Generation Enterprise, de Peter Weill, Stephanie Woerner.

Leading Digital: Turning Technology into Business Transformation, de George Westerman, Didier Bonnet, Andrew McAfee.

BIBLIOGRAPHIC RESOURCES:

- IEEE Database
- SPRINGER Database
- ELSEVIER Database
- Kaggle

Course Schedule

Week /Unid	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS
FACULTAD DE INGENIERIA

SYLLABUS

Page 7 de 8

UNIVERSIDAD DISTRITAL
FRANCISCO JOSE DE CALDAS

FACULTAD DE INGENIERÍA
Maestría en Ciencias de la
Información y las Comunicaciones

1. Introduction	X	X													
2. Big Data Architecture			X	X	X	X									
3. Data Analysis Process							X	X	x	X	X				
4.NLP												X	X	X	X	
5. Final Poject																X

Evaluation

ASPECTS TO EVALUATE

It is important to note the differences between evaluating and scoring. The first is a qualitative process and the second a quantitative terminal state. The evaluation criteria must be previously known by the students. To obtain the necessary information for the evaluation processes, it is required to design different specific formats of self-evaluation, coevaluation and hetero-evaluation. Competency training requires: 1. Assessment of student learning in its dimensions: individual / group, theory / practice, written oral. 2. Self-evaluation: the evaluation of student performance carried out by the student. 3. Co-evaluation of student performance between students and teacher. 4. Evaluation of teaching performance.

	TIPO DE EVALUACIÓN	**FECHA**	**PORCENTAJE**
First Grade	Theoretical Evaluation of the concepts of Big Data Architectures and NO-SQL Databases		20%
Second Grade	Theoretical evaluation of data analysis and machine learning		20%
Third Grade	Presentation of projects applied every week		30%
Fourth Grade	Final project		30%

TEACHER INFORMATION:

NAME: Jorge Mario Calvo L.