ADVANCED AI

Bachelor in Data and Business Analytics BDBA SEP-2023
AAI-DBA.1C.4.M.A

Area Others
Number of sessions: 25
Academic year: 23-24
Degree course: FOURTH
Number of credits: 5.0
Semester: 1º
Category: OPTIONAL
Language: English

Professor: **LAURA SANCHEZ GARCIA**

E-mail: lsanchezg@faculty.ie.edu

Laura Sánchez García, PhD

lsanchezg@faculty.ie.edu

Laura has a mixed academic and industry background. On the academic side, Laura has a PhD, Master and a Bachelor in Physics from Universidad Autónoma de Madrid. She did her PhD in the field of Nanotechnology with a strong focus on experimental data acquisition and analysis. She has published more than 10 peer-reviewed scientific publications and has been speaker at numerous international conferences.On the industry part, she is working as Advanced Analytics Lead in Sandoz Farmacéutica and previously worked for McKinsey & Company as a Senior Data Scientist. She has experience in a broad variety of Data Science use cases and has worked in several industries across the globe.

lsanchezg@faculty.ie.edu

PREREQUISITES

- Probability, Statistics, Calculus and Algebra
- Python programming
- AI - Machine Learning Foundations
- AI - Machine Learning & Analytics

Examples will be given to the students and the subject will be taugth in Python (mostly in a Notebook format). Therefore, students are expected to have a fluent level of Python.

SUBJECT DESCRIPTION

Artificial Intelligence (AI) started in the aftermath of World War II as one of the newest fields in science and engineering. Shortly afterwards, AI transformed into a hot topic as its applications spread across numerous industries and businesses, including automation, natural language processing, recommended systems, fraud detection and disease diagnosis in medicine.

In recent years, the development of new Deep Learning techniques, together with the availability of large amounts of data and increased computing power has lead to novel applications and AI systems which can perform a variety of tasks at human-performance level.

The general objective of this course is to understand the basis of the newest AI applications based on Machine Learning and Deep Learning using a "from theory to practice" approach. We will apply theory and knowledge to real practical examples across different industries and study state-of-the-art algorithms and technologies.

LEARNING OBJECTIVES

In this course, you will learn state-of-the-art methodologies in AI used for different applications. The subject covers three important blocks of the AI modern landscape. Namely Advanced Computer Vision & Object Detection; Natural Language Processing and Advanced Sequence Models; and Generative Models.

We expect that at the end of the course, students will obtain a solid understanding of the theory and practical implementations of modern AI methodologies, with special focus on Deep Learning technicques for the above mentioned applications. The course has a strong practical component in which we will use Python libraries such as PyTorch and Tensorflow. Therefore, it is expected that students have a medium to advanced level of programming in Python.

In order to fill the possible gap in programming knowledge, the course has a section devoted to reviewing important programming concepts that will be used along the course.

TEACHING METHODOLOGY

Learning Activity	**Weighting**	**Estimated time a student should dedicate to prepare for and participate in**
Lectures	30.0 %	37.5 hours
Discussions	10.0 %	12.5 hours
Exercises in class, Asynchronous sessions, Field Work	30.0 %	37.5 hours
Group work	30.0 %	37.5 hours
Individual studying	0.0 %	0.0 hours
TOTAL	100.0 %	125.0 hours

PROGRAM

The program addresses the following topics:

- Advanced Computer Vision Applications
- Advanced Natural Language Processing Applications
- Generative Models

SUMMARY

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.

SESSIONS 1 - 2 (LIVE IN-PERSON)

Course introduction and overview of Artificial Intelligence.

Review of programming concepts, setup and best practices.

PyTorch and Tensorflow for AI applications.

Book Chapters: Deep Learning with PyTorch. Chapters 1 & 3

SESSIONS 3 - 4 (LIVE IN-PERSON)

Computer Vision & Object Detection (I)

Review of CNNs & architectures

LeNet-5, AlexNet, ResNet and related concepts.

SESSIONS 5 - 6 (LIVE IN-PERSON)

Computer Vision & Object detection (II)

YOLO algorithm

Region proposals

Article: You Only Look Once. Unified, Real-Time Object Detection (Arxiv, 2015)

SESSIONS 7 - 8 (LIVE IN-PERSON)

Advanced Sequence Modelling and Natural Language Processing (NLP)

Review of autorregressive models for Advanced AI & NLP: RNNs, LSTMs & GRU, Deep RNNs and examples (text generation, music generation)

Word representations, embeddings, embedding matrix.

Book Chapters: Generative Deep Learning. Teaching Machines to Paint, Write, Compose and Play. Chapters 6-8

Book Chapters: Deep Learning with Python. Chapter 6

Book Chapters: Generative Deep Learning. Chapter 5

SESSION 9 (LIVE IN-PERSON)

Advanced Sequence Modelling and Natural Language Processing (NLP)

Encoder-Decoder architectures (e.g., text translation, image captioning)

Book Chapters: Generative Deep Learning. Teaching Machines to Paint, Write, Compose and Play. Chapter 3

Book Chapters: Generative Deep Learning. Chapter 9.

SESSION 10 (LIVE IN-PERSON)

Advanced Sequence Modelling and Natural Language Processing (NLP)

Autoencoders & Variational Autoencoders

Book Chapters: Generative Deep Learning. Teaching Machines to Paint, Write, Compose and Play. Chapter 3

SESSIONS 11 - 12 (LIVE IN-PERSON)

Transformers & Self-attention mechanisms

Book Chapters: Generative Deep Learning. Chapter 3

SESSION 13 (LIVE IN-PERSON)

Transformers and self-attention for NLP and other applications

GPT & ChatGPT

Book Chapters: Lewis Tunstall, et al.: Natural Language Processing with Transformers; O'Reilly 2022

SESSION 14 (LIVE IN-PERSON)

Review/Questions/Inquiries.

SESSION 15 (LIVE IN-PERSON)

Midterm exam.

SESSION 16 (LIVE IN-PERSON)

Student's group project follow-up

SESSION 17 (LIVE IN-PERSON)

Review of theory behind Generative models & applied concepts

SESSIONS 18 - 19 (LIVE IN-PERSON)

Diffussion models

Book Chapters: David Foster: Generative Deep Learning

SESSION 20 (LIVE IN-PERSON)

Applications of Generative Models

SESSION 21 (LIVE IN-PERSON)

Review of notebooks and explanations

SESSION 22 (LIVE IN-PERSON)

.Students Group projects follow-up

SESSION 23 (LIVE IN-PERSON)

Review/Questions/Inquiries.

SESSION 24 (LIVE IN-PERSON)

Student's Group Project presentations.

SESSION 25 (LIVE IN-PERSON)

Final exam

EVALUATION CRITERIA

During the course students will be required to read material prior to the sessions, and to participate in discussions during class. Students will be given assignments and quizzes to be completed individually. These will be solved and discussed during class sessions where students are expected to participate actively. The group project will consist of solving a given business or research problem using some of the AI methods learned. The group will present the solution to the problem. The overall grading will be based on the following criteria:

criteria	percentage	Learning Objectives	Comments
Final Exam	20 %		final exam
Individual Work	20 %		includes assignments, voluntary exercises, research.
Midterm Exam	20 %		midterm exam
Group Presentation	30 %		includes quality of group project and delivery
Class Participation	10 %		includes active engagement in class asking/answering questions

RE-SIT / RE-TAKE POLICY

PARTICIPATION

The following criteria will be used to make a judgment about your class participation:

Quality (not quantity) of in class and forum participation. Specially, high quality comments revealing deep insights and a good understanding of the subject will be very considered as positive scores. Concise and precise presentation of your ideas will reveal that you have a good understanding of what you are saying. Basically, the ability of explaining deep or difficult concepts in simple terms. Frequency of relevant interventions. Doing the right intervention at the right time, which includes asking questions that may also help your peers to understand better which is being explained.

ASSIGNMENTS

Assignments on previously covered material or prework to prepare for the upcoming class will be given periodically. These can be in the form of quizzes and or short programming exercises. These quizzes will help you assess your overall understanding of the subject being studied and identify any caveat in your learning. The dates of these assignments will be announced by the professor weekly. Students should be prepared to solve these assignments as they comes in.

All assignments need to be submitted online via Turnitin on Campus Online. All other methods of submission will not be considered.

FINAL GROUP PROJECT

The group project has the highest weight in the overall grade because it is an important part of this course. Each group (composed of 4 -5 students) will be asked to choose an advanced topic related to the course content. The professor will suggest several topics to pick up at the beginning of the course. The solution to the problem should constitute a complete and working solution (e.g., uploaded as a repository in GitHub). This will mimic how professionals work and familiarize students with software version control. This solution has to be accompanied with a description of the project and a presentation. There will be a special session where students will present their solutions.

All assignments need to be submitted online via Turnitin on Campus Online. All other methods of submission will not be considered.

MIDTERM AND FINAL EXAMS

There will be a midterm and a final exam, each of them covering part of the program. The final exam may include questions from the part covered in the midterm exam as well. For the exams, you will need your own laptop. Each of the exams is split in two parts, a theoretical and a practical one. For the theoretical part you are allowed to bring an (2-sides) A4 sheet with all the formulae you may consider. For the practical one you have to make use of your computer and be prepare to complete programming questions.

Disclaimer:

NO QUESTIONS ARE ALLOWED DURING THE EXAMS. THE CHEAT-SHEET ALONG WITH ANY SCRAP PAPER WILL BE COLLECTED AND STAPLED TO YOUR EXAMS.

In order to pass the course, you need a minimum grade of 5 in the final exam. If your grade in the final exam does not reach the threshold value of 5, you will fail the course, even in the case in which your weighted average (computed using the table above) exceeds 5.0.

Notice that the date of the midterm exams could change and needs to be considered with flexibility. The precise date will be communicated to students two weeks ahead of time. The date of the final and the retake exam CANNOT BE CHANGED under any circumstances.

There exists the possibility that both, midterm and final exams, will be in the form of take home exercise. In such a case, the professor will adapt the exams to that format.

BIBLIOGRAPHY

Recommended

- Eli Stevens, Luca Antiga, Thomas Viehmann. (2020). *Deep Learning with PyTorch.* Manning. ISBN 978161729526 (Digital)
- David Foster. (2022). *Generative Deep Learning. Teaching Machines to Paint, Write, Compose and Play.* O'Reilly. ISBN 978149204194 (Digital)
- L. Tunstall, L. von Werra, T. Wolf. (2022). *Natural Language Processing with Transformers. Building Language Applications with Hugging Face..* O'Reilly. ISBN 978109810324 (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.

ADVANCED TOPIC: FINANCIAL ANALYTICS

Bachelor in Data and Business Analytics BDBA SEP-2023
FA-DBA.4.M.A

Area Others
Number of sessions: 15
Academic year: 23-24
Degree course: FOURTH
Number of credits: 3.0
Semester: 2º
Category: COMPULSORY
Language: English

Professor: **MANOEL FERNANDO GADI ALONSO**

E-mail: mfalonso@faculty.ie.edu

Academic Background:

- Executive MBA, Business Administration and Management, IE Business School, Spain
- Master in Computer Science and Statistics, University of Sao Paulo, Brazil
- Computer Science Degree, University of Sao Paulo, Brazil

Academic Experience:

- Risk & Fraud Analytics - Master in Business Analytics & Big Data at IE Business School
- Statistical Programming Python - Master in Business Analytics & Big Data at IE Business School
- Building a Fintech - Master in Business Analytics & Big Data at Universidad de Alcalá de Henares
- Math and Stats for Big Data - Master in Business Analytics & Big Data at U-TAD

Corporate Experience:

- Founder and CEO of Suncaged Analytics Consulting Europe (Madrid: 2019-curr)
- Director of Analysis & Reporting, Altamira Real Estate (Madrid: 2018-2019)
- Director of Credit Rating, Big Data and Business Analytics, Bravo Capital (Madrid: 2015-2018)
- Head of R&D for Risk Analytics Area, Santander Bank Headquarters Spain (Madrid: 2012-2015)
- R&D Risk Analytics Area Manager, Santander Bank United Kingdom (Milton Keynes: 2008-2012)
- Credit Risk Modelling Manager, Santander Bank Brazil (Sao Paulo: 2007-2008)
- Credit Risk Supervisor, ibi bank (C&A group) (Sao Paulo: 2006-2007)
- Credit Card Risk analyst, Citibank Brazil (Sao Paulo: 2003-2005)

Office Hours

Office hours will be on request. Please contact at:

mfalonso@faculty.ie.edu

SUBJECT DESCRIPTION

Finance is a vast subject that involves the management, creation, and study of money and investments, including acquiring capital and spending or investing that money. Financial markets serve as the means to facilitate the flow of money through investments and other financial instruments between different parties. Financial instruments are monetary contracts that can be created, traded, modified, and settled. These instruments include cash, ownership interests in an entity, contractual rights to receive or deliver currency, debt such as bonds and loans, equity such as shares, and derivatives such as options, futures, and forwards. This diversity of financial instruments enables the flow of money through economies worldwide.

Recent technological advances have propelled finance into a new era, driven by two major forces. The first is programmatic access to real-time financial data, which generates vast amounts of data that humans alone cannot process. As a result, computational power and algorithms increasingly drive financial markets instead of human traders. The second force is the growing importance of artificial intelligence in finance. An increasing number of financial institutions are investing in machine learning (ML) and deep learning (DL) algorithms to improve their operations, trading, and investment performances.

This course covers essential topics in Financial Data Science that are at the core of these technological evolutions, including financial data and preprocessing, technical analysis, financial time series modeling, factor models and volatility modeling, Montecarlo simulations, asset allocation and portfolio optimization, and financial machine learning.

LEARNING OBJECTIVES

The course aims to cover the basic techniques involved in Financial Data Science with a practical emphasis on their application, starting with traditional methods such as time series analysis and factor models and progressing to more modern techniques such as machine learning. The objective is to provide students with an adequate understanding of theory and data analytics knowledge previously acquired to determine which technique best solves a specific problem.

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:

Learning Activity	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	26.67 %	20.0 hours
Discussions	6.67 %	5.0 hours
Exercises in class, Asynchronous sessions, Field Work	6.67 %	5.0 hours
Group work	46.67 %	35.0 hours
Individual studying	13.33 %	10.0 hours
TOTAL	100.0 %	75.0 hours

PROGRAM

SESSION 1 (LIVE IN-PERSON)

Financial Data Science

The session begins by introducing the course, its objectives, methodology, and evaluation. The second part of the session focuses on introducing Finance, including markets and instruments, and how data science is transforming the field.

SESSION 2 (LIVE IN-PERSON)

Financial data and preprocessing

In this session we explore how financial data is different from other types of data commonly used in machine learning tasks. We will learn the different techniques to access financial data from different sources and how to preprocess it for further analysis.

SESSION 3 (LIVE IN-PERSON)

Financial data and preprocessing (Practice)

In this session we put in practice what we have learnt in the previous session. We will download data for different instruments and preprocess it using Python.

SESSION 4 (LIVE IN-PERSON)

Technical analysis

Technical analysis is commonly used by traders and when creating trading systems. In this session we demonstrate some fundamental basics of technical analysis as well as how to create elegant dashboards in Python. We will be able to draw some insights into patterns emerging from a selection of the most commonly used metrics.

SESSION 5 (LIVE IN-PERSON)

Technical analysis (Practice)

In this session we put in practice what we have learnt in the previous session. Students will perform technical analysis of different instruments.

SESSION 6 (LIVE IN-PERSON)

Financial time series modelling

Traditionally financial applications have been addressed using methods from Econometrics. This session introduces the basics of time series modeling. Then, we look at two of the most widely used approaches of time series modeling, exponential smoothing methods and ARIMA class models. We present Python libraries to perform different statistical analyses and time series forecasting.

SESSION 7 (LIVE IN-PERSON)

Financial time series modelling (Practice)

In this session we put in practice what we have learnt in the previous session. We will use Python to perform financial time series analysis.

SESSION 8 (LIVE IN-PERSON)

Factors models and volatility modelling

In the first part of this session we will introduce factor models. Factor models are financial models that use factors (technical, fundamental, macroeconomic or alternate) to define a security's risk and returns. These models are linear, as they define the securities returns to be a linear combination of factor returns weighted by the securities factor exposures. We will estimate various factor models starting with the simplest one-factor model and then explain how to estimate more advanced three-, four-, and five-factor models.

In the second part of the session we address volatility modelling. We introduce the concept of volatility forecasting using (G)ARCH class models, how to choose the best-fitting model, and how to interpret your results.

Factors models and volatility modelling (Practice)

In this session we put in practice what we have learnt in the previous session. We will learn how to create factor models and model volatility with GARCH using Python.

SESSION 9 (LIVE IN-PERSON)

Factors models and volatility modelling

In the first part of this session we will introduce factor models. Factor models are financial models that use factors (technical, fundamental, macroeconomic or alternate) to define a security's risk and returns. These models are linear, as they define the securities returns to be a linear combination of factor returns weighted by the securities factor exposures. We will estimate various factor models starting with the simplest one-factor model and then explain how to estimate more advanced three-, four-, and five-factor models.

In the second part of the session we address volatility modelling. We introduce the concept of volatility forecasting using (G)ARCH class models, how to choose the best-fitting model, and how to interpret your results.

Factors models and volatility modelling (Practice)

In this session we put in practice what we have learnt in the previous session. We will learn how to create factor models and model volatility with GARCH using Python.

SESSION 10 (LIVE IN-PERSON)

Asset allocation and portfolio optimization

Investments in real life are done using a set of assets pulled in a portfolio to diversify risk. In this session we introduce Modern Portfolio Theory (MPT) and show how to obtain the Efficient Frontier and identify specific portfolios, such as minimum variance or the maximor Sharpe ratio. We also show how to evaluate the performance of such portfolios.

Asset allocation and portfolio optimization (Practice)

In this session we put in practice what we have learnt in the previous session. We will learn how to do portfolio analysis using Python.

SESSION 11 (LIVE IN-PERSON)

Asset allocation and portfolio optimization

Investments in real life are done using a set of assets pulled in a portfolio to diversify risk. In this session we introduce Modern Portfolio Theory (MPT) and show how to obtain the Efficient Frontier and identify specific portfolios, such as minimum variance or the maximor Sharpe ratio. We also show how to evaluate the performance of such portfolios.

Asset allocation and portfolio optimization (Practice)

In this session we put in practice what we have learnt in the previous session. We will learn how to do portfolio analysis using Python.

