

EMERGING TOPICS IN DATA ANALYSIS & MANAGEMENT

Dual Degree in Business Administration & Data and Business Analytics BBADBA SEP-2024 ETDAM-NBDA.5.M.A

Area Data Science

Number of sessions: 30

Academic year: 24-25

Degree course: FIFTH

Number of credits: 6.0

Semester: 2º

Category: COMPULSORY

Language: English

Professor: **ALEJANDRO VACA SERRANO**

E-mail: avaca@faculty.ie.edu

CEO & Founder @ LenguajeNatural.AI

CEO and Founder - LenguajeNatural.AI (since Feb. 2024)

Data Scientist at Instituto de Ingeniería del Conocimiento (IIC) from 2019-2024

Deep Learning, Computer Vision & Natural Language Processor in different institutions since 2021.

? Google Scholar profile: [shorturl.at/lnp58](https://scholar.google.com/citations?user=shorturl.at/lnp58)

??Best Poster Presentation Award at the NAACL 2022 LatinxAI Workshop.

??1st prize at the SomosNLP 2022 Hackathon for the BioMedIA project: <https://huggingface.co/spaces/hackathon-pln-es/BioMedIA>.

??Special award for project with the highest number of Likes for BioMedIA - Hackaton SomosNLP 2022.

??1st prize in Tasks 1 & 2 (2/2) at Exist22@IberleF22, a challenge focused on detecting sexism in texts in English and Spanish.

???Main developer of RigoBERTa, the most capable Spanish language model, reaching a new State of the Art (best model in 10/13 tasks). Link to the paper: [shorturl.at/ghjIW](https://arxiv.org/abs/2010.05958)

??Best Data Scientist Special Award - Hackaton SpainAI 2021

??1st prize @ SpainAI 2021 Hackathon Time Series challenge.

??1st prize @ Computer Vision challenge of the SpainAI Hackathon 2021.

??3rd prize @ NLP challenge of the SpainAI Hackathon 2021.

??1st Prize @ Minsait Land Classification challenge - Cajamar UniversityHack2020.

Office Hours

Office hours will be on request. Please contact at:

LinkedIn: <https://www.linkedin.com/in/alejandro-vaca-serrano/>

e-mail: alejandro_vaca0@hotmail.com

SUBJECT DESCRIPTION

This course is designed to explore the evolving landscape of AI & Data Analysis and its impact on organizations. We will delve into three key areas that are shaping the future of intelligent organizations:

- Firstly, we will focus on understanding Data Analysis, providing a foundational understanding of the field, covering essential concepts, methodologies, and tools. Students will explore the role of data in decision-making, learn fundamental statistical techniques, and gain hands-on experience with popular data analysis tools.
- Secondly, we will explore emerging trends in data analytics, embarking on a journey to discover the latest advancements in the field. From machine learning and artificial intelligence to interactive data visualization, students will explore cutting-edge techniques and technologies shaping the future of organizations. Practical applications and case studies will be used to showcase how these emerging trends are being employed in real-world scenarios.
- Finally, we will cover the matter of data-driven decision-making in business contexts. Students will examine specific industries and business areas where data analytics plays a pivotal role in driving informed decision-making and operational efficiencies. Case studies from various sectors will provide insights into how organizations leverage data to gain a competitive advantage, optimize processes, and innovate.

LEARNING OBJECTIVES

- Develop a foundational understanding of data analysis principles and techniques.
- Explore and critically evaluate emerging trends and technologies in the field of data analytics and AI.
- Analyze real-world applications of data analytics and AI in different business domains.
- Gain hands-on experience with relevant tools and technologies used in data analysis and AI.
- Enhance critical thinking and problem-solving skills through practical exercises and case studies.
- Foster an awareness of ethical considerations and challenges in the use of data and intelligent systems.

