DESIGNING ARTIFICIAL INTELLIGENCE & IMPLEMENTING SMART TECHNOLOGIES

Bachelor in Data and Business Analytics BDBA SEP-2024
DAI-DBA.3.M.A

Area Computer Science and AI
Number of sessions: 30
Academic year: 24-25
Degree course: THIRD
Number of credits: 6.0
Semester: 2º
Category: COMPULSORY
Language: English

Professor: **JUAN JOSE MANJARIN COLON**

E-mail: jjmanjarin@faculty.ie.edu

Juanjo Manjarín received the highest grade for his Ph.D. in Theoretical Physics in String Theory and M-Theory from the Universdad Autónoma de Madrid (UAM), after earning an Advanced Studies Diploma (DEA) in Theoretical Physics from the Universdad Autónoma de Madrid (UAM). His teaching experience includes different international universities where he lectured on Complex Variable Analysis, Theoretical Mechanics or Classical Electrodynamics and in the IE University where he has taught Mathematics, Statistics, Econometrics, Social Media Analytics and Programming in R and Python. He also teaches the course of "Math and Stats for Machine Learning" in the Bootcamp for Data Science in the IE Exponential Learning.

He has published a number of papers on international journals on mathematics and theoretical physics and he was reviewer for Mathematical Reviews from 2003 to 2005. His research interests are Quantum Information and Computing and Network Science together with their applications in Data Science.

He has also corporate experience on different TV and cinema production companies: Gestmusic Endemol, 7 y Acción S.L., Hill Valley S.L., 100 Balas S.L. or Zebra Producciones in TV shows such as “El Hormiguero” or “Esto es vida!”, receiving prizes Ondas and Rose d'Or in 2008 and 2009. He was director, producer and post-producer of different short films and now as director of E8 Producciones is recording a documentary film. He also worked in El Pais in the realization of some divulgative science materials.

Office Hours

Office hours will be on request. Please contact at:

jjmanjarin@faculty.ie.edu

SUBJECT DESCRIPTION

Artificial intelligence and smart technologies are one of the most common techniques applied to analyze and get business value from data. It is important to understand how Artificial Intelligence systems can be developed, why their goals are and which underlying techniques are more suitable.

This course will cover the following topics: (i) **AI Design**, (ii) **Natural Language Processing (NLP)**, (iii) **Computer vision**, (iv) **Reinforcement learning** and (v) **Machine Learning Operations (MLOps)**.

In AI design we will focus on the requirements and trade-offs when building AI systems in practical business contexts and how to maximize the investments.

The NLP section will cover the recent advances applying transformers to text processing use cases like text classification and summarization, question-answering and text generation.

Then, we will cover Computer vision use cases and provide an understanding of the underlying Deep learning technologies that are making possible achieving high levels of accuracy in novel use cases like Image classification and text to image.

We will follow with an understanding and study of reinforcement learning and how this technique has been successfully applied to achieve huge improvements in recent AI advances.

Finally, we will cover Machine Learning Operations also known as MLOps. MLOps consists of the set of techniques and best practices to not only deploy and operationalize machine learning models to drive business outcomes but also monitor, manage and govern those models over time.

Throughout the course students will have the opportunity to design and implement an end-to-end use case using the technologies covered.

LEARNING OBJECTIVES

The course will cover the following learning objectives:

- Understand the context and business value of Artificial Intelligence (AI) technologies
- Explore how AI technologies are disrupting traditional business models
- Acquire knowledge of the technologies and tools used to build AI systems
- Use Python and PyTorch to develop AI systems, with a focus on natural language processing, computer vision, and reinforcement learning
- Implement an end-to-end AI use case and understand the concepts and requirements for Machine Learning Operations (MLOps).

By the end of the course, participants will have a strong foundation in AI technologies and their practical applications in business

TEACHING METHODOLOGY

Students will have the opportunity to apply what they have learned through practical exercises using Python notebooks and popular deep learning frameworks such as PyTorch. This will provide a deeper understanding of the concepts and their real-world applications

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 indicated later.