SESSION 12 (LIVE IN-PERSON)

Financial machine learning

In this session we introduce advanced techniques in financial analytics. We will explore applications of Machine Learning in finance, such as price prediction or portfolio selection.

SESSION 13 (LIVE IN-PERSON)

Financial machine learning (Practice)

In this session we put in practice what we have learnt in the previous session. We will use machine learning to address different problems in finance.

SESSION 14 (LIVE IN-PERSON)

Final Exam (Quiz)

In this session we do an in-class exercise/exam with all the course content.

SESSION 15 (LIVE IN-PERSON)

Group work presentation

EVALUATION CRITERIA

Your final grade in the course will be based on both individual and group work of different characteristics that will be weighted in the following way:

criteria	percentage	Learning Objectives	Comments
Final Exam	40 %		
Group Presentation	40 %		
Class Participation	20 %		

RE-SIT / RE-TAKE POLICY

A. Class participation and discussion

Class participation will be evaluated based on the following criteria:

Quality (not quantity) of your participation in class discussion: The most important dimension of participation concerns what it is that you are saying. A high quality comment reveals depth of insight, rigorous use of case evidence, consistency of argument, and realism. Frequency refers to the attainment of a threshold quantity of contributions that is sufficient for making a reliable assessment of comment quality. The logic is simple: if contributions are too few, one cannot reliably assess the quality of your remarks. However, once threshold quantity has been achieved, simply increasing the number of times you talk does not automatically improve your evaluation. Beyond the threshold, it is the quality of your comments that must improve. In particular, one must be especially careful that in claiming more than a fair share of “airtime”, quality is not sacrificed for quantity. Finally, your attempts at participation should not be such that the instructor has to “go looking for you”. You should be attempting to get into the debate on a regular basis.

You might want to avoid being classified as one of the following types of students:

- Repeaters, i.e., students that, consciously or unconsciously, make comments that are really just repeats/rephrasing of what has already been said (by other students, or you). This wastes time and adds nothing to learning.
- Ramblers, i.e., students that take a lot of time to say simple things or they may tell long personal/professional stories, or they roam into topics that are not relevant, or simply make low-quality comments just to participate. They waste valuable time and prevent other students from being able to participate.
- Students that have been distracted (by social networks, etc.) or who have stopped paying attention and then, later on, when they realized they have missed a term or concept, they ask you about it.

B. Final exam

Students will be asked to complete quizzes to assess the overall understanding of the subject being studied.

C. Group Work

The group project is an integral part of this course. Each group will be asked to work on a project, prepare and deliver a presentation.

BIBLIOGRAPHY

Compulsory

- Yves Hilpisch. (2019). *Python for Finance. Mastering Data-Driven Finance.* 2nd edition. O'Reilly Media, Inc, USA. ISBN 1492024333 (Digital)

- Eryk Lewinson. (2020). *Python for Finance Cookbook: Over 50 recipes for applying modern Python libraries to financial data.* 1st. Packt. ISBN 1789618517

(Digital)

- Yves J. Hilpisch. (2010). *Artificial Intelligence in Finance: A Python-Based Guide.* First. O'Reilly Media, Inc, USA. ISBN 1492055433 (Digital)

Recommended

- Marcos M López de Prado. (2020). *Machine Learning for Asset Managers.* First. Cambridge University Press. ISBN 1108792898 (Digital)
- Matthew F. Dixon, Igor Halperin, Paul Bilokon. (2020). *Machine Learning in Finance: From Theory to Practice.* 1st. Springer. ISBN 3030410676 (Digital)
- Frederic Mishkin. (2018). *Financial Markets and Institutions.* Pearson. ISBN 1292215003 (Digital)
- Irene Aldridge. (2021). *Big Data Science in Finance.* First. John Wiley & Sons Inc. ISBN 9781119602989 (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.

AI APPLICATIONS IN HEALTHCARE

Bachelor in Data and Business Analytics BDBA SEP-2023
AIAH-DBA.1C.4.M.A

Area Others
Number of sessions: 25
Academic year: 23-24
Degree course: FOURTH
Number of credits: 5.0
Semester: 1º
Category: OPTIONAL
Language: English

Professor: **ADRIÁN SÁNCHEZ MORALES**

E-mail: adrians@faculty.ie.edu

Dr Sánchez holds a PhD in Artificial Intelligence from Technical University of Cartagena and Carlos III of Madrid, and a BSc in Telecommunication Engineering from the same university. He has taken part in research projects involving deep neural networks, computer vision and e-health technologies. He has several years of experience in private companies, working as a data scientist on all kind of sectors as well as project manager.

More details in https://www.linkedin.com/in/adrián-sánchez-morales/

If your question cannot be properly answered via email and/or you would prefer to meet in person, please make an appointment to meet with me on the university campus during my scheduled office hours. Office hours will be determined during the semester and posted on Campus Online.

PREREQUISITES

Some recommended prerequisites for this type of course include:

- Basic knowledge of programming: Students should have a basic understanding of programming concepts, such as variables, loops, and conditional statements. Knowledge of a programming language like Python is also recommended.
- Knowledge of statistics: Students should have a good understanding of statistical concepts like probability, mean, median, mode, standard deviation, and hypothesis testing. This will help them analyze and interpret data using statistical techniques.
- Understanding of machine learning concepts: Students should have an understanding of machine learning concepts like supervised and unsupervised learning, classification, regression, clustering, and neural networks. This will help them understand and implement AI

algorithms for healthcare applications.

- Familiarity with healthcare data is a plus: Students should have a basic understanding of healthcare data, including medical terminology, common healthcare data formats, and electronic health records (EHRs). This will help them work with healthcare data and understand the specific challenges associated with healthcare data.

Overall, a strong foundation in programming, statistics, machine learning, and healthcare data is essential for success in a course on AI applications in healthcare.

SUBJECT DESCRIPTION

This course is designed to provide students with an understanding of the potential applications of artificial intelligence (AI) in the field of healthcare. Topics covered will include AI techniques such as machine learning, deep learning, natural language processing, and robotics, and how these techniques can be applied to various healthcare scenarios such as disease diagnosis, treatment planning, drug discovery, and patient monitoring. The course will also explore ethical considerations related to the use of AI in healthcare.

LEARNING OBJECTIVES

Upon completion of this course, students will be able to:

- Understand the basic concepts and techniques of artificial intelligence, including machine learning, deep learning, natural language processing, and robotics.
- Apply AI techniques to healthcare scenarios such as disease diagnosis, treatment planning, drug discovery, and patient monitoring.
- Evaluate the potential benefits and limitations of AI applications in healthcare.
- Understand ethical considerations related to the use of AI in healthcare.

TEACHING METHODOLOGY

Learning Activity	**Weighting**	**Estimated time a student should dedicate to prepare for and participate in**
Lectures	30.0 %	37.5 hours
Discussions	10.0 %	12.5 hours
Exercises in class, Asynchronous sessions, Field Work	30.0 %	37.5 hours
Group work	20.0 %	25.0 hours
Individual studying	10.0 %	12.5 hours
TOTAL	100.0 %	125.0 hours

PROGRAM

SESSION 1 (LIVE IN-PERSON)

Course Introduction

In this session we will review the course logistics and organization together with basic concepts and requirements along the subject.

SESSION 2 (LIVE IN-PERSON)

Introduction to Healthcare Intelligence

In this session, we provide a comprehensive overview of the current and future uses of Artificial Intelligence in healthcare. We will start learning about the meaning of Healthcare Intelligence and its objectives.

SESSION 3 (LIVE IN-PERSON)

Introduction to AI and Machine Learning

In this session, we will start learning about the main ML models and types of learning algorithms in AI. We'll see some examples of learning models. Finally, we'll do some python exercises to review basic concepts of the programming language.

SESSION 4 (LIVE IN-PERSON)

Regression ML Models

In this session we will look at the details of regression, from linear to more complex regression models. Some problems will be included to be solved in the session.

SESSION 5 (LIVE IN-PERSON)

Practical Session

First practical session of the course, where we work on a real regression problem in Healthcare.

SESSION 6 (LIVE IN-PERSON)

Classification ML Models

In this session we will look at the details of classification models such as Logistic regression, SVMs and Decision Trees. Some problems will be included to be solved in the session.

SESSION 7 (LIVE IN-PERSON)

Practical Session

We work on a real classification problem in Healthcare.

SESSION 8 (LIVE IN-PERSON)

Unsupervised Learning in Healthcare

In these sessions we will learn the challenge of unsupervised learning and the way of dealing with missing data in healthcare applications.

SESSION 9 (LIVE IN-PERSON)

Final Project Introduction

Group Work Introduction

SESSION 10 (LIVE IN-PERSON)

Deep Learning

Introduction to deep learning methods and overview of the basic algorithms with neural networks: MLP and Autoencoders. Applications in Healthcare.

SESSION 11 (LIVE IN-PERSON)

Practical Session

Application of Deep Learning over a real AI-based problem in Healthcare.

SESSION 12 (LIVE IN-PERSON)

AI for Medical Imaging

After knowing the basics in neural networks, we will introduce the Convolutional Neural Networks. We will review their theory and applications in Healthcare.

SESSION 13 (LIVE IN-PERSON)

Practical Session

In this session, we will work on a real Computer Vision application in Healthcare.

SESSION 14 (LIVE IN-PERSON)

Sequence Modeling and NLP

After knowing the basics in neural networks, we will introduce the Recurrent Neural Networks and LSTMs. We will review their theory and applications in Healthcare.

SESSION 15 (LIVE IN-PERSON)

Introduction to Transformers

In this session we will introduce this novel architecture that aims to solve sequence-to-sequence tasks while handling long-range dependencies with ease. We will see some real applications too.

SESSION 16 (LIVE IN-PERSON)

Practical Session

Working on a real NLP application with a text dataset in healthcare.

SESSION 17 (LIVE IN-PERSON)

Generative AI in Healthcare

Introduction to GAI and its potential in healthcare applications.

SESSION 18 (LIVE IN-PERSON)

AI for Health Monitoring

Introduction to health monitoring, wearable devices, sensors and applications in telemedicine.

SESSION 19 (LIVE IN-PERSON)

AI for medical robotics

Review of the potential of AI and Robotics in healthcare.

SESSION 20 (LIVE IN-PERSON)

AI for Healthcare Ethics and Governance

Guidance on Ethics & Governance of Artificial Intelligence for Health.

SESSION 21 (LIVE IN-PERSON)

AI and Future Healthcare Trends

Conclussions of the course and review of current companies working on AI applied to Healthcare in the world.

SESSION 22 (LIVE IN-PERSON)

Advanced Real Application

David Gomez Ullate Oteiza (one of our internal professors and head of math for the SST) will present his research and prototype about computer vision devices to monitor babies.

SESSION 23 (LIVE IN-PERSON)

Group Work/Individual Presentations I

SESSION 24 (LIVE IN-PERSON)

Group Work/Individual Presentations II

SESSION 25 (LIVE IN-PERSON)

Exam

EVALUATION CRITERIA

Throughout this course, you will be asked to read material related to the sessions, participate in discussions, complete individual assignments, participate in in-class quizzes, implement a final project (creating a Python program which provides a solution to a business challenge) and present the outcome of this project in class. Specifically, grading will be based on the following criteria:

1. Class participation – discussion: you are expected to attend every class and participate in the discussions and class activities (problems, etc.). The basic criteria in grading your participation are: a) your presence in each session, b) your (quality) contributions to the group discussion. Lively discussions in the classroom are always encouraged, however, make sure that you provide constructive comments which contribute to the learning experience of the whole class.
2. Final project implementation: a major part of this course's learning experience consists of a final project (individual or in groups) which will be focusing on designing and implementing an algorithmic solution to a business problem. The definition of the problem as well as the main

requirements will be provided to you during the sessions, so that you have only to focus on designing and implementing the best solution. Remember, creativity is always rewarded!

3. Individual work: you are expected to solve and present a number of problems and exercises that will be presented along the practical sessions.
4. Final exam: at the end of the course, you will have to pass an individual exam. The minimum pass grade is 3.5.

criteria	percentage	Learning Objectives	Comments
Final Exam	20 %		
Individual Work	30 %		
Group Work	35 %		
Class Participation	15 %		

RE-SIT / RE-TAKE POLICY

BIBLIOGRAPHY

Recommended

- Adam Bohr and Kaveh Memarzadeh. *Artificial Intelligence in Healthcare.* ISBN 9780128184387 (Digital)

- Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani. *An Introduction to Statistical Learning.* ISBN 9781071614204 (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.

APPLIED MACHINE LEARNING USING GRAPHS

Bachelor in Data and Business Analytics BDBA SEP-2023
AMLUG-DBA.1C.4.M.A

Area Others

Number of sessions: 15

Academic year: 23-24

Degree course: FOURTH

Number of credits: 3.0

Semester: 1º

Category: OPTIONAL

Language: English

Professor: **MANUELE LEONELLI**

E-mail: mleonelli@faculty.ie.edu

Manuele Leonelli is an Assistant Professor in the School of Human Sciences and Technology at IE University. He obtained a PhD in Statistics from the University of Warwick in 2015 under the supervision of Jim Q. Smith. He then won a CAPES post-doctoral fellowship working at the Federal University of Rio de Janeiro, Brazil, under the direction of Dani Gamerman. Before joining IE University, he was a Lecturer in Statistics in the School of Mathematics and Statistics at the University of Glasgow. Manuele's research focuses on probabilistic graphical models for decision-making under uncertainty and inference over extreme values, with a focus on approximated inferential algorithms within the Bayesian paradigm. His PhD thesis "Bayesian decision support in complex systems: an algebraic and graphical approach," won the John Copas Prize for the best PhD Thesis in Statistics at the University of Warwick in 2015.

mleonelli@faculty.ie.edu

PREREQUISITES

Basic knowledge of mathematics, probability, and programming.

SUBJECT DESCRIPTION

Probabilistic graphical models (PGMs) are a rich framework for encoding probability distributions over complex domains: joint (multivariate) distributions over large numbers of random variables that interact with each other. These representations sit at the intersection of statistics and computer science, relying on concepts from probability theory, graph algorithms, machine learning, and more. They are the basis for the state-of-the-art methods in a wide variety of applications, such as medical diagnosis, image understanding, speech recognition, natural language processing, and many, many more. They are also a foundational tool in formulating many machine learning problems.

The module focus on a specific PGM called Bayesian network, which is the most commonly used to explore the dependence structure in datasets. The course discusses both its theoretical properties as well as its use in practice. The focus will then move to inferential problems, i.e. how a Bayesian network can be used to answer questions, and learning algorithms, i.e. how the structure of the Bayesian network can be inferred from data.

LEARNING OBJECTIVES

At the end of the course, you will be able to:

- Describe the relationship between variables by means of a Bayesian network.
- Use Bayesian networks to perform inferential tasks.
- Learn Bayesian networks from data.
- Implement Bayesian networks in R.

TEACHING METHODOLOGY

Learning Activity	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	20.0 %	15.0 hours
Discussions	13.33 %	10.0 hours
Exercises in class, Asynchronous sessions, Field Work	20.0 %	15.0 hours
Group work	26.67 %	20.0 hours
Individual studying	20.0 %	15.0 hours
TOTAL	100.0 %	75.0 hours

PROGRAM

SESSION 1 (LIVE IN-PERSON)

Overview of the module, discussion of assessment methods, introduction to Bayesian Networks.

SESSION 2 (LIVE IN-PERSON)

Using R to summarize data, exploratory data analysis in R.

SESSION 3 (LIVE IN-PERSON)

Review of probability and statistics, conditional independence and its axioms.

SESSION 4 (LIVE IN-PERSON)

1st presentation of students' projects.

SESSION 5 (LIVE IN-PERSON)

Introduction to graph theory, directed acyclic graphs and undirected graphs.

SESSION 6 (LIVE IN-PERSON)

Bayesian networks, conditional independence and d-separation.

SESSION 7 (LIVE IN-PERSON)

Elicitation of Bayesian networks and their construction in R.

SESSION 8 (LIVE IN-PERSON)

2nd presentation of students' projects.

SESSION 9 (LIVE IN-PERSON)

Exact inference in Bayesian networks, moralization and junction trees.

SESSION 10 (LIVE IN-PERSON)

Estimation of the probabilities of a Bayesian networks, implementation in R.

SESSION 11 (LIVE IN-PERSON)

Learning Bayesian networks from data, implementation in R.

SESSION 12 (LIVE IN-PERSON)

3rd presentation of students' projects.

SESSION 13 (LIVE IN-PERSON)

Parameter sensitivity analysis, distances between Bayesian networks.

SESSION 14 (LIVE IN-PERSON)

Review class.

SESSION 15 (LIVE IN-PERSON)

4th presentation of students' project.

EVALUATION CRITERIA

criteria	percentage	Learning Objectives	Comments
Individual Presentation of project	35 %		Individual performance in the presentation of the groupwork
Group Project	45 %		Quality of the group project
Class Participation	20 %		Individual participation in the classroom

RE-SIT / RE-TAKE POLICY

BIBLIOGRAPHY

Recommended

- Koller, D., & Friedman, N. *Probabilistic Graphical Models: Principles and Practice.* MIT Press. ISBN 9780262258357 (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.

AT THE EDGE OF NEUROSCIENCE: MARKETING DECISIONS

Bachelor in Data and Business Analytics BDBA SEP-2023
AEN-DBA.1C.4.M.A

Area Others

Number of sessions: 15

Academic year: 23-24

Degree course: FOURTH

Number of credits: 3.0

Semester: 1º

Category: OPTIONAL

Language: English

Professor: **ROSARIO PERICUESTA CAMACHO**

E-mail: rpericuesta@faculty.ie.edu

ROSARIO PERICUESTA CAMACHO

Charo Pericuesta holds a Bachelor in Economics from the Universidad de Salamanca, a Bachelor in Market Research from the Universidad Autónoma de Madrid and a Master in Marketing and Sales Management from ESIC University in Madrid.

She has been more than a decade in Nielseniq, advising multinational companies in their strategy of innovation of their product portfolio using consumer data analysis.

Prior to that she worked:

- In the consumer insights department of Mondelez (Kraft) developing research plans to explore the different marketing initiatives of the company and using different market research techniques to understand the consumer behaviour and support decision making for the strategy of the brand;

- In United Planet in Boston as Marketing Project Manager;

- In Research International as Market Research Executive being seconded to the Chicago Office;

- In Russell Bedford International in London as Marketing Assistant.

rpericuesta@faculty.ie.edu

PREREQUISITES

No specific academic background required. Passion to learn.This course is particularly useful for those interested in pursuing marketing and marketing research related careers.