TEACHING METHODOLOGY

IE University teaching method is defined by its collaborative, active, and applied nature. Students actively participate in the whole process to build their knowledge and sharpen their skills. Professor's main role is to lead and guide students to achieve the learning objectives of the course. This is done by engaging in a diverse range of teaching techniques and different types of learning activities such as the following:

- 30 live synchronous face-to-face sessions in a classroom.
- Each session will consist of synchronous and will have a theoretical and practical component. It

will combine live lectures, discussions, case studies and a final exam.

- Most of the sessions will have a focus on real applications of the theory we are covering.
- Bringing your laptop is mandatory to all sessions, although its use (or not) will be decided by the professor. There will be a brief reading list included also some articles and recommended media content which links will be provided to access on class notes.

Learning Activity	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	26.7 %	40.0 hours
Discussions	6.7 %	10.0 hours
Exercises in class, Asynchronous sessions, Field Work	6.7 %	10.0 hours
Group work	40.0 %	60.0 hours
Individual studying	20.0 %	30.0 hours
TOTAL	100.0 %	150.0 hours

AI POLICY

In this course, **the use of generative artificial intelligence (GenAI) is encouraged**, with the goal of developing an informed critical perspective on potential uses and generated outputs.

However, be aware of the limits of GenAI in its current state of development:

- If you provide minimum effort prompts, you will get low quality results. You will need to refine your prompts to get good outcomes. This will take work.
- Don't take ChatGPT's or any GenAI's output at face value. Assume it is wrong unless you either know the answer or can cross-check it with another source. You are responsible for any errors or omissions. You will be able to validate the outputs of GenAI for topics you understand.
- AI is a tool, but one that you need to acknowledge using. Failure to do so is in violation of academic honesty policies. Acknowledging the use of AI will not impact your grade.

Suggested format to acknowledge the use of generative AI tools:

I acknowledge the use of [AI systems link] to [specify how you used generative AI]. The prompts used include [list of prompts]. The output of these prompts was used to [explain how you used the outputs in your work].

If you have chosen not to include any AI generated content in your assignment, the following disclosure is recommended:

No content generated by AI technologies has been used in this assignment.

PROGRAM

DISCLAIMER

The following description of the material covered is tentative. An attempt will be made to cover all listed topics. However; the pace in the classes will depend on the group performance.

SESSION 1 (LIVE IN-PERSON)

Data Analysis Foundations I: fundamentals of data understanding. How knowledge is built from raw data. Differences between data, information, knowledge, and wisdom. Class examples illustrating the role of data in decision making. Review of emerging technologies cycle in the domain of data analytics and AI.

SESSION 2 (LIVE IN-PERSON)

Data Analysis Foundations II: sources, management and types of data. Current trends in processing of structured and unstructured data. Advances in data storage, retrieval, querying and visualization. Review of relevant data platforms for big data and related analytics services.

SESSION 3 (LIVE IN-PERSON)

Data Analysis Foundations III: maturity model in data analysis and life cycle of data analytics projects. Types of data analysis, from descriptive models to generative AI. Illustrative use cases of data analysis applications at various levels of maturity.

SESSION 4 (LIVE IN-PERSON)

Data Analysis Foundations IV: how a data analytics dream team looks like. Current organizational frameworks for the success of AI and data analysis projects. Building a career in AI and Data Science. Required skills in the data science team. The role of the Data Translator in bridging the gap between business requirements and technical capabilities.

SESSION 5 (LIVE IN-PERSON)

Data Analysis Tools and Methods I: trends in data analytics platforms, from Business Intelligence and Data Lakes to Edge AI. Types of platforms and tools. No-Code, Low-Code, AI Copilots and Data Science frameworks. Cloud based platforms. AI as a Service. Review of relevant platforms.

SESSION 6 (LIVE IN-PERSON)

Data Analysis Tools and Methods II: Trends in low-code and AI copilots. Hands-on practice with a selected modern low-code/no-code tool to create a data pipeline.