Exams:
There will just one formal test/exam. However, as an especial practice/assignment the students will have a series of test-exams to get them prepared for answering questions about ML and RL (multiple- choice, true-false or open questions). These test-exams will be either automatically scored by the machine or in some cases will require a self-evaluation or self-assessment. So, they are a type of simulation of an open-book exam on all material covered up to that point in the lectures, tutorials, required readings, and assignments, but eventually there could be topics not covered in any of these that will require some online investigation or other type of research to find out (or understand) the possible answers. The final scoring of this "assignment" will be based on the number of test- exams done (freely decided by each student), their average results, and the degree of progressive improvement/learning shown.

[Most of] The questions for the final test/exam will be selected from the ones used for the practice

Learning Activity	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	30.0 %	45.0 hours
Discussions	10.0 %	15.0 hours
Exercises in class, Asynchronous sessions, Field Work	33.3 %	50.0 hours
Group work	13.3 %	20.0 hours
Individual studying	13.3 %	20.0 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.

PROGRAM

SESSION 1 (LIVE IN-PERSON)

AI Design: Introduction to the course and reviewing AI Concepts

In this session, we will provide an overview of the course structure and learning objectives. We will also review the key AI concepts that will be covered in the course, including machine learning, natural language processing, computer vision, and reinforcement learning.

This session will serve as a foundation for the rest of the course, and it is important that participants have a good understanding of these concepts before moving on to more advanced topics. We encourage active participation and questions throughout the session, as this will help to clarify any areas of confusion and ensure that everyone has a strong foundation in AI design.

SESSION 2 (LIVE IN-PERSON)

AI Design: Challenges designing and developing AI Systems

In this session, we will explore the challenges that arise when designing and developing AI systems, with a focus on neural networks and deep learning models.

Some of the topics we will briefly discuss include

- Data preparation and feature engineering
- Model selection and optimization
- Overfitting and regularization
- Hyperparameter tuning
- Model evaluation and interpretation

We will also discuss strategies for addressing these challenges, such as using cross-validation and ensembles, and we will explore the trade-offs involved in different design choices. This session will provide participants with a deeper understanding of the technical considerations involved in building effective AI systems.

SESSION 3 (LIVE IN-PERSON)

AI Design: Challenges deploying AI Systems and getting business value

In this session, we will delve into the common challenges and obstacles that organizations face when deploying AI systems and realizing business value from them.

To begin, we will discuss the importance of aligning AI projects with business goals and objectives, and we will explore the various stakeholders that are typically involved in successful AI projects. We will also discuss the different roles and responsibilities that team members may have at different stages of the project, from ideation to deployment and ongoing maintenance.

Next, we will delve into the specific challenges that organizations may encounter when deploying AI systems, including issues related to data quality and quantity, model performance and reliability, and integration with existing systems and processes. We will also discuss strategies for overcoming these challenges, including best practices for project management, testing and validation, and stakeholder communication.

By the end of this session, participants should have a better understanding of the various challenges and considerations involved in deploying AI systems and realizing business value from them.

SESSION 4 (LIVE IN-PERSON)

Python Review and Pytorch introduction

In this section, we will review the essential Python concepts that are required for working with PyTorch, including setting up development environments, importing and using libraries, and working with data structures and control structures.

We will also introduce PyTorch, a powerful deep learning framework that is built on top of Python.

By the end of this session, participants should have a good understanding of Python and an introduction to PyTorch, and they should be ready to start using these tools to build machine learning and deep learning models.

SESSION 5 (LIVE IN-PERSON)

Understanding Pytorch

In this session, we will delve deeper into the PyTorch deep learning framework, exploring its key features and benefits in more detail.

We will discuss how to use PyTorch to define, train, and evaluate deep learning models. We will also cover advanced PyTorch concepts such as custom layers and loss functions, and we will demonstrate how to use PyTorch's built-in datasets and data loaders to efficiently load and process data.

Additionally, we will discuss the flexibility and efficiency of PyTorch, including its support for distributed training and GPU acceleration, and we will explore how to use PyTorch to build machine learning models.

By the end of this session, participants should have a good understanding of how to use PyTorch to build deep learning models.

SESSION 6 (LIVE IN-PERSON)

NLP: Introduction to Natural Language Processing with Transformers

In this session, we will explore the foundations of natural language processing (NLP) and introduce the transformers, a powerful deep learning architecture for solving a wide range of NLP tasks.

We will start by reviewing the basics of NLP, including common techniques for preprocessing and representing text data, and we will discuss the challenges of working with natural language data. We will then introduce the transformers architecture, which has achieved state-of-the-art results on a wide range of NLP benchmarks, and we will demonstrate how to use transformers to build and train NLP models using PyTorch.