SUBJECT DESCRIPTION

Neuroscience helps us to understand the decision making process that we undergo as consumers. For our brain, context matters and everything is relative. Consumers not always disclosure the drivers behind their decisions, that's why being able to measure such hidden drivers becomes so important.

This course provides a first approach to how Neuroscience is currently applied to Marketing. In particular understanding how attention, memory or emotion affect the decisions of consumers and how brands and value perceptions are built in consumers' minds.

LEARNING OBJECTIVES

The objectives are:

- Understanding how consumers make decisions.
- Understanding what's outside consumers' awareness and how to measure it.
- Leverage how our brain automatically works to make marketing experiences more powerful.

TEACHING METHODOLOGY

Learning Activity	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	13.33 %	10.0 hours
Discussions	13.33 %	10.0 hours
Exercises in class, Asynchronous sessions, Field Work	20.0 %	15.0 hours
Group work	26.67 %	20.0 hours
Individual studying	26.67 %	20.0 hours
TOTAL	100.0 %	75.0 hours

PROGRAM

SESSION 1 (LIVE IN-PERSON)

SESSION 1: Introduction

- Course overview
- Introduction to consumer Neuroscience
- Landscape of implicit tools and their possible use in marketing

SESSION 2 (LIVE IN-PERSON)

SESSION 2: How our rational decision-making is hindered by cognitive biases I: Illusions

SESSION 3 (LIVE IN-PERSON)

SESSION 3: How our rational decision-making is hindered by cognitive biases II: What factors affect our decisions?

SESSION 4 (LIVE IN-PERSON)

SESSION 4: How we think: System 1 and System 2 – Thinking Fast and Slow

Implicit and explicit: Two modes of thinking

- System 1: The automatic mode
- System 2: The lazy controller
- System 1 and System 2 in decision making

SESSION 5 (LIVE IN-PERSON)

SESSION 5: Practice: System 1 and System 2

SESSION 6 (LIVE IN-PERSON)

SESSION 6: Attention: Standing out from competitors. Modern view of attention in marketing.

SESSION 7 (LIVE IN-PERSON)

SESSION 7: Power of familiarity in marketing

- Memory: How familiarity impact on our decision-making process
- The Mere Exposure Effect
- The brain as an association machine

SESSION 8 (LIVE IN-PERSON)

SESSION 8: Practice: Brands as associative memory networks

SESSION 9 (LIVE IN-PERSON)

SESSION 9: Approach or Avoid? How emotions motivate our decision-making process

SESSION 10 (LIVE IN-PERSON)

SESSION 10: How the halo effect affects consumers' responses

SESSION 11 (LIVE IN-PERSON)

SESSION 11: Value beyond price. How non-conscious processes impact value calculations.

SESSION 12 (LIVE IN-PERSON)

SESSION 12: Neuroscience and Social Marketing

SESSION 13 (LIVE IN-PERSON)

SESSION 13: Ethics of Consumer Neuroscience

SESSION 14 (LIVE IN-PERSON)

SESSION 14: – Market Research today and tomorrow: the impact of Neuroscience

SESSION 15 (LIVE IN-PERSON)

Exam

EVALUATION CRITERIA

criteria	percentage	Learning Objectives	Comments
Final Exam	30 %		
Individual Work	20 %		
Group Presentation	30 %		
Class Participation	20 %		

RE-SIT / RE-TAKE POLICY

BIBLIOGRAPHY

Recommended

- Edited by Moran Cerf and Manuel Garcia-Garcia. (2017). *Consumer Neuroscience.* The MIT Press. ISBN 9780262036597 (Printed)
- Daniel Kahneman. *Thinking, fast and slow.* PENGUIN BOOKS LTD. ISBN 0141033576 (Printed)
- Richard H. Thaler and Cass R. Sunstein. *Nudge, the final edition..* Penguin Books. ISBN 9780143137009 (Printed)
- Dan Ariely. *Predictably Irrational.* Generic. ISBN 9780062018205 (Printed)
- Phil Rosenzweig. *The halo effect… and the Eight Other Business Delusions That Deceive Managers.* FREE PR. ISBN 9781476784038 (Printed)

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.

AUTOMATION AND SECURITY ENGINEERING

Bachelor in Data and Business Analytics BDBA SEP-2023
ASE-DBA.1C.4.M.A

Area Others
Number of sessions: 25
Academic year: 23-24
Degree course: FOURTH
Number of credits: 5.0
Semester: 1º
Category: OPTIONAL
Language: English

Professor: **PEDRO FIGUEIRAS VICENTE**

E-mail: pfigueirasv@faculty.ie.edu

Pedro Figueiras, Senior Data Expert

* BSc Physics by USC. MBA by EEN Business school. Master's degree in AI research by UIMP

* Executive education: Digital transformation by IE Business school. Internet of Things by MIOTI

* 15+ years experience in Technology, both in Business and Engineering positions

* C-level positions in Industrial companies

* Freelance consultant for Data Engineering and Data Science projects

* Private investor. Data & AI Startups mentor.

pfigueirasv@faculty.ie.edu

PREREQUISITES

SUBJECT DESCRIPTION

This course explores the principles, techniques, and tools used in securing automated systems and software. Students will learn how to design, implement, and maintain secure systems by leveraging the latest technologies and methodologies in security engineering. The course emphasizes the importance of incorporating security measures throughout the entire software development lifecycle, with a particular focus on continuous integration and continuous delivery (CI/CD) pipelines and the cloud.

LEARNING OBJECTIVES

Upon successful completion of this course, students will be able to:

- Utilize the CI/CD method in software development and its related tools
- Understand the fundamental principles of Security Engineering in the context of Automation systems
- Identify potential security threats and vulnerabilities
- Develop strategies for mitigating security risks in automated Cloud deployments
- Apply secure coding practices and guidelines

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:

Learning Activity	**Weighting**	**Estimated time a student should dedicate to prepare for and participate in**
Lectures	48.0 %	60.0 hours
Discussions	8.0 %	10.0 hours
Exercises in class, Asynchronous sessions, Field Work	16.0 %	20.0 hours
Group work	16.0 %	20.0 hours
Individual studying	12.0 %	15.0 hours
TOTAL	100.0 %	125.0 hours

PROGRAM

SESSION 1 (LIVE IN-PERSON)

Introduction

- Overview of automation systems and their security challenges
- Importance of security engineering in automation systems

SESSION 2 (LIVE IN-PERSON)

CI/CD concepts I

- Fundamentals of CI/CD. Basic tools.
- In-class workshop: Introducing Git

SESSION 3 (LIVE IN-PERSON)

CI/CD concepts II

- Process automation
- In-class workshop: Jenkins

SESSION 4 (LIVE IN-PERSON)

CI/CD concepts III

- Exercise 1: connecting Github and Jenkins
- In-class workshop: Docker

SESSION 5 (LIVE IN-PERSON)

CI/CD concepts IV

- The deployment stage
- In-class workshop: CI/CD working
- Assignment CI/CD

SESSION 6 (LIVE IN-PERSON)

Fundamentals of Security Engineering I

- Secure design principles
- Threat modeling and risk assessment

SESSION 7 (LIVE IN-PERSON)

Fundamentals of Security Engineering II

- Common attack vectors
- Identifying potential vulnerabilities in automation systems

SESSION 8 (LIVE IN-PERSON)

Fundamentals of Security Engineering III

- Types of security controls
- Implementing effective countermeasures

SESSION 9 (LIVE IN-PERSON)

Fundamentals of Security Engineering IV

- Input validation and output encoding
- Exercise: proposing a Risk assesment & countermeasures

- Assignment SE

SESSION 10 (LIVE IN-PERSON)

Group project I

- Definition and methodology
- Groups creation, case presentation and next steps

SESSION 11 (LIVE IN-PERSON)

Security in CI/CD I

- Environment isolation and hardening
- Code review and static analysis

SESSION 12 (LIVE IN-PERSON)

Security in CI/CD II

- Dynamic analysis and penetration testing
- Vulnerability management and patching

SESSION 13 (LIVE IN-PERSON)

Security in CI/CD III

- In-class workshop: security implementation in a CI process

SESSION 14 (LIVE IN-PERSON)

Security in CI/CD IV

- Secure deployment practices
- Configuration management and secrets management
- Assignment SCI

SESSION 15 (LIVE IN-PERSON)

Cloud Security I

- Shared responsibility model
- Cloud security best practices

SESSION 16 (LIVE IN-PERSON)

Cloud Security II

- IAM concepts and components
- Managing access to cloud resources
- In-class workshop: configuring Azure to use resources

SESSION 17 (LIVE IN-PERSON)

Cloud security III

- Infrastructure as Code (IaC)
- Security configurations and policies
- In-class workshop: AWS policies

SESSION 18 (LIVE IN-PERSON)

Cloud security IV

- Data encryption and key management
- Data loss prevention and recovery
- Assignment CS

SESSION 19 (LIVE IN-PERSON)

Emerging trends I

- Basics of cryptography
- Encryption, hashing, and digital signatures

SESSION 20 (LIVE IN-PERSON)

Emerging trends II

- Privacy Engineering
- In-class workshop: selective encryption in the cloud

SESSION 21 (LIVE IN-PERSON)

Emerging trends III

- Extending Security Engineering to decentralized architectures
- Blockchain Security in Automation Systems

SESSION 22 (LIVE IN-PERSON)

Emerging trends IV

- Artificial intelligence and machine learning in security
- The future of automation security

SESSION 23 (LIVE IN-PERSON)

Group project II

- Validation of project development

SESSION 24 (LIVE IN-PERSON)

Course recap

- Putting together the pieces
- In-class workshop: complete CI/CD pipeline

SESSION 25 (LIVE IN-PERSON)

Group project III

- Final presentation

EVALUATION CRITERIA

criteria	percentage	Learning Objectives	Comments
Intermediate tests	50 %		Total of the four individual assignments that have been proposed
Group Work	40 %		It includes all the presentations and delivered code
Class Participation	10 %		See below

RE-SIT / RE-TAKE POLICY

Class participation:

To be evaluated based on the following criteria:

- Quality (not quantity) of your participation in class discussion: The most important dimension of participation concerns what it is that you are saying. A high quality comment reveals depth of insight, rigorous use of case evidence, consistency of argument, and realism.
- Frequency refers to the attainment of a threshold quantity of contributions that is sufficient for making a reliable assessment of comment quality. The logic is simple: if contributions are too few, one cannot reliably assess the quality of your remarks. However, once threshold quantity has been achieved, simply increasing the number of times you talk does not automatically improve your evaluation. Beyond the threshold, it is the quality of your comments that must improve. In particular, one must be especially careful that in claiming more than a fair share of “airtime”, quality is not sacrificed for quantity.
- Finally, your attempts at participation should not be such that the instructor has to “go looking for you”. You should be attempting to get into the debate on a regular basis.

You might want to avoid being classified as one of the following types of students:

- Repeaters, i.e., students that, consciously or unconsciously, make comments that are really just repeats/rephrasing of what has already been said (by other students, or you). This wastes time and adds nothing to learning.
- Ramblers, i.e., students that take a lot of time to say simple things or they may tell long personal/professional stories, or they roam into topics that are not relevant, or simply make low-quality comments just to participate. They waste valuable time and prevent other students from being able to participate.
- Students that have been distracted (by Facebook, etc.) or who have stopped paying attention and then, later on, when they realized they have missed a term or concept, they ask you about it.

As per University Policy:

Each student has 4 chances to pass any given course distributed in two consecutive academic years (regular period and July period).

It is mandatory to attend 100% of the classes. Students who do not comply with at least 70% attendance will lose their 1st and 2nd chance, and go directly to the 3rd one (they will need to enroll again in this course the next academic year).

Grading for retakes will be subject to the following rules:

- Those students who failed the subject in the first regular period will have to do a retake in July (except those not complying with attendance rules who are banned from this possibility).
- Dates and location of the July retakes will be posted in advance and will not be changed. Please take this into consideration when planning your summer.
- The maximum grade that a student may obtain in the 2nd exam session is 8 out of 10. Those students in the 3rd call will be required to attend 50% of the classes. If due to schedule overlap, a different option will be discussed with the professor in order to pass the subject.

BIBLIOGRAPHY

Recommended

- Gene Kim, Jez Humble, Patrick Debois, and John Willis. *The DevOps Handbook: How to Create World-Class Agility, Reliability, and Security in Technology Orgs.* ISBN 978194278800 (Digital)
  https://ie.on.worldcat.org/oclc/1341896353
- Jez Humble and David Farley. *Continuous Delivery: Reliable Software Releases through Build, Test, and Deployment Automation.* ISBN 978032160191 (Digital)
  https://ie.on.worldcat.org/oclc/1100830377
- Julien Vehent. *Securing DevOps: Security in the Cloud.* ISBN 978161729413 (Digital)
  https://ie.on.worldcat.org/oclc/1105800614

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.

BAYESIAN STATISTICS

Bachelor in Data and Business Analytics BDBA SEP-2023
BS-DBA.1C.4.M.A

Area Others

Number of sessions: 15

Academic year: 23-24

Degree course: FOURTH

Number of credits: 3.0

Semester: 1º

Category: OPTIONAL

Language: English

Professor: **MANUELE LEONELLI**

E-mail: mleonelli@faculty.ie.edu

Manuele Leonelli is an Assistant Professor in the School of Human Sciences and Technology at IE University. He obtained a PhD in Statistics from the University of Warwick in 2015 under the supervision of Jim Q. Smith. He then won a CAPES post-doctoral fellowship working at the Federal University of Rio de Janeiro, Brazil, under the direction of Dani Gamerman. Before joining IE University, he was a Lecturer in Statistics in the School of Mathematics and Statistics at the University of Glasgow. Manuele's research focuses on probabilistic graphical models for decision-making under uncertainty and inference over extreme values, with a focus on approximated inferential algorithms within the Bayesian paradigm. His PhD thesis "Bayesian decision support in complex systems: an algebraic and graphical approach," won the John Copas Prize for the best PhD Thesis in Statistics at the University of Warwick in 2015.

mleonelli@faculty.ie.edu

PREREQUISITES

Basic knowledge of Calculus, Probability and Statistical Inference.

SUBJECT DESCRIPTION

We often use probabilities informally to express our information and beliefs about unknown quantities. However, the use of probabilities to express information can be made formal: In a precise mathematical sense, it can be shown that probabilities can numerically represent a set of rational beliefs, that there is a relationship between probability and information, and that Bayes' rule provides a rational method for updating beliefs in light of new information. The process of inductive learning via Bayes' rule is referred to as Bayesian Inference.

More generally, Bayesian methods are data analysis tools that are derived from the principles of Bayesian inference. In addition to their formal interpretation as a means of induction, Bayesian methods provide:

- parameter estimates with good statistical properties;
- parsimonious descriptions of observed data;
- predictions for missing data and forecasts of future data;
- a computational framework for model estimation, selection and validation.

Thus, the use of Bayesian methods goes beyond the formal task of induction for which the methods are derived. Throughout this module we will explore the broad uses of Bayesian methods for a variety of inferential and statistical tasks.

LEARNING OBJECTIVES

At the end of the course, you will be able to:

- Summarize the basic steps of a Bayesian inferential analysis.
- Perform prior to posterior updating of probabilities.
- Model real-world data with Bayesian methods.

TEACHING METHODOLOGY

Learning Activity	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	21.33 %	16.0 hours
Discussions	6.67 %	5.0 hours
Exercises in class, Asynchronous sessions, Field Work	18.67 %	14.0 hours
Group work	26.67 %	20.0 hours
Individual studying	26.67 %	20.0 hours
TOTAL	100.0 %	75.0 hours

PROGRAM

SESSION 1 (LIVE IN-PERSON)

Introduction to the course and to Bayesian statistics.

SESSION 2 (LIVE IN-PERSON)

The steps of Bayesian data analysis.

SESSION 3 (LIVE IN-PERSON)

A review of probability for Bayesian data analysis.

SESSION 4 (LIVE IN-PERSON)

Bayes rule and its use for inference: Estimating bias in binomial probabilities using a Bayesian approach.

SESSION 5 (LIVE IN-PERSON)

SESSION 6 (LIVE IN-PERSON)

Lab: Bayes rule in action.

SESSION 7 (LIVE IN-PERSON)

Inferring Binomial probabilities and the Beta distribution.

SESSION 8 (LIVE IN-PERSON)

The Beta posterior distribution and its use for inference.

SESSION 9 (LIVE IN-PERSON)

Lab: The Beta distribution in action.

SESSION 10 (LIVE IN-PERSON)

Random walks and their properties.

SESSION 11 (LIVE IN-PERSON)

The Metropolis algorithms.

SESSION 12 (LIVE IN-PERSON)

Lab: The Metropolis algorithm in action.

SESSION 13 (LIVE IN-PERSON)

The Gibbs algorithm.

SESSION 14 (LIVE IN-PERSON)

Accuracy and efficiency of Bayesian algorithms.

SESSION 15 (LIVE IN-PERSON)

Lab: The Gibbs algorithm in action.

EVALUATION CRITERIA

Your final grade in the course is based on a variety of activites. The weight of each one will be as follows:

criteria	percentage	Learning Objectives	Comments
Intermediate Tests	30 %		
Workgroups	50 %		
Class Participation	20 %		

RE-SIT / RE-TAKE POLICY

Class Participation (20%)

Your grade will be based on: active engagement in synchronous and asynchronous classes. For asynchronous activities, participation will be based on a within-group peer-evaluation of your commitment and contribution during labs.

Lab Reports (80%)

In small groups you will produce short reports for each of the labs, showing the code you developed and the results of your analysis.

Late Assignments/Presentation

Will be penalized 10% per 24-hour period, starting on the day they are due. Only in cases of emergency or illness can changes be made to due dates of assignments or projects. ALL such arrangements are the full responsibility of the student and must be made PRIOR to the due date. Failure to confirm any changes to the due date with the professor prior to the due date will result in a grade of zero.

Minimum passing grade

The overall passing course grade is 5.0.

BIBLIOGRAPHY

Recommended

- Kruschke, J.K. *Doing Bayesian Data Analysis.* ISBN 9780124059160 (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.

CLINICAL DATA AND LAW

Bachelor in Data and Business Analytics BDBA SEP-2023
CDL-DBA.1C.4.M.A

Area Others

Number of sessions: 15

Academic year: 23-24

Degree course: FOURTH

Number of credits: 3.0

Semester: 1º

Category: OPTIONAL

Language: English

Professor: **MARCOS GALLEGO LLORENTE**

E-mail: mgallegol@faculty.ie.edu

Marcos is currently a senior consultant at Vintura, a healthcare strategy boutique.