SESSION 7 (LIVE IN-PERSON)

Data Analysis Tools and Methods III: Trends in code intensive frameworks. Hands-on practice with a state-of-the-art code-based tool.

SESSION 8 (LIVE IN-PERSON)

Data Analysis Tools and Methods IV: Trends in cloud-based APIs. Hands-on practice with modern AI API services.

SESSION 9 (LIVE IN-PERSON)

Introduction to the group work assignment.

Current advances in Data Science I: Trends in Machine Learning and Deep Learning. From automated feature engineering to MLOps. Hands-on practice.

SESSION 10 (LIVE IN-PERSON)

Current advances in Data Science II: Trends in Machine Learning and Deep Learning. From latent spaces to LLMs. Hands-on practice.

SESSION 11 (LIVE IN-PERSON)

Generative AI I: Trends in foundational models LLMs and MMMs. Prompt Engineering. Prompting strategies. Hands-on practice.

SESSION 12 (LIVE IN-PERSON)

Generative AI II: Advanced Prompt Engineering. Meta prompting and few-shot learning strategies. Hands-on practice.

SESSION 13 (LIVE IN-PERSON)

Generative AI III: Building LLM-based apps, unstructured data driven application building. Hands-on practice.

SESSION 14 (LIVE IN-PERSON)

Generative AI IV: Building LLM-based intelligent agents. Agent frameworks and libraries. Hands-on practice.

SESSION 15 (LIVE IN-PERSON)

Review of group work proposals.

Advanced Data Retrieval: When data is not in a relational database. From vector stores to intelligent document ingestion. Hands-on practice.

SESSION 16 (LIVE IN-PERSON)

Retrieval Augmented Generation I: Empowering LLMs with real-time business data. Understanding the RAG pattern. Hands-on practice.

SESSION 17 (LIVE IN-PERSON)

Retrieval Augmented Generation II: Advanced RAG methods. Metadata and hybrid search. RAG as a tool for agents. Hands-on practice.

SESSION 18 (LIVE IN-PERSON)

Reasoning over data structures: Multi-stage and Multi-agent prompting. Chain of Thought reasoning for data intensive intelligent applications. Knowledge Graphs. Hands on practice.

SESSION 19 (LIVE IN-PERSON)

Affective Computing I: from usability to user affective experience. Using emotion data to make products and services appealing.

SESSION 20 (LIVE IN-PERSON)

Affective Computing II: user personality versus artificial personality. Using personality data to make products and services appealing. Building an artificial personality for AI assistants.

SESSION 21 (LIVE IN-PERSON)

Building data applications I: Review of reference architectures and best practices for building scalable and robust data applications.

SESSION 22 (LIVE IN-PERSON)

Building data applications II: Backend architecture design for building scalable and robust data applications. Data services “APIfication”. Hands on practice.

SESSION 23 (LIVE IN-PERSON)

Building data applications III: Frontend architecture design for building scalable and robust data applications. Serving predictive and generative models. Taking MLOps/AIOps seriously in the application design. Building the user experience. Hands on practice.

SESSION 24 (LIVE IN-PERSON)

Group Work Presentations by student teams.

SESSION 25 (LIVE IN-PERSON)

Data Privacy and Security: Understanding data privacy. Identifying data risks. Specific risks in emerging data intensive technologies. Compliance and regulations. Privacy by design. How data protection works. Anonymization. Tackling vulnerabilities and risks in Generative AI applications. Ethical hacking practice.

SESSION 26 (LIVE IN-PERSON)

Data Ethics: How to build ethical data products and services. AI Ethics. Data-driven decision making versus model human alignment. How AI models are tamed to behave appropriately. Debate: AI and Data Analysis in conflicts, wars, intelligence services and military applications.

SESSION 27 (LIVE IN-PERSON)

Data & AI Use Cases and Strategic Planning I: Use case discovery, ideation and prioritization. Use Case Ideation workshop.

