ACADEMIC PROGRAM

ADVANCED DATABASES

B.F.A. IN **COMPUTER SCIENCE**

MODALITY: ON CAMPUS

ACADEMIC YEAR: 2023-2024

Name of the course:	Advanced Databases
Degree :	Computer Science
Location:	Centro Universitario de Tecnología y Arte Digital
Area:	Data Engineering
Year:	3º
Teaching period:	1
Type:	OBM
ECTS credits:	6
Teaching modality:	On campus
Language:	English
Lecturer / Email	-
Web page:	http://www.u-tad.com/

SUBJECT DESCRIPTION

Area description

The contents of the subject allow students to understand the flow of searching, ingesting, storing, processing and analyzing data information and brings students closer to the techniques and technologies necessary for managing large amounts of data.

Subject description

The subject Database Expansion focuses on the study of non-relational databases and their main currents. Take advantage of this tour to present basic knowledge and techniques that are applicable to all databases and data management systems in general.

To do this, we work from a theoretical and practical point of view with the standard databases of the four main currents of non-relational modeling. Likewise, massive databases (Big Data) and their main management and analysis techniques are studied.

The objective of this subject is for the student to obtain sufficient knowledge to deal with relational and non-relational database management problems from the point of view of a designer and/or programmer, as well as the management and analysis of large volumes of data. information.

COMPETENCIES AND LEARNING OUTCOMES

Competencies

BASIC AND GENERAL SKILLS

CG1 - Ability to understand, schedule and solve problems trough software development

CG3 - Knowledge of the scientific fundamentals applicable to the resolution of computer problems

CG4 - Ability to simplify and optimize computer systems by understanding their complexity

CG9 - Ability to learn, modify and develop new software solutions

CG10 -Use of creative techniques to carry out computer projects

CG11 - Ability to search, analyze and manage information for insights capture

BC1: Students should demonstrate knowledge in an area of study that builds upon the foundation of general secondary education and goest beyond at a level that, while supported by advanced textbooks, also encompasses certain aspects derived from the cutting edge of their field of study.

BC2: Students should be able to apply their knowledge to their work or vocation in a professional manner, and they should possess the competencies typically demonstrated through the development and defence of arguments as well as problem-solving within their field of study.

BC3: Students must possess the ability to gather and interpret relevant data (usually within their field of study) in order to make judgments that involve reflection on socially, scientifically, or ethically significant issues.

BC4: Students should be capable of conveying information, ideas, problems, and solutions to both specialized and non-specialized audiences.

BC5: Students should have developed the learning skills necessary to pursue further studies with a high degree of autonomy.

TRANSVERSAL SKILLS

CT1 - Knowledge of the definition, scope and implementation of the fundamentals of project management methodologies for technology projects

CT2 - Knowledge of the main sectorial players and the life cycle of a digital content development and commercialization project

CT4 -Ability to update the knowledge acquired in the management of digital tools and technologies according to the current state of affairs of the sector and the technological solution

CT5 -Development of the basic skills for digital entrepreneurship.

SPECIFIC SKILLS

CE3 - Knowledge of relational algebra and querying in procedural languages for the design of standardized database schemas based on entity-relationship models

CE10 - Ability to work with a release manager and generate application documentation automatically.

Learning outcomes

Upon completion of the degree, the graduate will be able to:

- To know and develop storage procedures and data management in distributed environments.

- To know and apply supervised, unsupervised and semisupervised learning techniques.

- To know and apply Deep Learning techniques

- To be able to retrieve information using web scraping or standard APIs

- To know and understand Natural Language Processing techniques

- To be able to analyze social networks contents.

- To understand the nature and representation of digital images.

- To know the applications of neural networks to the analysis and generation of sound, static images and video.

- To develop software solutions for computer vision.

- To develop a fully-fledged data project applying iterative methodology from design to delivery.

CONTENTS

Non-relational systems, classification and properties

Integrity restrictions and normalization

Query optimization in distributed environments

Transactions, concurrency control, consistency

Distributed, parallel, heterogeneous and adaptative databases.

SUBJECT SYLLABUS

Topic 1.- Non-relational databases

Introduction to databases and data persistence. Evolution of databases. Standardized model: Relational vs non-relational modeling. Concurrent access and transactions: ACID (atomicity, consistency, isolation and durability). Database integration.

Topic 2.- Document-oriented databases

MongoDB. Document-oriented non-relational modeling. MVC pattern. Design and implementation of models (MVC). JSON and integration with web systems. Concurrency control. Query optimization. Indexed. Iterators.