Marcos brings extensive experience working on the technology, business and policy sides of healthcare. He holds a PhD from the Universiity of Cambridge, has worked as a consultant at Cambridge Innovation Consulting and Vintura, has collaborated with the OECD, at the Science, Technology and Innovation Directorate on the Future of AI in Healthcare, and also works as research lead at the IE School of Global and Public Affairs CGC program on the Future of Healthcare.

He holds 6+ years of experience at the interface between science, technology, healthcare, business and policymaking, helping companies and governments identify technology-enabled opportunities, exploit emerging market gaps, anticipate the future needs of different skeholders, and develop the right use cases for developing innovative products and services.

mgallegol@faculty.ie.edu or +34 684168105.

PREREQUISITES

No format prerequisits to enter this course.

SUBJECT DESCRIPTION

This course explores the intersection of clinical data and law. Through a series of interactive discussions and case studies, students will learn how to navigate the complex ethical, legal, and social issues surrounding the collection, analysis, and use of clinical data.

LEARNING OBJECTIVES

Course Goals:

- To develop a critical understanding of the ethical, legal, and social issues surrounding the collection, analysis, and use of clinical data.
- To provide students with the skills and knowledge necessary to navigate the complex legal and regulatory landscape surrounding clinical data.
- To encourage students to engage in critical thinking and analysis of the ethical and social implications of the use of clinical data in healthcare.

Full bibliography and reading lists will be included at the beginning of the course.

TEACHING METHODOLOGY

Learning Activity	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	20.0 %	15.0 hours
Discussions	6.67 %	5.0 hours
Exercises in class, Asynchronous sessions, Field Work	8.0 %	6.0 hours
Group work	18.67 %	14.0 hours
Individual studying	46.67 %	35.0 hours
TOTAL	100.0 %	75.0 hours

PROGRAM

SESSION 1 (LIVE IN-PERSON)

Introduction to Clinical Data and Law.

This session provides an overview of the course and introduces students to the key ethical, legal, and social issues surrounding the collection, analysis, and use of clinical data.

Article: Evolving regulatory perspectives on digital health technologies for medicinal product development (Seya Colloud et al. Digit. Med. 6, 56; 2023) (NPJ)

SESSION 2 (LIVE IN-PERSON)

The The Legal Framework & Ethics of Clinical Data Collection

This session explores the ethical considerations involved in the collection of clinical data, including issues related to privacy, consent, and confidentiality. This session also examines the legal framework surrounding the collection of clinical data, including the HIPAA Privacy Rule, the Common Rule, and state laws governing data collection.

SESSION 3 (LIVE IN-PERSON)

[Guest lecture]

This session will bring an external speaker to showcase the in-hospital challenges for data collection in terms of regulatory, ethics, and feasibility. It will wrap-up with success examples of best practices for data collection and its usefulness for research and model training.

SESSION 4 (LIVE IN-PERSON)

Using Clinical Data: Research and Public Health

This session explores the use of clinical data in research, including issues related to informed consent, risk, and benefit. The session will also examine the use of clinical data in public health, including issues related to surveillance, outbreak investigation, and emergency response.

SESSION 5 (LIVE IN-PERSON)

Clinical Data and Quality Improvement

This session explores the use of clinical data in quality improvement initiatives, including issues related to data analysis, benchmarking, performance measurement, as well as including issues related to interoperability, data security, and data sharing.

SESSION 6 (LIVE IN-PERSON)

Clinical Data and Precision Medicine

This session explores the use of clinical data in precision medicine, including issues related to data analysis, risk stratification, and patient-centered care.

SESSION 7 (LIVE IN-PERSON)

Clinical Data in Artificial Intelligence

This session examines the use of artificial intelligence in the collection, analysis, and use of clinical data, including issues related to algorithmic bias, transparency, and accountability.

SESSION 8 (LIVE IN-PERSON)

Clinical Data and Genomics

This session explores the use of clinical data in genomics, including issues related to privacy, informed consent, and genetic discrimination.

SESSION 9 (LIVE IN-PERSON)

Clinical Data and Medical Device Regulation

This session examines the regulation of medical devices that use clinical data, including issues related to safety, efficacy, and post-market surveillance.

SESSION 10 (LIVE IN-PERSON)

Clinical Data and Drug Development

This session explores the use of clinical data in drug development, including issues related to clinical trial design, safety monitoring, and regulatory approval.

SESSION 11 (LIVE IN-PERSON)

[Guest speaker]

This session will bring in a guest speaker to reflect on the uses of data in the settings discussed in the previous lectures: Personalised Health, AI, and Genomics.

SESSION 12 (LIVE IN-PERSON)

Future Directions for Regulation of Clinical Data

This session provides a summary of the course and encourages students to reflect on the ethical, legal, and social implications of the use of clinical data in healthcare.

SESSIONS 13 - 14 (LIVE IN-PERSON)

Student presentations

SESSION 15 (LIVE IN-PERSON)

Exam

EVALUATION CRITERIA

Evaluation will include:

- Coursework 1: Secondary research overview. Students will research a policy topic and identify the latest regulation developments in different geographies (EU, USA, and others)
- Coursework 2: Policy Proposal: Students will research and develop an op-ed with a a policy proposal addressing a specific issue in clinical data and law. They will identify gaps or challenges in existing policies, analyze their implications, and propose recommendations for improving the legal and ethical framework surrounding clinical data collection, use, and protection.
- Group work: Each group can research and prepare a presentation on their assigned topic (tbd). Students will analyze real-world cases involving the collection, analysis, and use of clinical data, and identify the key ethical, legal, and social issues involved.
- Written exam. Mocks will be provided.

criteria	**percentage**	**Learning Objectives**	**Comments**
Final Exam	50 %		Full overview of the course
Coursework 2	15 %		Op-ed writing, compelling policy-making, issue analysis
Coursework 1	15 %		Secondary research and policy reviews
Group Presentation	10 %		Case study analysis, best practice sharing
Class Participation	10 %		Participation in discussions

RE-SIT / RE-TAKE POLICY

All students who fail the course will be given a second chance as a re-sit.

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.

DATA GOVERNANCE

Bachelor in Data and Business Analytics BDBA SEP-2023
DG-DBA.4.M.A

Area Others

Number of sessions: 15

Academic year: 23-24

Degree course: FOURTH

Number of credits: 3.0

Semester: 2º

Category: COMPULSORY

Language: English

Professor: **ALFONSO FERNÁNDEZ REVENGA**

E-mail: alfonsofernandezr@faculty.ie.edu

In over 15 years of experience in the discipline of data architectures, I have played several roles in the field of Data Management and Big Data. During my time in KPMG and EY, I helped to build and consolidate the corporete Data Lake environment as the corporate decision-making system and as a manager of Big Data architecture. In 2017 I joined as Product Owner and Head of Data Governance at Stratio, where we built a full Augmented Data Fabric producto that covers the whole lifecycle of data management: Auto-discovering the data, virtualizing it, providing it with metadata and intelligently giving it semantic meaning and, finally, mining it with MLOps. All with a unique platform that ensures the best integration among the different parts.

Data Management and Data Governance expert, based on DAMA Framework (DMBOK2).

Specialized in project management with agile methodologies (scrum and kanban).

Office Hours

Office hours will be on request. Please contact at:

On demand, via zoom mmeting

Before or after the classes

Professor: **SILVINA ARCE GIL**

E-mail: sarce@faculty.ie.edu

Experienced Chief Data Officer, establishing a robust and integral Data Strategy & Governance model, with a demonstrated history of working in the banking&retail industries. She has been involved in Banking transformations for more than 20 years, projects related to Core Banking and Digital transformation. She has been leading this kind of projects in different financial institutions worldwide.

Office Hours

Office hours will be on request. Please contact at:

On demand, via zoom mmeting

Before or after the classes

SUBJECT DESCRIPTION

In 2017, The economist publishes “The world’s most valuable resource is no longer oil, but data”. In 2021, a report from the World Economic Forum states that “data now constitute a new kind of economic asset, such as cash or gold “.

Since then, all organizations have tried to extract the value from this strategic asset to take better decisions regardless the size of the sector or their business purpose. Data, instead of intuition, will lead them to the right decisions. This is what we call Data Driven Companies.

The key driver to the adoption of data and intelligence to 100% in the organizations is to have an efficient management of the data in its life cycle. So what do the organizations do to manage such an important asset? A great amount of them invest to increase their capabilities in one or various of the stages of the data life cycle (collection, integration, storage, exploitation, etc.) rather than undertaking the issue under a holistic point of view.

Data Management and governance are the main drivers to ensure data is used for decision-making. While Data Governance establishes policies, procedures, roles, and responsibilities around data, Data Management applies those policies and procedures along the whole life cycle of data.

During this course, terms as Data Governance, Data Quality, Metadata Managment, Data Strategy, Security, etc. and the relationships between them will be explained to get a full knowledge of data Managment.

LEARNING OBJECTIVES

The objective of the course is to provide students with necessary skills that wolud allow them to efficiently manage the life cycle of an organization's data.
At the end of the course, students will be able to:

- Apply Data Management as a whole, using one of the most recognized frameworks in the sector
- Know through real cases, the best practices carried out in organizations that allow obtaining value from their data assets
- Understand the best practices in different disciplines related to data management
- Avoid the most common mistakes and correctly apply the knowledge in this matter

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:

Learning Activity	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	26.67 %	20.0 hours
Discussions	6.67 %	5.0 hours
Group work	26.67 %	20.0 hours
Individual studying	40.0 %	30.0 hours
TOTAL	100.0 %	75.0 hours

PROGRAM

SESSION 1 (LIVE IN-PERSON)

Introduction to Data Managment & Governance:

- Presentation
- Data as an asset
- Key Business Questions and KPIs

SESSION 2 (LIVE IN-PERSON)

Data Managment fundamentals:

- Introduction to Data Managment and fundamentals
- Essential concepts
- Data Managment frameworks

SESSION 3 (LIVE IN-PERSON)

Data governance:

- Business drivers
- Goals and principles
- Activities
- Implementation guidelines
- Metrics

SESSION 4 (LIVE IN-PERSON)

Data Strategy

- Introduction

- Business drivers
- Data maturity level with dataMat
- Data Strategy with dataToolkit

SESSION 5 (LIVE IN-PERSON)

Data Arquitecture:

- Business drivers
- Activities, tools and techniques
- Data architecture governance: metrics

SESSION 6 (LIVE IN-PERSON)

Data maturity assesmet and Data Strategy

A group research project in which students need to make a data matrity assesment using dataMat, and make a Data Strategy proposal using dataToolkit.

Reference material (Club CDO):

- dataMat
- dataToolkit

SESSION 7 (LIVE IN-PERSON)

Data Modeling & Design:

- Business drivers
- Goals and principles
- Activities and tools
- Best practices in database design

SESSION 8 (LIVE IN-PERSON)

Data Warehousing & Business Intelligence:

- Business drivers
- Goals and principles
- Essential concepts
- Activities, tools and techniques
- Implementation guidelines
- DW/BI Governance

SESSION 9 (LIVE IN-PERSON)

Designing a Multidimensional model

Students will have to design a Star Schema or a Snowflake Schema for a given business model

SESSION 10 (LIVE IN-PERSON)

Data integration and Interoperability

- Business drivers

- Goals and principles
- Activities and tools
- Best practices in database design

Reference material:

SESSION 11 (LIVE IN-PERSON)

Data Security and Ethics :

- Business drivers
- Goals and principles
- Essentials concepts
- Activities, tools and techniques
- Implementation guidelines

SESSION 12 (LIVE IN-PERSON)

Metadata Managment:

- Business drivers
- Goals and principles
- Essentials concepts
- Activities, tools and techniques
- Implementation guidelines
- Metadata Governance

SESSION 13 (LIVE IN-PERSON)

Data Quality:

- Business drivers
- Goals and principles
- Essentials concepts
- Activities, tools and techniques
- Implementation guidelines
- Data Quality and Data Governance

SESSION 14 (LIVE IN-PERSON)

Setting up the best Data Quality framework

In this session, students have to deploy a Data Quality Solution for the data model created in session number seven

SESSION 15 (LIVE IN-PERSON)

Final exam

EVALUATION CRITERIA

Your final grade in the course will be based on both individual and group work of different characteristics that will be weighted in the following way:

criteria	percentage	Learning Objectives	Comments
Final Exam	40 %		
Workgroups	15 %		Each group will send the proposal to the teacher within a maximum of seven days from the session #4
Individual Work	15 %		Each student must send the desing up to seven days from the finalization of session #7
Workgroups	15 %		Each group will send the proposal to the teacher within a maximum of seven days from the session #13
Class Participation	15 %		

RE-SIT / RE-TAKE POLICY

BIBLIOGRAPHY

Compulsory

- Deborah Henderson, Susan Early, Laura Sebastian-Coleman, elena Sykora, Eva Smith. *DAMA - DMBOK.* Second edition. Techincs publications. BASKING RIDGE, NEW JERSEY. ISBN 1634622340 (Digital)

This guide provides information on data governance, data architecture, data development, database operations, data security, reference and master data, data warehousing and business intelligence, document and content management, meta data management, data quality and professional development. DAMA-DMBOK2 provides data management and IT professionals, executives, knowledge workers, educators, and researchers with a framework to manage their data and mature their information infrastructure

Recommended

- Robert S. Seiner. *NON-INVASIVE DATA GOVERNANCE.* ISBN 9781935504856 (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.

DATA IN THE HEALTHCARE INDUSTRY

Bachelor in Data and Business Analytics BDBA SEP-2023
DHI-DBA.1C.4.M.A

Area Others
Number of sessions: 15
Academic year: 23-24
Degree course: FOURTH
Number of credits: 3.0
Semester: 1º
Category: OPTIONAL
Language: English

Professor: **DAVID GONZALEZ PISANO**

E-mail: dgonzalezp@faculty.ie.edu

David G Pisano holds a MSc in Applied Statistic, a MSc in Bioinformatics PhD and a BSc in Fundamental (Cellular and Molecular) Biology from Oviedo University. Research interests include integrative and multimodal approaches to improve processes in the healthcare industry, with special focus on the usage of omics data in diagnostics, patient segmentation and improvement of health outcomes. He has published 114 papers in peer-reviewed scientific journals, with more than 11,000 acumulated citations and h-index 46. Lecturer in genomics, bioinformatics, biomedical data handling and advance analytics in the healthcare industry. Previous jobs include ISP owner and CTO, process and technology consultant, project manager and service delivery manager, CTO, lab director, and head of healthcare analytics. He is currently the Clinical Bionformatics Director at Genomocore.

David is an accomplished senior level manager with solid background in technology consulting, professional service delivery, cancer genomics research, and data science/advanced analytics within insurance/healthcare sector. He has in-depth knowledge of processes and requisites related with healthcare, pharmaceutical, manufacturing, veterinary, in-vitro diagnostics, biotechnology, R&D, cancer research, and high performance computing. Well-versed in devising strategies for the achievement of short and long-term organisational objectives. Articulate communicator and speaker with proven track record or fostering strong relations with multi-disciplinary functions, stakeholders, and C-suite executives. Skilled problem-solver, supporting the achievement of individual and shared goals.

CORPORATE EXPERIENCE

- Current position: Clinical Bioinformatics Director, Made of Genes (Barcelona: 2020-present)
- Health Analytics Head, Sanitas (Madrid: 2016-2019)
- Head of Bioinformatics, Spanish National Cancer Research Centre (CNIO) (Madrid: 2006-2016)
- CTO, National Bioinformatics Institute (Madrid: 2004-2009)
- Head Bioinformatician, GENOMICA (Zeltia) (Madrid: 2001-2003)
- Service Delivery Manager for Iberia and Latin America, Thermo Fisher Scientific (Madrid: 2000-2001)

- Project Manager, Thermo Fisher Scientific (Madrid: 1999)
- Project Consultant, Thermo Fisher Scientific (Madrid: 1998)
- Founder & CTO, CSSC (Oviedo: 1995-1997)

ACADEMIC BACKGROUND

- Accelerate: Building Business from Science & Technology, IE Business School
- MSc Statistics, UNED
- MSc Bioinformatics, GNA
- MSc Biochemistry, Univ. Oviedo
- BSc Molecular Biology, Univ. Oviedo
- Corporate Experience

ACADEMIC EXPERIENCE

- Master in Business Analytics & Big Data (IE University)
- Master in Research Methodology in Molecular Biosciences (Univ. Autonoma Madrid)
- Master in Bioinformatics (Univ. Complutense de Madrid, National Health School)
- Master in Health Biotechnology, Technology and Biotechnology Management (CESIF, Madrid)
- Summer School in Bioinformatics (Univ. Complutense, Madrid)
- Biotechnology Degree (Univ. Francisco de Vitoria, Madrid)
- Biology Degree (IE University, Segovia)
- Bichemistry Degree (Univ. Oviedo)

dgonzalezp@faculty.ie.edu

PREREQUISITES

SUBJECT DESCRIPTION

This is an introductory course for data in the healthcare industry, from the perspective of data analysts. Through this course, you will gain highly valuable skills in the healthcare sector by understanding how the healthcare system records information about each patient and medical encounter. You will also learn the specific features of health data that enable you to perform more insightful analyses. You will also be able to communicate more effectively with clinical and analytic colleagues. You will be empowered to improve care processes and make a difference to many people's health and lives.