SESSION 28 (LIVE IN-PERSON)

Data & AI Use Cases and Strategic Planning II: Review of use cases in healthcare, marketing, banking, insurance, energy, utilities, defense, public sector, retail, logistics, and transportation. Trends review. Successfully integrating advanced data analytics initiatives in the organization. The role of the CDO/CAIO/CDTO/CDSO.

SESSION 29 (LIVE IN-PERSON)

Exam Preparation Questions.

Near Future Technologies in Data & AI. Review and predictions of upcoming trends in data analysis and management.

SESSION 30 (LIVE IN-PERSON)

Final Exam.

EVALUATION CRITERIA

Your final grade in the course will be based on class participation, an exam and a group work. The weight of each one will be as follows:

criteria	percentage	Learning Objectives	Comments
Final Exam	40 %		
Group Presentation	10 %		
Group Work	40 %		
Class Participation	10 %		

RE-SIT / RE-TAKE POLICY

Each student has four chances to pass any given course distributed over two consecutive academic years: ordinary call exams and extraordinary call exams (re-sits) in June/July.

Students who do not comply with the 80% attendance rule during the semester will fail both calls for this Academic Year (ordinary and extraordinary) and have to re-take the course (i.e., re-enroll) in the next Academic Year.

Evaluation criteria:

- Students failing the course in the ordinary call (during the semester) will have to re-sit the exam in June / July (except those not complying with the attendance rule, who will not have that opportunity and must directly re-enroll in the course on the next Academic Year).
- The extraordinary call exams in June / July (re-sits) require your physical presence at the campus you are enrolled in (Segovia or Madrid). There is no possibility to change the date, location or format of any exam, under any circumstances. Dates and location of the June / July re-sit exams will be posted in advance. Please take this into consideration when planning your summer.
- The June / July re-sit exam will consist of a comprehensive exam. Your final grade for the course will depend on the performance in this exam only; continuous evaluation over the semester will not be taken into consideration. Students will have to achieve the minimum passing grade of 5 and can obtain a maximum grade of 8.0 (out of 10.0) - i.e., "notable" in the re-sit exam.
- Retakers: Students who failed the subject on a previous Academic Year and are now re-enrolled as re-takers in a course will be needed to check the syllabus of the assigned professor, as well as contact the professor individually, regarding the specific evaluation criteria for them as retakers in the course during that semester (ordinary call of that Academic Year). The maximum grade that may be obtained in the retake exam (3rd call) is 10.0.

After ordinary and extraordinary call exams are graded by the professor, you will have a possibility to attend a review session for that exam and course grade. Please be available to attend the session in order to clarify any concerns you might have regarding your exam. Your professor will inform you about the time and place of the review session. Any grade appeals require that the student attended the review session prior to appealing.

- Students failing more than 18 ECTS credits in the academic year after the June-July re-sits will be asked to leave the Program. Please, make sure to prepare yourself well for the exams in order to pass your failed subjects.
- In case you decide to skip the opportunity to re-sit for an exam during the June / July extraordinary call, you will need to enroll in that course again for the next Academic Year as a re-taker and pay the corresponding extra cost. As you know, students have a total of four allowed calls to pass a given subject or course, in order to remain in the program.

BIBLIOGRAPHY

Recommended

- Stuart Russell, Peter Norvig. *Artificial Intelligence: A modern approach*. ISBN 0136042597 (Digital)
- Denis Rothman. *Transformers for Natural Language Processing: Build, train, and fine-tune deep neural network*. ISBN 978180324733 (Digital)
- Aurelien Geron. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques*. ISBN 978149203264 (Digital)
- Martin Kleppmann. *Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable*. ISBN 978144937332 (Digital)

BEHAVIOR RULES

Please, check the University's Code of Conduct [here](#). The Program Director may provide further indications.

ATTENDANCE POLICY

Please, check the University's Attendance Policy [here](#). The Program Director may provide further indications.

ETHICAL POLICY

Please, check the University's Ethics Code [here](#). The Program Director may provide further indications.