By the end of this session, participants should have a good understanding of NLP and the transformers architecture, and they may be able to apply these concepts to build and deploy NLP models using PyTorch.

SESSION 7 (LIVE IN-PERSON)

NLP: Text Classification

In this session, we will dive deeper into the topic of text classification, which is a common task in natural language processing (NLP). Text classification involves assigning labels or categories to text data, and it can be applied to a variety of use cases such as tagging customer feedback, routing support tickets, and performing sentiment analysis.

By the end of this session, participants should have a good understanding of text classification and how to apply it to a variety of NLP tasks.

SESSION 8 (LIVE IN-PERSON)

NLP: Text generation and summarization

In this session, we will delve into the exciting topics of text generation and summarization, which are two important natural language processing (NLP) use cases.

Text generation involves creating new text that is coherent and meaningful, and it has a wide range of applications such as content creation, chatbots, and machine translation.

Summarization, on the other hand, involves extracting the most important information from a text and condensing it into a shorter form, and it has numerous applications such as news summarization, document summarization, and social media analytics.

By the end of this session, participants should have a good understanding of text generation and summarization and how to apply these techniques to solve real-world NLP problems using PyTorch.

SESSION 9 (LIVE IN-PERSON)

NLP: Question-Answering

In this session, we will explore the question answering problem, which involves generating a response to a question based on a given text or a set of texts.

Question answering has a wide range of applications, including chatbots, customer service, and information retrieval systems.

SESSION 10 (LIVE IN-PERSON)

NLP: Advanced topics with Transformers

In this session, we will delve into advanced topics in natural language processing (NLP) using transformers, a powerful deep learning architecture for solving various NLP tasks.

We will start by discussing how to train transformers from scratch, including how to preprocess the data, how to define the model architecture, and how to optimize the model using various techniques such as data augmentation and regularization. We will also demonstrate how to train transformers using PyTorch.

Next, we will cover how to deal with few or no labels in your dataset, which is a common challenge in NLP. We will introduce the concept of self-supervised learning.

Finally, we will discuss the current state of the art in NLP and the future directions of this field.

By the end of this session, participants should have a good understanding of advanced topics in NLP using transformers and how to apply these techniques to solve real-world NLP problems.

SESSION 11 (LIVE IN-PERSON)

Hands-On Working Session: Individual Assignment

We will introduce and work with the individual assignment work in this session.

SESSION 12 (LIVE IN-PERSON)

Computer Vision: How Deep Learning and Convolutional Neural Networks improved the state of the art

In this session, we will explore the significant impact that deep learning and convolutional neural networks (CNNs) have had on the field of computer vision.

We will start by discussing how deep learning and CNNs revolutionized the way we solve computer vision tasks such as image classification, object detection, and segmentation. We will cover the key concepts and techniques of deep learning and CNNs, including convolution, pooling, and activation functions, and we will demonstrate how to implement and train CNNs using PyTorch.

We will also introduce the concept of generative adversarial networks (GANs), which are a type of deep learning model that can generate synthetic images that are indistinguishable from real images. We will discuss the applications of GANs in computer vision.

By the end of this session, participants should have a good understanding of how deep learning and CNNs have significantly improved the state of the art in computer vision.

SESSION 13 (LIVE IN-PERSON)

Computer Vision: Using Pytorch in Computer Vision

This section will introduce Pytorch examples applied to Computer Vision. By the end of this session, participants should have a good understanding of how to use PyTorch to solve computer vision tasks.

SESSION 14 (LIVE IN-PERSON)

Computer Vision: Object Detection examples

In this session, we will focus on object detection, which is a computer vision task that involves identifying and localizing objects in images from defined classes.

We will provide a variety of examples and hands-on exercises to help participants understand how to apply object detection techniques to solve real-world problems. By the end of this session, participants should have a good understanding of object detection and how to apply it to solve practical problems.

SESSION 15 (LIVE IN-PERSON)

Computer Vision: Image Classification examples

In this session, we will focus on image classification, which is a computer vision task that involves assigning a label or class to a given image.

We will provide a variety of examples and hands-on exercises to help participants understand how to apply image classification techniques to solve real-world problems. By the end of this session, participants should have a good understanding of image classification and how to apply it to solve practical problems.