Topic 3.- Key-value oriented databases

Redis. Key-value oriented non-relational modeling. High-performance in-memory databases. Databases as high-capacity cache systems. Model optimization (MVC) with cache systems. Concurrency and consistency control. Centralized communication systems (publish & subscribe). Iterators.

Topic 4.- Graph-oriented databases

Neo4j. Non-relational graph-oriented modeling. Query optimization. Functional programming in queries. Indexed.

Topic 5.- Big Data: Databases and distributed systems

Hadoop: Storage of unstructured or semi-structured information in distributed systems. Cassandra: Column-oriented databases in distributed systems. Architecture and logic of distributed systems. Consistency and durability in distributed systems. Fundamentals of the query system in distributed systems: MapReduce. Query optimization in distributed systems.

TRAINING ACTIVITIES AND TEACHING METHODOLOGIES

TRAINING ACTIVITIES

LEARNING ACTIVITIES	Total hours	Hours of presence
Theoretical / Expository classes	29,38	29,38
Practical classes	23,25	23,25
Tutorials	4,00	2,00
Independent study and autonomous work of the student	50,00	0,00
Elaboration of work (group or individual)	31,88	0,00
Evaluation Activities	5,25	5,25
Project Follow-Up	6,25	6,25
TOTAL	150	66,13

Teaching methodologies

Expository method or master lesson

Case learning

Learning based on problem solving

Project based learning

Cooperative or collaborative learning

inquiry learning

Flipped classroom methodology

Gamification

Just in time Teaching (JITT) or classroom on time

Expository method or master lesson

Case method

Learning based on problem solving

Project based learning

Cooperative or collaborative learning

inquiry learning

Flipped classroom methodology

Gamification

TEMPORAL DEVELOPMENT

DIDACTIC UNITS / TOPICS TIME PERIOD

Topic 1. Non-relational databases. 1 week

Topic 2. Document-oriented databases. 3 weeks

Topic 3. Key-value oriented databases. 3 weeks

Topic 4. Graph-oriented databases. 3 weeks

Topic 5. Big Data: Databases and distributed systems 3 weeks

EVALUATION SYSTEM

ASSESSMENT SYSTEM	MINIMUM SCORE RESPECT TO THE FINAL ASSESSMENT (%)	MAXIMUM SCORE RESPECT TO THE FINAL ASSESSMENT (%)
Assessment of participation in class, exercises or projects of the course	10	30
Assessment of assignments, projects, reports, memos	40	80
Objective test	10	60

GRADING CRITERIA

ASSESSMENT SYSTEM	ORDINARY EVALUATION	EXTRAORDINARY EVALUATION
Assessment of participation in class, exercises or projects of the course	20	20

Assessment of assignments, projects, reports, memos	60	60
Objective test	20	20

General comments on the evaluations/assessments

For SE1 Knowledge acquisition: Evaluation of the student's effort through assessment of participation, attendance and student work during class. Rated from 0 to 10

For SE2. 4 projects focused on the evaluation of specific knowledge. Evaluated from 0 to 10. It is mandatory to obtain at least a 5 in each of the projects to be evaluated.

Cleanliness, level of detail, precision, demonstration of assimilation of concepts, quality of the projects developed, effort made, defense.

Negative grading of a defense may result in invalidation of the practice for said student.

For SE3 Final test focused on the global evaluation of the material taught in class and the projects developed. Evaluated from 0 to 10. It is mandatory to obtain at least a 5 to be evaluated. This minimum is reduced to a 4 if you attend more than 80% of the classes and have obtained more than a 7.5 in Knowledge Acquisition.

• To pass the subject it is necessary to obtain a weighted average grade, as expressed in the table corresponding to the type of evaluation, equal to or greater than 5 points out of 10.

• In order to be able to take the weighted average, all the projects and the final test must have been passed separately with a grade equal to or greater than 5 points out of 10.

• In the case of obtaining a grade higher than 7.5 in knowledge acquisition and having attended 80% of the classes or more, the minimum required in the final test will be reduced to 4.

• Those students who have not attended 80% of the classes may be classified as failed regardless of the work done in class or the deliveries and tests taken. If attendance is expected to be less than 80%, notify the teacher indicating the reason.

Projects

• There will be two deadlines for the delivery of the projects.

• Normal deadline: Delivery of the project within the indicated deadline. It will be graded on the entire possible grade.