LEARNING OBJECTIVES

- To gain an understanding of healthcare data as fundamental basis for health analytics
- To explore the generation and usage of data in the healthcare industry
- To recognize the kind of datasets produced in the health system, their sources, the tools to manage and explore them and understand them, and their limits and applications in distinct scenarios

- To understand –omics and other big data technologies, and to realise that they drive the profound changes to be experienced by biomedical data sciences

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:

Learning Activity	**Weighting**	**Estimated time a student should dedicate to prepare for and participate in**
Lectures	6.67 %	5.0 hours
Discussions	6.67 %	5.0 hours
Exercises in class, Asynchronous sessions, Field Work	26.67 %	20.0 hours
Group work	33.33 %	25.0 hours
Individual studying	26.67 %	20.0 hours
TOTAL	100.0 %	75.0 hours

PROGRAM

SESSION 1 (LIVE IN-PERSON)

Introduction. Historical perspective. Health vs Healthcare Data. Characteristics and Challenges

SESSION 2 (LIVE IN-PERSON)

Healthcare Data Standards: Standardized vocabularies, controlled vocabularies, terminologies, and ontologies. CPT, ICD, LOINC, RxNORM, SNOMED, UMLSs... Data interchange and representation

SESSION 3 (LIVE IN-PERSON)

EHR Data + Case Study

SESSION 4 (LIVE IN-PERSON)

Health Insurance Claims Data + Case Study

SESSION 5 (LIVE IN-PERSON)

Laboratory Data + Individual Task (blood test redesign)

SESSION 6 (LIVE IN-PERSON)

Unstructured Data: medical images, medical notes, signals…

SESSION 7 (LIVE IN-PERSON)

Genomics and other -omic data + Hands-on: Genomics Data Analysis

SESSION 8 (LIVE IN-PERSON)

Clinical Trials Data

SESSION 9 (LIVE IN-PERSON)

Biomedical research and open data. RWD, Wearables and other data sources

SESSION 10 (LIVE IN-PERSON)

Healthcare Data Models: operations, supported use cases, normalization, integration, data quality

SESSION 11 (LIVE IN-PERSON)

Dataviz in Healthcare: monitors & dashboards

SESSION 12 (LIVE IN-PERSON)

Hands-on: Dataviz in Healthcare

SESSIONS 13 - 14 (LIVE IN-PERSON)

Final Group Presentations

SESSION 15 (LIVE IN-PERSON)

Exam

EVALUATION CRITERIA

criteria	percentage	Learning Objectives	Comments
Final Exam	40 %		
Group Presentation	30 %		
Individual work	20 %		
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 70% 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 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

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.

DEEP REINFORCEMENT LEARNING

Bachelor in Data and Business Analytics BDBA SEP-2023
DRL-DBA.1C.4.M.A

Area Others

Number of sessions: 25

Academic year: 23-24

Degree course: FOURTH

Number of credits: 5.0

Semester: 1º

Category: OPTIONAL

Language: English

Professor: **JOSÉ MANUEL REY GONZÁLEZ**

E-mail: jreyg@faculty.ie.edu

Consultant Director of numerous projects of strategic planning and business development, market studies and quality improvement, revenue growth and cost optimization, business processes reengineering, information systems design and implementation, for companies in virtually all industries, in both public and private sectors, with a systemic approach based on the identification of innovative competitive strategies, efficient organizational and process transformation and a firm leverage on the impulse of new technologies

Education

He is a Civil Engineer graduated with honors, special End Of Career Award, Bachelor of Economics and Business Administration, PhD student (unfinished) in Applied Quantitative Economics, and holds a CISA Title from the EDPAA (now ISACA) and a Master in Corporate Finance IESE, University of Navarra.

Professional Background

He became International Partner in Arthur Andersen in 1997, led the area of Innovation and Business Transformation, responsible for comprehensive quality management services, shared services centers, process reengineering and technological integration and the eBusiness area for strategic planning and development of interactive business and electronic commerce nationwide.

International Director of the World Excellence Center for Business Process Reengineering, member of the eBusiness leadership team in EMEIA and coordinator of this practice in the Mediterranean region.

He was a Managing Director in KPMG Consulting/BearingPoint, responsible for the Technology industry within the Telecommunications, Media and Content Sector and leader of the area of Process Innovation Solutions and Advanced Technology

He has served as Member of Board of Directors, Advisory boards or Management Consultant in matters of strategy and business development for several Spanish companies and Advisor of Venture Capital Companies and M&A Firms.

Arbitrator of the Madrid Arbitration Court, specialist in the Technology, Information and Communications sector and Developer of an advanced software platform for modelling business risks and economic projections, valuation of companies & portfolios and for building advanced risk & value information systems for Directors.

Public Activity and Teaching

He has been part of the Advisory Committees of eMobility and SIMO. He was a member of the Jury of the National eMobility Prize and participated in the Competitiveness Forums of Madrid and Castilla y León.

Instructor and speaker in numerous courses and seminars, national and international, of a general nature and related to his areas of specialization. Lecturer in various Business Management Master courses in subjects related to advanced economic and statistical analysis as well as organizer of the Master eBusiness of the Universidad Pontificia de Comillas, Speaker in the Executive Education Program of ESADE and in the Executive MBA of AEDE Business School in the subjects of "Information Systems", "Technological Environment" and "Systems Architecture".

He has published articles on business and technical content topics in newspapers such as Expansión, Cinco Días, Gaceta de los Negocios, Actualidad Económica, Dinero, etc., and directed or participated, as author or coordinator, in the publication of various books on issues of quality, organization and information technology (Las empresas del click / The Clickable Corporation [1999, Actualidad Económica], La Calidad en España/Quality in Spain [1995, Cinco Días]) and collaborated in the development of others (The organization in the information age: Learning, Innovation and Change [1995, IESE])

He was listed in the year 2000 by Actualidad Económica in the ranking of the 100 most relevant people on the Internet and eBusiness in Spain.

E-Mail: jreyg@faculty.ie.edu

PREREQUISITES

To optimize your chances of success in this program, we recommend having a good Python programming knowledge with at least 120hrs of programming experience, familiarity with data structures and experience with libraries. This is especially important for the assignments, and for getting the most out of the F2F practice/tutorials along the course. Additionally, the use of some standard Machine Learning libraries such as scikit-learn, Tensorflow and Pytorch will be required. The new environment programming and tools (such as OpenAI gym) will be introduced and covered along the course.

In a few sections, some of the theoretical concepts will be presented with quite a bit of rigor and a good college-level knowledge of linear algebra, calculus and basic inferential statistics would be desirable for a thorough understanding. For example, it is expected that you know about standard probability distributions (Uniform, Gaussians, Poisson…), and also understand the operation and interpretation of derivatives and the basic mathematics underlying probability models. Some concepts about dynamic programming and markov decision processes will be revisited.

SUBJECT DESCRIPTION

Reinforcement Learning (RL) is a melting pot where excellent computational techniques of dynamic programming, heuristic search in problem spaces and the new techniques of functional representation of neural networks have converged with great success to build a highly scalable area of discipline that is allowing to efficiently attack a wide range of problems previously considered as practically unapproachable.

Throughout the course we will analyze each of these components in a structured way and we will see the types of approach combinations currently in use, from basic Q-Value models, to Deep QL and methods based on policy gradients, actor-critic, curiosity driven and other innovations in the dynamic definition of reward functions.

We will try to understand in which problem scenarios they are most advisable, their strengths and limitations and to anticipate some of the future trends that will progressively lead to advances and breakthroughs in this field.

LEARNING OBJECTIVES

The main objective of this course is to introduce students to the exciting field of RL and to set up a framework of knowledge that will allow them to make informed analysis of the opportunities and challenges for its successful application in Business.

To achieve this, we will focus on helping students grasp a sound understanding of RL techniques and their variations (model-free, model-based, online/offline RL, behavioural cloning, etc) and a good intuition of how to apply them to tasks of their own design.

It encompasses knowledge about:

- the fundamentals and theoretical principles of this discipline,
- the most relevant algorithms used in practice,
- the engineering process for the development and management of this type of projects and
- the understanding of the difficulties, obstacles and challenges to be faced in this path.

Therefore, it will combine both a theoretical and conceptual approach with a hands-on technical understanding of the different stages of these type of RL projects keeping always a clear perspective of the business issues and opportunities and the future evolution and innovations from the current state of the art.

To streamline the technical/computational requirements, we will concentrate on the implementations using Python, and a level of complexity ranging from basic to intermediate.

For some advanced examples we could also touch on some other development environments (such as LUA for robotics or Matlab for process-control) but just for showcase purposes. No programming activities will be required in these cases.

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 program will try to put you in a good starting point for becoming a well-rounded RL practitioner, experiencing how to build RL models and apply them to problems in various fields.

We have tried to organize the different knowledge blocks integrating the conceptual and theoretical background first, then moving into practice/tutorials examples and then framing assignments of assorted levels of complexity.

We expect students to work 3 to 8 hours/week on average. This is a rough estimation of total hours the average student may take to complete all required coursework, including lecture and project time (actual hours may vary depending on the individual student and each specific stage of the course).

The course will be both lecture and example-based, and will include group in-class discussions to promote learning and understanding of the course material in a variety of formats.

The course will have 6 main elements:

- **Lectures**: We will explain the theoretical ideas, concepts and methods involved and will try to check on-time the correct understanding of the key elements. Questions and feedback will be the basic tool for this interactive dialog, in order to guarantee that the message gets across.
- **Examples/Tutorials/Cases:** We will use profusely cases and examples related to the theory (and preparatory for the Assignments). Questions and what-if interactive analysis will be encouraged.
- **Discussions:** as one critical skill for your future is how you present your work, we will encourage from time to time, for the relevant concepts in each section, some group discussions. A few of them will be announced in advance and will require some preparation from each student.
- **Assignments:** practical exercises for experimenting with RL concepts, algorithms and techniques.
- **Exams:** There will just one formal test/exam, and a “test-exam” assignment will be included to get students used to answer questions about RL in different types of formats.
- **Group work:** For the final part of the course the student will have to work in a small group for preparing a final project presentation.

Assignments:

The best way to learn about a machine learning method is to program it yourself and experiment with it. So, the assignments will generally involve implementing machine learning algorithms, and experimentation to test your algorithms on some data. You will be asked to summarize your work, and analyze the results.

The implementations will be done in Python, creating Jupyter notebooks or pure Python programs (in some cases, numba or cython could be required for performance optimization), and the structure and review criteria will be reported for each specific job.

Though they are considered as individual work, the "individual assignments" will be realized (and delivered jointly) in small groups that will be defined by an special algorithm that will dynamically form the groups for each assignment ensuring that no two students will be assigned more than once together and trying to optimize and combine the skills of each one based on their previous results to maximize cross-learning, Therefore, collaboration on the assignments is only allowed (and enforced) within each group. Each student/group is responsible for his or her own work.

Discussion of assignments and programs should be limited to clarification of the handout itself, and should not involve any sharing of pseudocode or code or simulation results.

The only exception will be for the Final Group Assignment, for which another specific algorithm will try to create optimized groups based on the preferences informed by each student.

The schedule of assignments is included in the syllabus and in general will correspond to the sessions planned as "with assignment" except for exceptions required for the best development of the content.

Assignments are due at the beginning of class/tutorial on the due date. Because they may be discussed in class that day, it is important that you have completed them by that day.

Optionally, some of the assignments could have the form of a “bake-off” (a competition between machine learning algorithms). We will give everyone some data for training a RL system, and you will try to develop the best method. We will then determine which system performs best on some unseen test data.

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 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.

The questions for the final test/exam will be selected from the ones used for the practice.

Participation:

We expect students to attend all lectures, and all practice/tutorials. Attention, interaction and participation during these lectures is especially important because we will cover material in class that is not included in the reference readings. Also, the tutorials will not only be for review, practicing and answering questions, but new material and concepts will also be covered or explained.

The oreintative expected distribution of times for these different types of learning activities is such as the following:

Learning Activity	**Weighting**	**Estimated time a student should dedicate to prepare for and participate in**
Lectures	16.0 %	20.0 hours
Discussions	8.0 %	10.0 hours
Exercises in class, Asynchronous sessions, Field Work	52.0 %	65.0 hours
Group work	12.0 %	15.0 hours
Individual studying	12.0 %	15.0 hours
TOTAL	100.0 %	125.0 hours

PROGRAM

SESSION 1 (LIVE IN-PERSON)

INTRODUCTION TO THE COURSE (SYLLABUS)

Procedural Learning and RL motivation. Overview of contents and analysis of the approaches and the relevant issues in this field.

SESSION 2 (LIVE IN-PERSON)

RL INTRODUCTION

Building blocks of current RL algorithms and approaches.

Different convergent routes to current RL (from MPC, Markov Chains, Genetic algorithms, etc...).

RL BASIC CONCEPTS

Definition of intelligent agents and their types. Evaluative vs Instructive feedback. Associative vs Non-associative models.

Book Chapters: Reinforcement Learning: An Introduction; Chapter 1, Section 1.7 Early history of Reinforcement Learning Chapter 1, Section 1.3 Elements of Reinforcement Learning (See Bibliography)

SESSION 3 (LIVE IN-PERSON)

MULTI-ARMED BANDITS

Introduction to contextual bandits and its understanding as a basis for full RL.

Interesting "low-hanging fruits" practical examples for its direct application in some type of business problems.

Book Chapters: Reinforcement Learning: An Introduction; Chapter 2, Tabular Solution Methods (See Bibliography)

SESSION 4 (LIVE IN-PERSON)

GRAPHS, SEARCH, MDP's

The understanding of the search space and its representation and the possible dimensions of heuristic approaches for search, prediction and control. Introduction to Markov Decision Processes.

Review and apply concepts about MDP's.

Practical cases of MDP's. Intuition and analytical solutions (for simple cases). Understanding of techniques for its representation (backup diagrams).

DYNAMIC PROGRAMMING, CSP's, GAME THEORY

The iterative value function and the Generalized Policy Improvement techniques.

Book Chapters: Reinforcement Learning: An Introduction; Chapter 3, Tabular Solution Methods, Finite Markov Decision Processes Chapter 4, Tabular Solution Methods, Dynamic Programming (See Bibliography)

SESSION 5 (LIVE IN-PERSON)

MONTE CARLO METHODS [Assignment: WARM-UP ASSIGNMENT]

Explanation of Monte Carlo methods and the mathematical foundations behind the (extremely versatile in ML) Importance Sampling technique. The merge of these approaches with the TD ones: n-step bootstrapping.

Book Chapters: Reinforcement Learning: An Introduction; Chapter 5, Mote Carlo Methods, Importance Sampling (See Bibliography)

SESSION 6 (LIVE IN-PERSON)

TEMPORAL DIFFERENCES

The integration of the previous different concepts in the current cornerstone of RL (for finite-MDP-like problems). Bellman Optimality Equation.

Book Chapters: Reinforcement Learning: An Introduction; Chapter 6, Temporal-Difference Learning Chapter 7, n-step Bootstrapping (See Bibliography)

SESSION 7 (LIVE IN-PERSON)

A FIRST PRACTICE - INTEGRATION OF CONCEPTS

The Snake Game. Application of the principles studied to an initial problem. Comparison with the other possible solutions (harcoded, genetic algorithm, hamiltonian paths, etc). Live comparison and charting of all of them and discussion of pros and cons. Rationale behind the use of RL approaches.

SESSION 8 (LIVE IN-PERSON)

LEARNING AND PLANNING

A deeper comparative study of the characteristics and differences of model-based vs model-free methods.

Book Chapters: Reinforcement Learning: An Introduction; Chapter 8, Planning & Learning with tabular methods (See Bibliography)

SESSION 9 (LIVE IN-PERSON)

EVOLUTIONARY COMPUTATION [Assignment: ASSIGNMENT]

Genetic algorithms and beyond, Non-dominated multi-objective optimization, the Pareto frontier and NEAT.

Exploration of a sometimes considered independent field (genetic algorihtms and other non-derivative-based methods). Practical examples of their use in practical problems such as multi-objective optimization. Analysis of the non-dominated genetic algorithms (introduction to the concept of Pareto frontier). Exploration of Neuro-Evolution techniques and the current and future lines of research for connecting these types of solutions with the mainstream RL approaches.

SESSION 10 (LIVE IN-PERSON)

RL FRAMEWORKS

Overview of main RL Frameworks.

Comparison of some currently used RL frameworks (from OpenAI gym, stable baselines, etc to Ray and RLib). Pros, cons and requisites for their use.

SESSION 11 (LIVE IN-PERSON)

INTERMEDIATE PRACTICE

Analysis and comparison of different algorithmic methods for the training of these types of models.

SESSION 12 (LIVE IN-PERSON)

APPROXIMATION METHODS

DEEP-Q NETWORKS

The key change in the strategy: the calculus of approximative functions (for value, q(s,a) or policy functions). The main advantages of this approach. Focus in one of the main (most relevant in practice) approaches: ANN function approximators.

Book Chapters: Reinforcement Learning: An Introduction; Chapter 9, On-policy prediction with approximation. Section 9.7 Nonlinear Function Approximation: Artificial Neural Networks Chapter 10, Chapter 11 and Chapter 12 (Eligibility traces) (See Bibliography)

SESSION 13 (LIVE IN-PERSON)

POLICY GRADIENTS [Assignment: ASSIGNMENT]

Theory behind PGM's and the "Policy Gradient Theorem".

Book Chapters: Reinforcement Learning: An Introduction; Chapter 13: Policy Gradient Methods (See Bibliography)

SESSION 14 (LIVE IN-PERSON)

ACTOR-CRITIC

Definition and analysis of the actor-critic methods. The concept of "advantage". Main algorithms.

Book Chapters: Reinforcement Learning: An Introduction; Chapter 13, Section 13.5 Actor-critic methods (See Bibliography)

SESSION 15 (LIVE IN-PERSON)

PRACTICE REVIEW

Dissection and discussion of sample solutions. Exploration of main difficulties.

SESSION 16 (LIVE IN-PERSON)

PROBLEM APPROACH FRAMEWORK

Exhaustive analysis of a classical problem (The Travelling Salesman Problem) under all possible lines of solutions (from Christofides to genetic algorithm, ant-colony optimization and simple or complex Reinforcement Learning) Discussion of pros and cons and initial understanding of "When, Why and How" of their possible uses.

SESSION 17 (LIVE IN-PERSON)

DEEP RL PRACTICE [Assignment: ASSIGNMENT]

Connecting the dots between RL and approximation (ANN) functions. Application to a more complex learning problem. Comparison of results.

SESSION 18 (LIVE IN-PERSON)

SUMMARY RECAP

Recap session for all the concepts and algorithms introduced so far. Introductory extension to some new concepts (Rainbow algorithms, etc).

SESSION 19 (LIVE IN-PERSON)

RESEARCH PAPERS, INNOVATION TRENDS [Assignment: EXAMINATION PRACTICE. BE PREPARED TO ANSWER QUESTIONS ABOUT RL]

Study and discussion of some of the most relevant papers published in the field. This will include the new developments in Parametric Model Based RL, Off-line RL, Inverse RL, introduction to the novel field of IRL (learning an agent’s objectives, values, or rewards by observing its behavior). Connection to related concepts of Behavioural Cloning, Imitation Learning and Teacher Agents, etc and a summary of new approaches (reward shaping, finite automata rewards, hierarchical RL...)

Analysis of obstacles and challenges in the field and in its practical application to business problems.