SESSION 16 (LIVE IN-PERSON)

Computer Vision: Image Generation examples (Text to image)

In this session, we will focus on image generation, which is a computer vision task that involves generating new images from a user input, such as a text description.

We will provide a variety of examples and hands-on exercises to help participants understand how to apply image generation techniques to solve real-world problems. By the end of this session, participants should have a good understanding of image generation.

SESSION 17 (LIVE IN-PERSON)

Computer Vision: Advanced Topics

We will discuss advances and new topics about Computer Vision with Deep Learning.

SESSION 18 (LIVE IN-PERSON)

Hands-On Working Session

This will be a working session to solve any questions related to the individual and Group assignment.

SESSION 19 (LIVE IN-PERSON)

RL: Introduction to Reinforcement Learning

In this session, we will introduce the basic concepts of reinforcement learning (RL), which is a type of machine learning that involves training agents to make decisions in dynamic environments in order to maximize a reward signal.

We will start by discussing the RL framework and how it differs from other types of machine learning. We will cover the key components of RL, including agents, environments, actions, and rewards, and we will discuss how RL algorithms work and what types of problems they are suitable for

By the end of this session, participants should have a good understanding of the basics of RL and how it can be applied to solve practical problems.

SESSION 20 (LIVE IN-PERSON)

RL: Fundamentals of Reinforcement Learning

In this session, we will delve deeper into the anatomy of reinforcement learning (RL) problems and the scenarios where RL can be applied.

We will start by reviewing the key components of RL, including agents, environments, actions, and rewards, and we will discuss how these components interact in an RL system. We will also introduce the RL loop, which is the fundamental structure of an RL algorithm, and we will discuss how RL algorithms learn from experience and adapt to changing environments.

We will also introduce some of the key challenges and considerations in designing and implementing RL systems, such as exploration-exploitation trade-offs, temporal credit assignment, and generalization across tasks and environments.

By the end of this session, participants should have a good understanding of the fundamentals of RL and how to apply RL to solve practical problems.

SESSION 21 (LIVE IN-PERSON)

RL: Policy Optimization in Reinforcement Learning and Deep Q-Learning algorithm

In this session, we will cover the algorithms behind policy optimization in reinforcement learning (RL) and the deep Q-learning algorithm, which is a popular and widely-used RL algorithm.

We will start by discussing the concept of policy optimization in RL, which is the process of finding the optimal policy that maximizes the expected reward for an RL agent. We will introduce the different types of policy optimization algorithms, including value-based and policy-based algorithms, and we will discuss the pros and cons of each type of algorithm.

Next, we will focus on the deep Q-learning algorithm, which is a value-based RL algorithm that uses a deep neural network to approximate the action-value function. We will discuss how the deep Q-learning algorithm works and how it can be used to solve a variety of RL problems. We will also cover some of the key challenges and considerations in implementing the deep Q-learning algorithm, such as stability, convergence, and exploration-exploitation trade-offs.

By the end of this session, participants should have a good understanding of the algorithms behind policy optimization in RL and how to apply the deep Q-learning algorithm to solve practical problems.

SESSION 22 (LIVE IN-PERSON)

RL: Agent-based modeling and OpenAI Gym introduction

In this session, we will dive into practical examples of reinforcement learning (RL) with Python and the OpenAI Gym framework.

We will introduce the OpenAI Gym framework, which is a toolkit for developing and comparing RL algorithms. We will discuss the key features and benefits of the Gym framework, including its wide range of environments, its support for multiple programming languages, and its flexibility in terms of problem setup and evaluation. We will also demonstrate how to use the Gym framework to develop and evaluate RL algorithms in Python.

By the end of this session, participants should have a basic understanding of agent-based modeling and the OpenAI Gym framework, and how to use these tools to solve practical RL problems.

SESSION 23 (LIVE IN-PERSON)

RL: Practical examples with OpenAI Gym

In this session, we will continue exploring practical examples of reinforcement learning (RL) with the OpenAI Gym framework.

Throughout the session, we will work through a series of hands-on exercises and examples that showcase the power and versatility of the Gym framework. We will cover topics such as policy optimization, value estimation, exploration-exploitation trade-offs, and how to design and implement RL algorithms for different types of problems.