• Extended deadline: Delivery of the project outside the indicated deadline. They will be graded out of 9 out of 10. For each day of delay in delivery, the grade will be reduced by 0.2 points. (It is necessary to obtain a grade equal to or greater than 5 to pass the project.)

• The last practice will have a limited extended delivery time. There is no extended delivery period in extraordinary calls.

Copy

• Any suspicion of copying between two or more practices or the use of code obtained on the Internet will invalidate the delivery and will lead to the assignment of a fail in said call and the following one to all the students involved.

• The use of virtual assistants such as ChatGPT or Copilot is not prohibited. However, the student must be able to demonstrate the knowledge acquired when required by the teacher.

Extraordinary Call

• Those parts passed in the ordinary evaluation will be considered passed in the extraordinary one.

• The parts not passed must be repeated in an extraordinary evaluation.

• The evaluation percentages will be the same as in the ordinary call.

Other considerations

• No grades of any kind will be kept between different academic years, nor between different calls.

• The use of mobile phones in the classroom is not allowed during the continuous evaluation period, unless expressly indicated otherwise by the teacher. Laptops may only be used for activities related to the subject. The teacher may withdraw the right to use the computer from those students who use it for activities that are not related to the subject (checking emails, news or social networks, consulting or preparing activities for other subjects, etc.).

• It is not allowed to consume drinks or food in the classroom. The presence of any type of drink on the tables is also not permitted, even in closed containers.

• The student will be required to participate actively, necessary for the development of the classes.

• The student will be required to behave well at all times during classes. Bad behavior that prevents the normal development of the class may lead to expulsion from the classroom for a period of time to be determined by the teacher.

LIST OF REFERENCES (BOOKS, PUBLICATIONS, WEBSITES):

asic Bibliography:

• Date, C.J. (2004). An Introduction to Database Systems. Pearson Education India.

• Sadalage, P. J., & Fowler, M. (2013). NoSQL distilled: a brief guide to the emerging world of polyglot persistence. Pearson Education.

• MongoDB Manual, MongoDB Inc., accessed 9/10/2019, https://docs.mongodb.com/manual/

• Redis documentation, Redis Labs, accessed 9/10/2019, https://redis.io/documentation

• Neo4j Documentation, Neo4j Inc., accessed 9/10/2019, https://neo4j.com/docs/

• Apache Cassandra Documentation, Apache Software Foundation, accessed 9/10/2019, https://cassandra.apache.org/doc/latest/

• Apache Hadoop, Apache Software Foundation, accessed 9/10/2019, http://hadoop.apache.org/docs/current/

Recommended Bibliography:

• Codd, E.F. (1990). The relational model for database management: version 2. Addison-Wesley Longman Publishing Co., Inc..

• Han, J., Haihong, E., Le, G., & Du, J. (2011). Survey on NoSQL database. In 2011 6th international conference on pervasive computing and applications (pp. 363-366). IEEE.

• Cattell, R. (2011). Scalable SQL and NoSQL data stores. Acm Sigmod Record, 39(4), 12-27.

• Moniruzzaman, A. B. M., & Hossain, S. A. (2013). Nosql database: New era of databases for big data analytics-classification, characteristics and comparison. arXiv preprint arXiv:1307.0191.

• Leavitt, N. (2010). Will NoSQL databases live up to their promise? Computer, 43(2), 12-14.

• Abadi, D. (2012). Consistency tradeoffs in modern distributed database system design: CAP is only part of the story. Computer, 45(2), 37-42.

• DataStax Documentation, DataStax Inc, accessed 9/10/2019, https://docs.datastax.com/

• List of NoSQL Databases, accessed 9/10/2019, http://nosql-database.org/

• DB-Engines Ranking, solid IT gmbh, accessed 9/10/2019, https://db-engines.com/en/ranking

• Python Documentation, Python Software Foundation, accessed 9/10/2019, https://www.python.org/doc/

REQUIRED MATERIALS, SOFTWARE AND TOOLS

Type of classroom
Theory classroom

Board and projection system

Materials:
Personal computer with Unix-based OS

o

Personal computer with Windows and virtual machine with Unix-based OS

Software:
SOFTWARE:

Lenguaje de programación

- Python 3.X

- Conda o pip para instalar librerías Bases de datos

- Mongo DB- Redis

- Neo4j

Entorno de trabajo.

Cualquier entorno de programación como:

- Visual Studio Code

- Atom

- PyCharm

- Etc.