Book Chapters: Reinforcement Learning: An Introduction; Chapter 17: Frontiers (See Bibliography)

SESSION 20 (LIVE IN-PERSON)

MULTI-AGENT RL

Incursion in the multi-agent reinforcement learning (MARL) sub-field. Possible dynamics for the behavior of multiple learning agents that coexist in a shared environment. Extension (or generalization) of the basic principles to these type of complex models.

SESSION 21 (LIVE IN-PERSON)

RL APPLICATIONS [Assignment: GROUP ASSIGNMENT]

Exploration of business cases of practical application across the different industries.

SESSION 22 (LIVE IN-PERSON)

TOP CASES & COMPANIES

Company leaders in this field. Study of breakthrough examples (from AlphaGo, AlphaFold-2, Fusion Reactor Modeling, and beyond).

SESSION 23 (LIVE IN-PERSON)

TEST-EXAM

Concept summary review examination. 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 24 (LIVE IN-PERSON)

GROUP PROJECTS PRESENTATIONS (A)

Presentations, discussion and Q&A of group projects.

SESSION 25 (LIVE IN-PERSON)

GROUP PROJECTS PRESENTATIONS (B), WRAP-UP

Presentations, discussion and Q&A of group projects.

EVALUATION CRITERIA

We will behave as professionals. This means that I expect you to come to class prepared to discuss as if this was a meeting in your company. This is a small group, so we will take advantage of it doing a lot of direct interaction. Come prepared to class and be inquisitive.

The followng table summarizes the Ongoing grading (maximum) for each of the sessions and assignments:

SESSION	LECTURE	Participation	I.Assignment	G.Assignment	Test/Exam	TOTAL
DRL_01	INTRODUCTION TO THE COURSE (SYLLABUS)	4				4
DRL_02	RL INTRODUCTION RL BASIC CONCEPTS	4				4
DRL_03	MULTI-ARMED BANDITS	4				4
DRL_05	MONTE CARLO METHODS [Assignment: WARM-UP ASSIGNMENT]	4	40			44
DRL_06	TEMPORAL DIFFERENCES	4				4
DRL_07	A FIRST PRACTICE - INTEGRATION OF CONCEPTS	4				4
DRL_07	A FIRST PRACTICE - INTEGRATION OF CONCEPTS	4				4
DRL_08	LEARNING AND PLANNING	4				4
DRL_09	EVOLUTIONARY COMPUTATION	4				4
DRL_10	RL FRAMEWORKS	4				4
DRL_11	INTERMEDIATE PRACTICE	4				4
DRL_12	APPROXIMATION METHODS	4				4
DRL_13	POLICY GRADIENTS [Assignment: ASSIGNMENT]	4	40			44
DRL_14	ACTOR-CRITIC	4				4
DRL_15	PRACTICE REVIEW	4				4
DRL_16	PROBLEM APPROACH FRAMEWORK	4				4
DRL_17	DEEP RL PRACTICE [Assignment: ASSIGNMENT]	4	40			44
DRL_18	SUMMARY RECAP	4				4
DRL_19	MODEL-BASED, OFF-LINE, INVERSE RL RESEARCH PAPERS, INNOVATION TRENDS (A) [Assignment: EXAMINATION PRACTICE. BE PREPARED TO ANSWER QUESTIONS ABOUT RL]	4	55			59
DRL_20	MULTI-AGENT RL	4				4
DRL_21	RL APPLICATIONS [Assignment: GROUP ASSIGNMENT]	4		100		104
DRL_22	TOP CASES & COMPANIES	4				4
DRL_23	TEST-EXAM	4			75	79
DRL_24	GROUP PROJECTS PRESENTATIONS (A)	4		50		54
DRL_25	GROUP PROJECTS PRESENTATIONS (B), WRAP-UP	4				4

Summary of Criteria:

CRITERIA	POINTS	PERCENTAGE
Participation	100	20%
I.Assignment	175	35%
G.Assignment	150	30%
Test/Exam	75	15%
TOTAL	500	100%

criteria	percentage	Learning Objectives	Comments
Final Exam	15 %		
Individual Assignments	35 %		
Group Assignment	20 %		
Group Presentation	10 %		
Participation	20 %		

RE-SIT / RE-TAKE POLICY

BIBLIOGRAPHY

Compulsory

- Richard S. Sutton and Andrew G. Barto. (2018). *Reinforcement Learning: An Introduction.* 2nd edition. AbeBooks. ISBN 9780262039246 (Digital)

This book covers the ground essential to understanding much of the work out there published on RL. Could be hard going for the students without a relatively solid mathematical background (especially the Bellman equations and monte-carlo sections) but it is worth it. It is a must-read for anyone doing graduate research in reinforcement learning in order to get to grips with these matters. It talks about most important topics to get you started in the required direction.

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.

DEVOPS ENGINEERING

Bachelor in Data and Business Analytics BDBA SEP-2023
DE-DBA.1C.4.M.A

Area Others
Number of sessions: 25
Academic year: 23-24
Degree course: FOURTH
Number of credits: 5.0
Semester: 1º
Category: OPTIONAL
Language: English

Professor: **IVÁN MAURICIO DÍAZ LEIVA**

E-mail: idiazl@faculty.ie.edu

Ivan Diaz, Senior Analytics Manager at Kearney Analytics since May 2022, has 14 years of international experience in product & project management tohether with strategy consulting at prominent firms such as Kearney, BCG, Bain & Co, and Deloitte S&O. At Kearney he is the lead for the Data & Analytics transformation hub, and during his tenure at BCG GAMMA, Ivan served as the Head of the Iberia node, supervising 18 analysts across various disciplines, including Data Engineering, Data Science, Geo-Analytics, and Software Development. He holds an MSc in Business Analytics and Big Data from IE Business School (Dean's list) and an MBA (Cum Laude) from Stellenbosch Business School, along with certifications in PRINCE 2, Scrum Master, and Product Owner.

Since November 2018, Ivan has been sharing his expertise as an instructor for Data Visualization, Communication & Storytelling with Data courses at IE University's Master's in Business Analytics and Big Data, Bachelor's in Business Analytics and Big Data, and the ieXL BootCamp. His teaching has been recognized with academic awards for all three cohorts.

Few selected software development projects:

- Retail
  - Development and deployment of a cross country, harmonized software solution, for Trade Terms & Promo performance improvement (Azure, Databricks, Delta Lake, Power BI)
- Oil & Gas (Retail)
  - Software development and deployment of a Personalization Engine for fuel and convenience store retailers. Managing a cross-regional team of 6 Developers, including DevOps, Back & Front Ends, and Data Science and Engineering Modules (AWS, Spark, Docker, Django, React)
- Infrastructure
  - Development and deployment of a data-intensive platform for Project Planning optimization and prediction of large infrastructure project deviation for a large multinational construction firm (AWS, Spark, Django, React)

- Pharma
- Developed of a SaaS Control Tower solution, for the coordination and planning of Pharmaceutical studies across Europe (AWS, Postgres, Python, Tableau)

1. Teams or Slack channels set up for this purpose
2. idiazl@faculty.ie.edu

PREREQUISITES

Students should have good understanding of software development, and some IT general operations concepts. Including:

- Some good understanding of Python, and familiarity with the command line/ shell scripting language (e.g. bash) to perform basic operations such as file manipulation, directory navigation, and executing commands
- Familiarity with operating systems e.g. Windows, macOS, and Linux
- Basic knowledge of database, and basic understanding of concepts such as tables, queries, and SQL
- Some exposure to version control e.g. Git, Github or GitLab
- Students will be provided with access to Azure Resource groups, students are expected to familiarize themselves with this environment

Familiarity and scan reading to some reference of software (potentially) covered in class:

- Git: https://git-scm.com/
- Jenkins: https://www.jenkins.io/
- GitLab CI: https://docs.gitlab.com/ee/ci/
- Terraform: https://www.terraform.io/
- Azure DevOps: https://aws.amazon.com/cloudformation/
- Docker: https://www.docker.com/
- Kubernetes: https://kubernetes.io/
- Prometheus: https://prometheus.io/
- Grafana: https://grafana.com/

SUBJECT DESCRIPTION

The rapidly evolving field of software development has led to the rise of DevOps, a **set of practices that combine development and operations to streamline the software development process**. This course aims to equip students with the knowledge and skills required to understand and implement DevOps principles and practices, **enhancing their employability in a very competitive software industry.**

DevOps is a critical discipline, essential for modern software development and operations. In today's fast-paced and highly competitive digital landscape, organizations need to be able to **deliver software and services quickly, reliably, and at scale.** DevOps provides a set of principles, practices, and tools that enable organizations to achieve these goals by breaking down silos between development and operations teams, automating processes, and **building a culture of collaboration and continuous improvement (Agile).**

This course will introduce you to the fundamentals of DevOps and how it can help organizations to accelerate their software delivery and increase their operational efficiency. You will learn about key DevOps concepts, and how these concepts can be applied to build modern software applications. Including:

- Coud
- Continuous integration, continuous delivery (CI/CD)
- Infrastructure as code (IaaC)
- Containerization,
- Monitoring

DevOps is not only important now, **but it will also become increasingly important in the future**. As organizations continue to embrace digital transformation, they will need to be able to **deliver software and services more quickly, reliably, and securely than ever before**. DevOps provides a framework for achieving these goals by enabling organizations to automate their software delivery process, increase their operational efficiency, and improve their overall quality of service.

Caveat emptor: Faithfully adhering to the agile spirit, this syllabus will be modified and adapted to the specific requirements, needs, and pace of the class. Lead this journey, provide feedback, and let's get the most out of this experience together!

LEARNING OBJECTIVES

This course introduces the fundamentals of DevOps, a software development methodology that combines software development and operations to enhance collaboration and communication between teams. The course will cover topics related to DevOps tools and practices that enable continuous integration, continuous delivery, and continuous deployment. The course is designed for undergraduate students who have a basic understanding of software development. Between others you will be able to:

- Develop a solid understanding of DevOps principles and practices
- Gain hands-on experience with essential DevOps tools and technologies
- Understand the importance of communication and collaboration in the DevOps process
- Apply DevOps practices to enhance software development and delivery, through your own project

Disclaimer: The following description of the material covered is tentative. An attempt will be made to cover all listed topics.

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:

Learning Activity	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	20.0 %	25.0 hours
Discussions	8.0 %	10.0 hours
Exercises in class, Asynchronous sessions, Field Work	0.0 %	0.0 hours
Group work	48.0 %	60.0 hours
Individual studying	24.0 %	30.0 hours
TOTAL	100.0 %	125.0 hours

PROGRAM

SESSIONS 1 - 2 (LIVE IN-PERSON)

Introduction to DevOps (History and evolution)

- Understanding DevOps methodology
- The benefits of DevOps
- Differences between traditional and DevOps approaches
- DevOps principles

Activity: Define a simple software application project and establish its requirements (e.g. Generative AI app)

Extra (optional) reference:

- "What is DevOps?" by Atlassian
- "What is DevOps?" by Azure DevOps Web Services
- "What is DevOps?" by Azure
- "What is DevOps?" by IBM

SESSIONS 3 - 4 (LIVE IN-PERSON)

Agile Methodologies and DevOps Culture

- Agile principles and frameworks
- Relationship between Agile and DevOps
- Creating an Agile culture
- Communication and collaboration
- Agile methodologies in DevOps
- DevOps team structure

Activity: Plan and set up a Kanban board for the software application project

Extra (optional) reference:

"What is Agile?" by Atlassian

"Agile Development" by ThoughtWorks

"What is Agile Software Development?" by Agile Alliance

SESSIONS 5 - 6 (LIVE IN-PERSON)

Continuous Integration (CI)

- Introduction to continuous integration
- CI tools and best practices
- Build automation and testing
- Continuous integration tools
- Version Control systems (Git)
- Jenkins, GitHub Actions and its features

Activity: Set up a Git repository for the software application and configure a CI pipeline using Jenkins or GitHub Actions

Extra (optional) reference:

- "Continuous Integration vs Continuous Deployment vs Continuous Delivery" by Atlassian
- "What is Continuous Integration and Continuous Deployment?" by CircleCI
- "What is Continuous Deployment?" by IBM
- "What is Continuous Integration" by Azure

SESSIONS 7 - 8 (LIVE IN-PERSON)

CI/ CD Pipelines and Delivery

- Introduction to continuous Delivery
- Continuous delivery tools
- Building CI/CD pipelines
- Pipeline best practices

Activity: Extend the CI pipeline to include continuous delivery for the software application project

Extra (optional) reference:

- "Continuous Delivery" by Martin Fowler
- "Continuous Delivery vs. Continuous Deployment vs. Continuous Integration: What's the Difference?" by CircleCI

SESSIONS 9 - 10 (LIVE IN-PERSON)

Continuous Deployment (CD)

- Introduction to Continuous Deployment
- CD best practices and principles
- Deployment strategies (rolling, blue-green, canary)
- CD tools (Jenkins, GitLab, GitHub Actions)

Activity: Implement a continuous deployment strategy for the software application using a chosen CD tool

Extra (optional) reference:

- "Continuous Deployment - What, Why, and How" by Atlassian
- "Continuous Deployment 101: Best Practices for the Popular Engineering Approach" by Lucidchart
- "What is continuous deployment" by Atlassian
- "Continuous Deployment and Release Management" by Azure

SESSIONS 11 - 12 (LIVE IN-PERSON)

Configuration Management and IaaC

- Introduction to Configuration Management
- IaC concepts and principles
- Configuration Management tools (Ansible, Puppet, Chef)
- IaC tools (Terraform, CloudFormation)

Activity: Set up configuration management and IaC for the software application using a chosen toolset

Extra (optional) reference:

- "What is Configuration Management?" by Atlassian
- "Configuration Management: A Comprehensive Guide" by UpGuard
- "Configuration Management" by Azure
- "Infrastructure as Code" by Azure

SESSIONS 13 - 14 (LIVE IN-PERSON)

Containerization and Orchestration

- Introduction to containerization
- Benefits of using containers
- Docker basics (Dockerfile, Docker Compose, Docker Swarm)
- Container orchestration (Kubernetes, Azure AKS)

Activity: Containerize the software application using Docker and deploy it using Kubernetes or Azure AKS

Extra (optional) reference:

- "Introduction to Docker" by Docker
- "Kubernetes Tutorial" by Kubernetes
- "Azure Container Instances and container orchestrators" by Azure

SESSIONS 15 - 16 (LIVE IN-PERSON)

Monitoring and Logging

- Importance of monitoring and logging in DevOps
- Monitoring and logging principles
- Monitoring tools (Prometheus, Grafana, Zabbix)
- Logging tools (ELK Stack, Graylog, Splunk)

Activity: Set up monitoring and logging for the software application using chosen monitoring and logging tools

Extra (optional) reference:

- "Monitoring and Logging" by Martin Fowler
- "Introduction to the ELK Stack" by Elastic
- "Design a solution to log and monitor Azure resources" by Azure

SESSIONS 17 - 18 (LIVE IN-PERSON)

Security in DevOps (DevSecOps)

- Introduction to DevSecOps
- Integrating security in the DevOps lifecycle
- Security tools (Snyk, SonarQube, OWASP)
- Compliance and security best practices

Activity: Integrate security tools and practices into the software application's CI/CD pipeline

Extra (optional) reference:

- "Security in DevOps (DevSecOps)" by Azure
- "From DevOops! to DevSecOps" by SANS Institute
- "Integrating Security in DevOps: Best Practices, Tools, and Challenges" by TechStack

SESSIONS 19 - 20 (LIVE IN-PERSON)

Cloud Computing and DevOps

- Introduction to cloud computing
- Cloud service models (IaaS, PaaS, SaaS)
- Popular cloud providers (AWS, Azure, Google Cloud)
- Cloud-native DevOps practices

Activity: Migrate the software application to a cloud provider and apply cloud-native DevOps practices

Extra (optional) reference:

- "What Is Cloud Computing?" by Microsoft Azure
- "Cloud service models IaaS vs. PaaS vs. SaaS" by Red Hat
- "What's the Difference Between AWS vs. Azure vs. Google Cloud?" by Coursera

SESSIONS 21 - 22 (LIVE IN-PERSON)

Microservices and DevOps

- Introduction to microservices architecture
- Benefits and challenges of microservices
- Implementing DevOps in a microservices environment
- Service mesh (Istio, Linkerd, Consul)

Activity: Refactor the software application into microservices and deploy them using a service mesh

Extra (optional) reference:

- "Introduction to Microservices" by NGINX
- "What Are Microservices?" by Martin Fowler
- "Microservices Architecture Style" by Azure

SESSIONS 23 - 24 (LIVE IN-PERSON)

Future of DevOps

- Trends in DevOps
- Emerging technologies
- Serverless computing
- Machine learning and AI in DevOps

Activity: Explore and implement a serverless component or machine learning service in the software application

Extra (optional) reference:

- "The Future of DevOps: Trends to Watch in 2023 and Beyond" by Neil Sha
- "Future Scope of DevOps – 9 Reasons To Learn DevOps in 2023" by upGrad
- "Future of DevOps: Trends to Watch" by DevOps.com

SESSION 25 (LIVE IN-PERSON)

Course Review and Conclusion

- Review of the course content
- Project presentations
- Conclusion and next steps for DevOps mastery

Activity: Present the developed software application, showcasing the DevOps principles and practices used throughout the project

- "Advanced DevOps Practices" by Azure
- "DevOps for advanced teams"
- "5 reasons Advanced DevOps is a step up from “just” DevOps"

EVALUATION CRITERIA

criteria	percentage	Learning Objectives	Comments
Intermediate tests	10 %		
Individual work	25 %		
Group Work	25 %		
Group Presentation	20 %		
Class Participation	20 %		

RE-SIT / RE-TAKE POLICY

BIBLIOGRAPHY

Recommended

- Gene Kim. *A Novel about IT, DevOps, and Helping Your Business Win.* ISBN 0988262509 (Digital)
- Gene Kim. *The DevOps Handbook: How to Create World-Class Agility, Reliability, and Security in Tech Orgs.* ISBN 1942788002 (Digital)
- Gene Kim. *Accelerate: The Science of Lean Software and DevOps: Building and Scaling High Performing Tech Orgs.* ISBN 1942788339 (Digital)
- Jez Humble. *Continuous Delivery: Reliable Software Releases through Build, Test, and Deployment Automation.* ISBN 0321601912 (Digital)
- Betsy Beyer. *Site Reliability Engineering: How Google Runs Production Systems.* ISBN 149192912X (Digital)
- Sean P. Kane. *Docker: Up & Running: Shipping Reliable Containers in*

Production. ISBN 1491917571 (Digital)
- Brendan Burns. *Kubernetes: Up and Running: Dive into the Future of Infrastructure.* ISBN 1491935677 (Digital)
- Jesse Keating. *Mastering Ansible: Master the Ins and Outs of Advanced Operations with Ansible.* ISBN 1784395487 (Digital)
- Yevgeniy Brikman. *Terraform: Up & Running: Writing Infrastructure as Code.* ISBN 1491977086 (Digital)
- Mitesh Soni. *Implementing DevOps with Microsoft Azure: Automate your deployments and incorporate the DevOps cultu.* ISBN 1789610545 (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.