By the end of this session, participants should have a good understanding of how to use the OpenAI Gym framework to solve practical RL problems, and how to apply the concepts and techniques covered in this course to real-world situations.

SESSION 24 (LIVE IN-PERSON)

Group Assignment Working Session

In this section we willl work with the Group assignment and discuss relevant questions.

SESSION 25 (LIVE IN-PERSON)

MLOps: Introduction of MLOps

In this session, we will introduce the concepts of Machine Learning Operations (MLOps) and discuss why MLOps is an important topic in the field of data science and machine learning.

We will start by defining MLOps and discussing the key challenges and considerations involved in operationalizing ML models and pipelines. We will cover topics such as model deploying, monitoring, and governance, and we will discuss the importance of these activities in ensuring the reliability, scalability, and maintainability of ML systems.

Next, we will introduce the concept of model deploying and discuss the different approaches and technologies that can be used to deploy ML models in production environments. We will also cover the importance of monitoring ML models in production and the different metrics and tools that can be used to track the performance and accuracy of ML systems.

By the end of this session, participants should have a good understanding of the MLOps landscape and the key considerations and best practices involved in operationalizing ML models and pipelines.

SESSION 26 (LIVE IN-PERSON)

MLOps: Advanced concepts

In this session, we will continue exploring the concepts of MLOp and cover advanced topics such as model management and governance.

We will start by discussing the importance of model management in MLOps, including version control, documentation, testing, and validation. We will review the different tools and techniques that can be used to manage ML models and pipelines, and we will demonstrate how to implement best practices such as continuous integration, delivery, and deployment (CI/CD).

Next, we will introduce the concept of ML governance and discuss the challenges and best practices involved in managing and governing ML systems in production. We will cover topics such as risk assessment, compliance, ethics, and transparency, and we will discuss the role of ML governance in ensuring the responsible and ethical use of ML.

By the end of this session, participants should have a good understanding of advanced MLOps concepts such as model management and governance, and how to apply these concepts and techniques to real-world ML projects.

SESSION 27 (LIVE IN-PERSON)

MLOps: Technologies and examples

In this session, we will review the available technologies and solutions for MLOps and provide examples of how to use these technologies in practice using open source products like weights and bias.

By the end of this session, participants should have a good understanding of the available technologies and solutions for MLOps.

SESSION 28 (LIVE IN-PERSON)

Group Assignment: Working and Q/A session

This will be a working session and a Q/A session for the Group Assignment

SESSION 29 (LIVE IN-PERSON)

Assessment: Test-Exam

In this session we will have a concept summary review examination based on the topics covered in the classes.

Not to be considered as a final exam. Its score will be added to the rest of the evaluation items. No minimum passing grade required.

SESSION 30 (LIVE IN-PERSON)

Assessment: Group assignment presentation

In this session we will have the group assignment presentation.

EVALUATION CRITERIA

criteria	percentage	Learning Objectives	Comments
Final Exam	30 %		
Individual Assignment	25 %		Individual Assignment
Group Assignment	30 %		Group Assignment
Class Participation	15 %		

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 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

- Richard S. Sutton, Andrew G. Barto.. *Reinforcement Learning.* ISBN 9780262039246 (Digital)
- Chip Huyen. (2022). *Designing Machine Learning Systems..* O'Reilly. ISBN 01234567890 (Digital)
- Delip Rao. (2019). *Natural Language Processing with Pytorch..* ISBN 01234567890 (Digital)
- Lewis Tunstall, Leandro Von Werra and Thomas Wolf.. (2022). *Natural Language Processing with Transformers.* O'Reilly. ISBN 01234567890 (Digital)
- Lakshmanan. (2021). *Practical Machine Learning for Computer Vision.* O'Reilly. ISBN 01234567890 (Digital)
- Kulkarni. *Computer Vision projects with Pytorch. Design and develop production grade models.* ISBN 01234567890 (Digital)

- Laura Graesser. *Foundations of Deep Reinforcement Learning.* 2019. ISBN 0135172381 (Digital)
- Marke Trevel. (2020). *Introducing MLOps: How to scale machine learning in the enterprise.* O'Reilly. ISBN 1492083291 (Digital)
- Noah Gift & Alfredo Deza. (2021). *Practical MLOps: Operationalizing Machine Learning Models.* O'Reilly. ISBN 1098103017 (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.