FINTECH: REVOLUTIONIZING FINANCIAL SERVICES VIA DESIGN AND TECHNOLOGY

Bachelor in Data and Business Analytics BDBA SEP-2023
FRFS-DBA.1C.4.M.A

Area Finance
Number of sessions: 30
Academic year: 23-24
Degree course: FOURTH
Number of credits: 6.0
Semester: 1º
Category: OPTIONAL
Language: English

Professor: **MANOEL FERNANDO GADI ALONSO**

E-mail: mfalonso@faculty.ie.edu

MANOEL FERNANDO GADI ALONSO

Academic Background:

- Executive MBA, Business Administration and Management, IE Business School, Spain
- Master in Computer Science and Statistics, University of Sao Paulo, Brazil
- Computer Science Degree, University of Sao Paulo, Brazil

Academic Experience:

- Risk & Fraud Analytics - Master in Business Analytics & Big Data at IE Business School
- Statistical Programming Python - Master in Business Analytics & Big Data at IE Business School
- Building a Fintech - Master in Business Analytics & Big Data at Universidad de Alcalá de Henares
- Math and Stats for Big Data - Master in Business Analytics & Big Data at U-TAD

Corporate Experience:

- Founder and CEO of Suncaged Analytics Consulting Europe (Madrid: 2019-curr)
- Director of Analysis & Reporting, Altamira Real Estate (Madrid: 2018-2019)
- Director of Credit Rating, Big Data and Business Analytics, Bravo Capital (Madrid: 2015-2018)
- Head of R&D for Risk Analytics Area, Santander Bank Headquarters Spain (Madrid: 2012-2015)
- R&D Risk Analytics Area Manager, Santander Bank United Kingdom (Milton Keynes: 2008-2012)
- Credit Risk Modelling Manager, Santander Bank Brazil (Sao Paulo: 2007-2008)
- Credit Risk Supervisor, ibi bank (C&A group) (Sao Paulo: 2006-2007)

- Credit Card Risk analyst, Citibank Brazil (Sao Paulo: 2003-2005)

mfalonso@faculty.ie.edu

PREREQUISITES

- Basic Stats
- Python programming (R or other programming language)

SUBJECT DESCRIPTION

Digitization has changed the way citizens and entrepreneurs interact with banks, their savings and financing.

The main actors of this disruption are the Financial technology companies, known as Fintech companies. Fintechs are companies with an important technological component, seeking new solutions within the financial sector through the use of new design and user experience via technology, apps and big data.

The Fintech field is a relatively new industry that applies technology to streamline financial processes. It gains terrain against traditional banks by simplifying bank services, by better controlling risks, by making improving trading decisions, or by simply by offering new generation the power control their finances.

LEARNING OBJECTIVES

The general objective of this course are:

- Explore how Fintech are reshaping the banking and payments industry,
- how Fintech are using technology to improve user experience focused on unbundling and untangling financial services,
- demystify blockchain and cryptocurrency from a technological and investment standing point,
- build a "Minimun Viable Fintech" and thus face the design dilemmas of the solution that a Fintech faces on a daily basis regarding design, user experience and technology,
- Finally, explore the future of Fintechs.

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:

Learning Activity	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	30.0 %	45.0 hours
Discussions	6.67 %	10.0 hours
Exercises in class, Asynchronous sessions, Field Work	26.67 %	40.0 hours
Group work	26.67 %	40.0 hours
Individual studying	10.0 %	15.0 hours
TOTAL	100.0 %	150.0 hours

PROGRAM

SESSION 1 (LIVE IN-PERSON)

Fintech 101 - Course Opening

SESSION 2 (LIVE IN-PERSON)

History of Banking

SESSION 3 (LIVE IN-PERSON)

Module 1: Reshaping the banking and payments industry

Mandatory to read the case study: "Bank 3.0: Why Banking Is No Longer Somewhere You Go But Something You Do (CHAPTER 1)"

Book Chapters: Bank 3.0: Why Banking Is No Longer Somewhere You Go But Something You Do (CHAPTER 1) (CED)

SESSION 4 (LIVE IN-PERSON)

Module 1: Reshaping the banking and payments industry

Mandatory to read the case study: "Bank 3.0: Why Banking Is No Longer Somewhere You Go But Something You Do (CHAPTER 1)" available on previous session

SESSION 5 (LIVE IN-PERSON)

Module 2.1: User Experience for unbundling and untangling financial services

Mandatory to read the case study:

Nubank case study.

Article: Nubank case study

SESSION 6 (LIVE IN-PERSON)

Module 2.1: User Experience for unbundling and untangling financial services

SESSION 7 (LIVE IN-PERSON)

Module 2.2: Fintech - Case Study Research and Writting

Mandatory to read the blog post:

Other / Complementary Documentation: How to Write a Convincing Case Study in 7 Steps

SESSION 8 (LIVE IN-PERSON)

Module 2.2: Fintech - Case Study Research and Writting

SESSION 9 (LIVE IN-PERSON)

Module 3.1: Harnessing data with Artificial Intelligence and Machine Learning for Fintech

SESSION 10 (LIVE IN-PERSON)

Module 3.1: Harnessing data with Artificial Intelligence and Machine Learning for Fintech

Mandatory to read the case study:

Multimedia Material: An introductory journey through Business Analytics (Youtube)

Practical Case : Do you fear an analytics driven competitor? You should! (IST010071-U-ENG-WOD)

SESSION 11 (LIVE IN-PERSON)

Module 3.2: Harnessing data with Artificial Intelligence and Machine Learning for Fintech

SESSION 12 (LIVE IN-PERSON)

Module 3.2: Harnessing data with Artificial Intelligence and Machine Learning for Fintech

Practical Model Development using Python, Pandas and Sklearn

SESSION 13 (LIVE IN-PERSON)

Module 4.1: Demystifying blockchain and cryptocurrency - Crypto 101

Mandatory to read the paper:

Article: The inefficiency of Bitcoin (Economics Letters 148 (2016) 80–82) (CED)

SESSIONS 14 - 15 (LIVE IN-PERSON)

Module 4.1: Demystifying blockchain and cryptocurrency - Crypto 101

****

Midterm Group Presentation (Case Study Presentation)

SESSION 16 (LIVE IN-PERSON)

Module 4.2: Demystifying blockchain and cryptocurrency - Advanced Crypto

Mandatory to read the paper:

Article: Bitcoin price prediction using machine learning: An approach to sample dimension engineering (Journal of Computational and Applied Mathematics 365 (2020)) (CED)

SESSION 17 (LIVE IN-PERSON)

Module 4.2: Demystifying blockchain and cryptocurrency - Advanced Crypto

Mandatory to read the paper:

SESSION 18 (LIVE IN-PERSON)

Module 4.3: Demystifying blockchain and cryptocurrency - Advanced Crypto - Part 3

Mandatory to read the paper: "Bitcoin price prediction using machine learning: An approach to sample dimension engineering" available on previous session.

SESSION 19 (LIVE IN-PERSON)

Module 4.3: Demystifying blockchain and cryptocurrency - Advanced Crypto - Part 3

Mandatory to read the paper: "Bitcoin price prediction using machine learning: An approach to sample dimension engineering" available on previous session.

SESSION 20 (LIVE IN-PERSON)

Module 4.1: Transforming personal finance with Technology - Introduction to Web Development

Mandatory to watch the videos:

Multimedia Material: HTML, CSS, JavaScript Explained (Youtube)

Multimedia Material: A brilliant introduction to Flask (Youtube)

Multimedia Material: Running Your First Flask Application (Youtube)

SESSION 21 (LIVE IN-PERSON)

Module 4.1: Transforming personal finance with Technology - Introduction to Web Development

SESSION 22 (LIVE IN-PERSON)

Module 4.2: Transforming personal finance with Technology - API Development and Model Deployment via API

SESSION 23 (LIVE IN-PERSON)

Module 4.2: Transforming personal finance with Technology - API Development and Model Deployment via API

Multimedia Material: Frameworks - Web Development (Youtube)

Multimedia Material: Intro to Flask Series (Youtube)

SESSION 24 (LIVE IN-PERSON)

Module 4.2: Transforming personal finance with Technology - API Development and Model Deployment via API

SESSION 25 (LIVE IN-PERSON)

Module 4.2: Transforming personal finance with Technology - API Development and Model Deployment via API

SESSION 26 (LIVE IN-PERSON)

Module 5: Forging the future of FinTech

Mandatory to read the article: "How banks can compete against an army of Fintech Startups"

Technical note: How Banks Can Compete Against an Army of Fintech Startups (HBS H03MS6-PDF-ENG)

SESSION 27 (LIVE IN-PERSON)

Module 5: The Digitial Banking Transformation Framework

Digital Banking Transformation Framework - Case Study.

Article: Digital Banking Transformation Framework - Case Study

SESSIONS 28 - 29 (LIVE IN-PERSON)

Final Group Presentation (Elevator Pitch - Fintechs)

SESSION 30 (LIVE IN-PERSON)

Final Exam

EVALUATION CRITERIA

criteria	percentage	Learning Objectives	Comments
E. Final Exam	30 %		
B. Quizzes	20 %		
C. Group project (40%/60%)	30 %		
A. Class Participation	20 %		

RE-SIT / RE-TAKE POLICY

A. Class participation and discussion

Class participation will be evaluated based on the following criteria:

- One participation point for attending class.
- Extra participation point for sharing experience or for star comment during class - quality (not quantity) of your participation in class discussion: The most important dimension of participation concerns what it is that you are saying.
- In case the professor requests participation in forums: 0,1 or 2 participation points are given depending on the quality of the participation.

B. Quizzes

For every module, throughout the course, a short online-quiz based on previously covered material will take place (unless otherwise specified by the professor at the beginning of the session). These quizzes will help you assess your overall understanding of the subject being studied and identify any caveat in your learning. NO MAKE UP FOR QUIZZES WILL BE PERMITTED.

C. Group project_Student Series

The group project is an integral part of this course. Each group (the ones defined by IE or randomly composed of N students defined by the professor) will be asked to work on a project, prepare and deliver a presentation.

D. Final Exam

There will be one final exam (open book). You can refer to any class material and also access the internet during the exam. HOWEVER, NO COMMUNICATION BETWEEN STUDENTS ARE ALLOWED DURING THE EXAMS (IN PERSON OR ONLINE). MOBILE PHONES ARE NOT ALLOWED.

As per University Policy:

Each student has 4 chances to pass any given course distributed in two consecutive academic years (regular period and July period).

It is mandatory to attend 100% of the classes. Students who do not comply with at least 70% attendance will lose their 1st and 2nd chance, and go directly to the 3rd one (they will need to enroll again in this course the next academic year).

Grading for retakes will be subject to the following rules:

1. Those students who failed the subject in the first regular period will have to do a retake in July (except those not complying with attendance rules who are banned from this possibility).

2. Dates and location of the July retakes will be posted in advance and will not be changed. Please take this into consideration when planning your summer.

3. The maximum grade that a student may obtain in the 2nd exam session is 8 out of 10. Those students in the 3rd call will be required to attend 50% of the classes. If due to schedule overlap, a different option will be discussed with the professor in order to pass the subject.

BIBLIOGRAPHY

Recommended

- Brett King. (2012). *Bank 3.0: Why Banking Is No Longer Somewhere You Go But Something You Do.* 1st edition. Wiley. ISBN 1118589637 (Digital)
- Miguel Grinberg. (2018). *Flask Web Development: Developing Web Applications with Python.* 2nd Edition. O'REILLY. ISBN 1491991739 (Digital)
- Yves Hilpisch. *Python for Finance: Mastering Data-Driven Finance, 2nd edition.* O'Reilly Media. ISBN 1492024333 (Printed)
- Raymond Anderson. *The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk Management and Decision Autom.* Oxford Press. ISBN 0199226407 (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.

FROM DATA TO DECISION MAKING IN MARKETING

Bachelor in Data and Business Analytics BDBA SEP-2023
FDTDM-DBA.1C.4.M.A

Area Data Science
Number of sessions: 30
Academic year: 23-24
Degree course: FOURTH
Number of credits: 6.0
Semester: 1º
Category: OPTIONAL
Language: English

Professor: **MARCUS JOSEPH SHERWIN**

E-mail: mjoseph@faculty.ie.edu

Marcus Sherwin holds a Bachelor in History and Classical Studies from University College Dublin, and an MBA in Business from Universidad Carlos III.

He has 9 years experience working in the Tech sector with leading multi national companies on their digital media strategies.

- Criteo Account strategist helping multinationals on their performance marketing strategy in Retail & Travel.
- Yiedlr Key Account Manager advising revenue management and marketing teams on SaaS integrations in the Aviation sector.
- TikTok Brand Partnership Manager working with Fintech clients on how they can invest with TikTok to reach audiences at scale.
- LinkedIn Senior Entreprise Client Solutions Manager working with B2B clients on how they can leverage LinkedIn to engage with their target audience.

mjoseph@faculty.ie.edu

Professor: **ROSARIO PERICUESTA CAMACHO**

E-mail: rpericuesta@faculty.ie.edu

ROSARIO PERICUESTA CAMACHO

E-mail: rpericuesta@faculty.ie.edu

Charo Pericuesta holds a Bachelor in Economics from the Universidad de Salamanca, a Bachelor in Market Research from the Universidad Autónoma de Madrid and a Master in Marketing and Sales Management from ESIC University in Madrid.

She has been more than a decade in Nielseniq, advising multinational companies in their strategy of innovation of their product portfolio using consumer data analysis.

Prior to that she worked:

- In the consumer insights department of Mondelez (Kraft) developing research plans to explore the different marketing initiatives of the company and using different market research techniques to understand the consumer behaviour and support decision making for the strategy of the brand;
- In United Planet in Boston as Marketing Project Manager;
- In Research International as Market Research Executive being seconded to the Chicago Office;
- In Russell Bedford International in London as Marketing Assistant.

rpericuesta@faculty.ie.edu

PREREQUISITES

No specific academic background required. Passion to learn.This course is particularly useful for those interested in pursuing marketing and marketing research related careers.

SUBJECT DESCRIPTION

Marketing professionals are nowadays exposed to an infinite number of data arising from an ever increasing number of sources. This generates an unpararel opportunity to understand consumer behavior but only to the extent that those professionals are able to correctly read and interpret those data.

This course provides an overview of the sources of information that Marketing teams are most exposed to and how to analyze and interpret such information to identify and solve marketing challenges.

LEARNING OBJECTIVES

- Identify the sources of information most used in Marketing.
- Learn to interpret the information gathered from those sources
- Learn to build a story telling from the data gathered by putting together a report which will be used as the basis for decision making
- Have a solid understanding of what is needed to create a digital marketing plan.
- Learn how to measure the effectivness of digital marketing campaigns.

TEACHING METHODOLOGY

The course will be divided into 30 sessions.

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 presentation assigment.

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.

Students are expected to create and execute a marketing plan targeting to specific audience to promote their product/service. Students will need to apply what they are learning not only to their group projects but also to the broader real-world context.

There will be a brief reading list included also some articles and recommended short Pod casts which links will be provided to access on class notes.

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:

Learning Activity	**Weighting**	**Estimated time a student should dedicate to prepare for and participate in**
Lectures	6.67 %	10.0 hours
Discussions	6.67 %	10.0 hours
Exercises in class, Asynchronous sessions, Field Work	20.0 %	30.0 hours
Group work	46.67 %	70.0 hours
Individual studying	20.0 %	30.0 hours
TOTAL	100.0 %	150.0 hours

PROGRAM

SESSION 1 (LIVE IN-PERSON)

SESSION 1 (MARCUS)

Course Intro

Introducing Digital - Marketing landscape, Key players in the space.

SESSION 2 (LIVE IN-PERSON)

SESSION 2 (CHARO)

Introduction to Marketing Research

- The marketing research process
- Research problem definition: typical bussiness questions

SESSION 3 (LIVE IN-PERSON)

SESSION 3 (MARCUS)

Your marketing strategy for your startup

Three options, B2C product, B2B product or service,

Resources will be provided on how to research your idea, in season 3, there will be a brief 5 minute presentation on why you chose the product/service, what is your thought process around the decision.

SESSION 4 (LIVE IN-PERSON)

SESSION 4 (CHARO)

Introduction to Story Telling

- The Minto Pyramid Principle
- How to write Headlines
- How to write Summaries

SESSION 5 (LIVE IN-PERSON)

SESSION 5 (MARCUS)

Your company vision & mission statement

Who is your audience, how are you servicing them, how will you target them.

This session will be around Qualitative info, writing your value proposition.

SESSION 6 (LIVE IN-PERSON)

SESSION 6 (CHARO):

Decision making using Cualitative Research: Focus groups and In-depth interviewing

SESSION 7 (LIVE IN-PERSON)

SESSION 7 (MARCUS)

Focus groups validating your idea

Focus solutions:

- Focus groups, how can you find users of the B2C or B2B service to validate your idea

SESSION 8 (LIVE IN-PERSON)

SESSION 8 (CHARO)

Decision making using Social Media and Mystery Shopper

SESSION 9 (LIVE IN-PERSON)

Session 9 (MARCUS)

Secondary sources:

- Competitor research, who is the leader in your industry, what gaps are there in the market:
- Survey analysis: Can you validate your audience by survey data.

SESSION 10 (LIVE IN-PERSON)

SESSION 10 (CHARO)

Decision making using Cuantitative Research: Surveys

SESSION 11 (LIVE IN-PERSON)

SESSION 11 (MARCUS)
Creating an ideal buyer
Who is this person
What do we know about them ?
How are we going to find them ?
Here we will look : Hubspot and how you can use it as the foundation for Marketing efforts.

SESSION 12 (LIVE IN-PERSON)

SESSION 12 (CHARO)
Qualitative Research - Group Presentations I

SESSION 13 (LIVE IN-PERSON)

SESSION 13 (MARCUS)
Building a marketing funnel:
What are the key steps

- Awareness key KPIs
- Consideration what are the core KPIs
- Lower funnel KPIs

SESSION 14 (LIVE IN-PERSON)

Qualitative Research - Group Presentations II

SESSION 15 (LIVE IN-PERSON)

SESSION 15 (MARCUS)
Marketing to your audience and how to do it per platform:
TikTok overview
Google Ad words/Youtube

SESSION 16 (LIVE IN-PERSON)

Syndicated sources of secondary data: Retail and Purchase Panel
Decision making using information derived from individuals or households on their buying behavior

SESSION 17 (LIVE IN-PERSON)

Session 17 (MARCUS)
Marketing to your audience Social platforms part 2
LinkedIn overview
Facebook/Instagram overview

SESSION 18 (LIVE IN-PERSON)

SESSION 18 (CHARO)

Understanding Price to improve Decision Making

SESSION 19 (LIVE IN-PERSON)

SESSION 19 (MARCUS)

Positioning digital on platforms

Competitors pricing: premium, mid level , mass markets.

Organic vs Paid content

SESSION 20 (LIVE IN-PERSON)

SESSION 20 (CHARO)

Data and Brand Growth:

Double jeopardy law

Pareto's Law

The duplication of purchase law

Natural Monopoly law

SESSION 21 (LIVE IN-PERSON)

Session 21 (MARCUS)

Getting set up

Linkedin ad account

TikTok Ad account

Implementing conversion tracking

Setting up your audiences lookalike audiences, retargeting audiences

SESSION 22 (LIVE IN-PERSON)

Group Presentations: Quantitative Research I

SESSION 23 (LIVE IN-PERSON)

SESSION 23 (MARCUS)

User Identity.

Identity cookies, md5, walled gardens vs the open internet.

(Last 15 minutes) Building out your marketing plan, analysing, what platforms you will find your audience.

SESSION 24 (LIVE IN-PERSON)

SESSION 24 (CHARO)

Group Presentations: Quantitative Research II

Diagnostic Consulting: Decision making using Correlations, Quadmaps, Brand positioning maps etc

SESSION 25 (LIVE IN-PERSON)

Session 25 (MARCUS)

1p 2p 3p data GDBR & Data privacy, measurement platforms, Dataroma, Google data hub.

Google Analytics

Analyzing attribution

SESSION 26 (LIVE IN-PERSON)

Session 26 (CHARO)

Market innovation breakthrough

SESSION 27 (LIVE IN-PERSON)

Session 27 (MARCUS)

Best D2C examples

Understanding campaign performance, live demos

What can be done to improve performance, creative, bid or messaging.

SESSION 28 (LIVE IN-PERSON)

Session 28 (CHARO)

Final exam

SESSION 29 (LIVE IN-PERSON)

SESSION 29 (MARCUS)

Presentation of your marketing plan, where you will invest your seed capital on what channels, what are your expected returns, how will you allocate what & how much.

SESSION 30 (LIVE IN-PERSON)

SESSION 30 (CHARO)

Presentation of your marketing plan, where you will invest your seed capital on what channels, what are your expected returns, how will you allocate what & how much.

EVALUATION CRITERIA

criteria	percentage	Learning Objectives	Comments
Final Exam	30 %		
Group Work	20 %		
Group Presentation	30 %		
Class Participation	20 %		

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 70% 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** 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

- Daniel Nunan, David F. Birks, Naresh K. Malhotra. *Marketing Research. Applied Insight..* Pearson Education Limited. ISBN 9781292308722 (Printed)
- Barbara Minto. *The Pyramid Principle: Logic in Writing and Thinking.* Pearson Education Limited. ISBN 9781292372266 (Printed)
- Byron Sharp. *How Brands grow, what marketers don't know.* OUP Australia & New Zealand. ISBN 9780195573565 (Printed)

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.

INTRODUCTION TO HEALTHCARE

Bachelor in Data and Business Analytics BDBA SEP-2023 IH-DBA.1C.4.M.A

Area Others

Number of sessions: 25

Academic year: 23-24

Degree course: FOURTH

Number of credits: 5.0

Semester: 1º

Category: OPTIONAL

Language: English

Professor: **ANA MARIA GLAVAN**

E-mail: aglavan@faculty.ie.edu

Ana Maria holds a Ph.D. in Mathematical Engineering from Carlos III University of Madrid. She taught different topics of Mathematics at Carlos III University, University of Navarra and IE University. Her actual research interests are on mathematical analysis and numerical simulations of biological and medical systems, and new trends in the fintech industry.

Academic Qualifications

- Ph.D. in Mathematical Engineering, Carlos III University, 2016
- Master in Biosensors for Environmental Monitoring, Perpignan University and Bucharest University, 2003
- B.Sc. in Chemistry, Bucharest University, 2001

Academic Experience

- Adjunct Professor, IE Business School, 2016-present
- Invited Professor, Carlos III University, Ph.D. Programme in Mathematical Engineering, 2016-2017
- Assistant Professor, University of Navarra, 2010-2012
- Teaching Assistant, Carlos III University, 2004-2009
- Research Assistant, BIOMEM, Centre de Phytopharmacie, 2002-2003

Fields of Interest

- Computational methods in Biology and Finance
- Applied and computational mathematics
- Numerical modelling algorithms

aglavan@faculty.ie.edu

PREREQUISITES

The course is designed to introduce the healthcare system and to help identifying the opportunities of using data science for improving and personalizing it. No previous experience is needed.

SUBJECT DESCRIPTION

Introduction to Healthcare course provides an interdisciplinary environment designed for students interested in healthcare, technology, and data science.

This course provides an in-depth overview of the healthcare ecosystem, the economic tools and concepts used in the analysis of healthcare system, and identifies the opportunities and challenges for data science applications in this booming sector.

LEARNING OBJECTIVES

The course is designed for students interested in Life Sciences and Biomedical disciplines, and its main objective is to see how they can use new technologies and data science to improve the quality of healthcare services.

Specifically, this course is designed to achieve several objectives:

- Describe the differences between public, private and non-profit health organizations
- Recognize the stakeholders in the healthcare system, their roles and responsibilities
- Identify the main challenges of the healthcare system
- Understand the structure and functions of pharmaceutical entities, the drug discovery process and future trends in the pharmaceutical sector
- Overview the values and applications of bioengineering in solving actual health issues
- Understand the impact of biomedical research in modern healthcare and identify the challenges faced by scientists in the field
- Examine in detail the Economics theories applied in healthcare, and analyze the economic implications of actual healthcare reforms
- Explain the difference between private insurance plans and public insurance programs
- Identify the professionals delivering support and proficiency in different facets of the pharmaceutical and healthcare businesses
- Understand the essential concepts, uses and innovations in biotechnology and health technologies
- Analyze how the actual status of the healthcare industry can be improved through the use of data science and analytics
- Investigate the actual industry environment covering personalized and precision medicine
- Examine and improve groundbreaking business models for health tech startups in creating value-based healthcare 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:

Learning Activity	**Weighting**	**Estimated time a student should dedicate to prepare for and participate in**
Lectures	9.6 %	12.0 hours
Discussions	6.0 %	7.5 hours
Exercises in class, Asynchronous sessions, Field Work	12.0 %	15.0 hours
Group work	18.4 %	23.0 hours
Individual studying	54.0 %	67.5 hours
TOTAL	100.0 %	125.0 hours

PROGRAM

SESSION 1 (LIVE IN-PERSON)

Overview of the Healthcare Industry:

- Basic terminology
- Introduction to healthcare systems: structure and stakeholders
- Public vs private health systems
- Discussion about characteristics of healthcare systems in different countries
- Case studies

SESSION 2 (LIVE IN-PERSON)

Overview of the Healthcare Industry:

- Basic terminology
- Introduction to healthcare systems: structure and stakeholders
- Public vs private health systems
- Discussion about characteristics of healthcare systems in different countries
- Case studies

SESSION 3 (LIVE IN-PERSON)

Overview of the Healthcare Industry:

- Basic terminology
- Introduction to healthcare systems: structure and stakeholders
- Public vs private health systems
- Discussion about characteristics of healthcare systems in different countries
- Case studies

SESSION 4 (LIVE IN-PERSON)

Overview of the Healthcare Industry:

- Basic terminology
- Introduction to healthcare systems: structure and stakeholders
- Public vs private health systems
- Discussion about characteristics of healthcare systems in different countries
- Case studies

SESSION 5 (LIVE IN-PERSON)

Introduction to Life Sciences

SESSION 6 (LIVE IN-PERSON)

Introduction to Pharmaceutical sector. Case study

SESSION 7 (LIVE IN-PERSON)

Introduction to Pharmaceutical sector. Case study

SESSION 8 (LIVE IN-PERSON)

Introduction to Bioengineering (challenges and applications)

SESSION 9 (LIVE IN-PERSON)

Introduction to Bioengineering (challenges and applications)

SESSION 10 (LIVE IN-PERSON)

The importance of biomedical research. Case study

SESSION 11 (LIVE IN-PERSON)

The importance of biomedical research. Case study

SESSION 12 (LIVE IN-PERSON)

Consultants for pharmaceuticals and hospitals

SESSION 13 (LIVE IN-PERSON)

Quiz after Module 1 and Final Project Task Description

SESSION 14 (LIVE IN-PERSON)

Healthcare Economics:

- Health economics
- Health insurance; Insurtech

- Operations management in healthcare: tradeoffs between costs, patients experience, and clinical quality
- Operations management in healthcare: pricing of procedures and their margins

SESSION 15 (LIVE IN-PERSON)

Healthcare Economics:

- Health economics
- Health insurance; Insurtech
- Operations management in healthcare: tradeoffs between costs, patients experience, and clinical quality
- Operations management in healthcare: pricing of procedures and their margins

SESSION 16 (LIVE IN-PERSON)

Healthcare Economics:

- Health economics
- Health insurance; Insurtech
- Operations management in healthcare: tradeoffs between costs, patients experience, and clinical quality
- Operations management in healthcare: pricing of procedures and their margins

SESSION 17 (LIVE IN-PERSON)

Biotech and health technologies sectors

SESSION 18 (LIVE IN-PERSON)

Health Tech startups landscape overview: opportunities for data science. Case study

SESSION 19 (LIVE IN-PERSON)

Health Tech startups landscape overview: opportunities for data science. Case study

SESSION 20 (LIVE IN-PERSON)

Improving healthcare quality – characteristics and new strategies. Personalized medicine

SESSION 21 (LIVE IN-PERSON)

Improving healthcare quality – characteristics and new strategies. Personalized medicine

SESSION 22 (LIVE IN-PERSON)

Future Trends and Challenges in Health Tech. New business models

SESSION 23 (LIVE IN-PERSON)

Group Presentations Part I

SESSION 24 (LIVE IN-PERSON)

Group Presentations Part II

SESSION 25 (LIVE IN-PERSON)

Final Exam

EVALUATION CRITERIA

The final grade of the course will be based on both individual and group work and it will be weighted in the following way:

- **Class Participation 10%**
- **Case Studies 20%**
- **Intermediate Quiz 20%**
- **Final Group Presentation 25%**
- **Final Exam 25%**

Case Studies represent a fundamental learning mechanism and are also crucial for the well-functioning of the class. The pre-class preparation and in-class participation will make the course pleasant and intellectually stimulating.

Final Presentation is a group presentation where the students present their proposal for the creation of a Tech company delivering service to the healthcare industry.

The Intermediate Quiz and Final Exam will comprise both theoretical and applications of the topics covered during the course.

criteria	percentage	Learning Objectives	Comments
Final Exam	25 %		
Intermediate tests	20 %		
Group Work	20 %		Case Studies
Group Presentation	25 %		
Class Participation	10 %		Class Discussions

RE-SIT / RE-TAKE POLICY

Grading for retakes will be subject to the following rules:

1. Those students who failed the subject in the first regular period will have to do a retake in July (except those not complying with attendance rules who are banned from this possibility).
2. Dates and location of the July retakes will be posted in advance and will not be changed. Please take this into consideration when planning your summer.
3. The maximum grade that a student may obtain in the 2nd exam session is 8 out of 10.

BIBLIOGRAPHY

Recommended

- Dakota Mitchell; Lee Haroun. *Introduction to Health Care.* 4th. ISBN 1305575075 (Digital)

- Christensen, Clayton M., Jerome H. Grossman M.D., and Jason Hwang M.D.. *The Innovator's Prescription: A Disruptive Solution for Health Care.* McGraw-Hill. ISBN 0071592083 (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.

MACHINE LEARNING FOR HEALTHCARE

Bachelor in Data and Business Analytics BDBA SEP-2023
MLFH-DBA.1C.4.M.A

Area Others

Number of sessions: 20

Academic year: 23-24

Degree course: FOURTH

Number of credits: 4.0

Semester: 1º

Category: OPTIONAL

Language: English

Professor: **JUAN CARLOS IBAÑEZ RODRIGUEZ**

E-mail: jibanezr@faculty.ie.edu

Dr Ibáñez has a PhD in Applied Computational Statistics from Lancaster University (UK); MSc in Mathematical Engineering from Universidad Carlos III (Spain), and Ingeniero Superior en Geodesia by Universidad Politécnica de Madrid (Spain)

Currently he is Chief Data and Analytics Officer at Wecity, CDO and founder of Real Estate AI; and lecturer in Data Science for several institutions in Madrid: IE, AFI y UCJC. Recent roles are CDO for Atresmedia , CDO for Urban Data Analytics, Academic Director at AFI, and previously head in analytics for multinational companies like Orange, SAS or Accenture

The office hours will be carried on online and under demand (email) of the student.

PREREQUISITES

Machine Learning.

SUBJECT DESCRIPTION

Today rigorous statistical models are applied to problems in medicine and pharmacy. During the last two decades, advances in Machine Learning have enabled the resolution of complex problems in the health industry.

In this course we cover real life models used in Health industry, from GLM models for clinical trials, to ML models for patient clustering or survival analysis.

LEARNING OBJECTIVES

- Master the most coomonly used GLM and ML models in health.
- Understand the clinical and diagnostic processes and use GLM+ML tools for their improvement/resolution.
- Solve real-life practical examples.

TEACHING METHODOLOGY

Learning Activity	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	30.0 %	30.0 hours
Discussions	0.0 %	0.0 hours
Exercises in class, Asynchronous sessions, Field Work	30.0 %	30.0 hours
Group work	10.0 %	10.0 hours
Individual studying	30.0 %	30.0 hours
TOTAL	100.0 %	100.0 hours

PROGRAM

SESSIONS 1 - 2 (LIVE IN-PERSON)

Statistical Modelling in Medicine

SESSION 3 (LIVE IN-PERSON)

Use Case: regression analysis - Predicting Blood Pressure

SESSION 4 (LIVE IN-PERSON)

Use Case: classification - Risk Stratification

SESSION 5 (LIVE IN-PERSON)

Statistical Modelling: Time series and panel data

SESSION 6 (LIVE IN-PERSON)

Statistical Modelling: Time series and panel data

SESSION 7 (LIVE IN-PERSON)

Use Case: time series - Health demand forecasting

SESSION 8 (LIVE IN-PERSON)

Use Case: survival analysis - Disease Progression

SESSION 9 (LIVE IN-PERSON)

Use Case: clinical trials

SESSION 10 (LIVE IN-PERSON)

Midterm exam

SESSION 11 (LIVE IN-PERSON)

Machine Learning Methods for Health Data Analysis

SESSION 12 (LIVE IN-PERSON)

ML fundamentals

SESSION 13 (LIVE IN-PERSON)

Use Case: patient clustering

SESSION 14 (LIVE IN-PERSON)

Use Case: regression analysis - Predicting Length of Hospital Stay

SESSION 15 (LIVE IN-PERSON)

Use Case: survival analysis - survival outcomes of cancer patients

SESSION 16 (LIVE IN-PERSON)

Use Case: time series - Health insurance claims

SESSION 17 (LIVE IN-PERSON)

Use Case: clinical trials

SESSION 18 (LIVE IN-PERSON)

Use Case: clinical trials

SESSION 19 (LIVE IN-PERSON)

Evaluation

SESSION 20 (LIVE IN-PERSON)

Evaluation

EVALUATION CRITERIA

criteria	percentage	Learning Objectives	Comments
Final Exam	30 %		Practical
Intermediate Tests	30 %		Multiple Choice test
Group Work	30 %		Group project
Class Participation	10 %		Class Participation

RE-SIT / RE-TAKE POLICY

BIBLIOGRAPHY

Compulsory

- Ewen Harrison, Riinu Pius. *R for Health Data Science.* Chapman and Hall/CRC. ISBN 0367428199 (Printed)

- Ruth Etzioni, Micha Mandel, Roman Gulati. *Statistics for Health Data Science: An Organic Approach.* Springer. ISBN 978303059888 (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.

MANAGING PROJECTS AND TEAMS IN DATA SCIENCE

Bachelor in Data and Business Analytics BDBA SEP-2023
MPTDS-DBA.1C.4.M.A

Area Operations and Business Analytics
Number of sessions: 15
Academic year: 23-24
Degree course: FOURTH
Number of credits: 3.0
Semester: 1º
Category: OPTIONAL
Language: English

Professor: **ALVARO ROMERO MIRALLES**

E-mail: aromerom@faculty.ie.edu

Álvaro Romero is Data Science Technical Director in the group of Energy Predictive Analytics of Instituto de Ingeniería del Conocimiento (IIC). He holds a Master's Degree in Computer Engineering, a Master's Degree in ICT Research and Innovation in Computational Intelligence and an EMBA from IE. His thesis on prediction of electricity prices in Spain has been published in the International Journal of Interactive Multimedia and Artificial Intelligence. He graduated in Mathematics and Computer Engineering from Universidad Autónoma de Madrid. He has experience in fraud detection, predictive maintenance, optimization problems, forecasting of demand and renewable energy production, etc. He also collaborates as a professor in different Masters and Bachelors of Data Sciece, BI and Big Data in business schools such as MBIT School and ENAE Business School or IE.

aromerom@faculty.ie.edu

PREREQUISITES

No prerequisites are needed.

SUBJECT DESCRIPTION

Artificial Intelligence and Big Data are cutting edge technologies that are changing the world we live. According to the McKinsey Global Survey of the last few years, AI adoption has steadily increased and a majority of executives report that this technology is generating returns (revenue increase or cost decrease).

The latest McKinsey report also indicates that there are at least six sets of practices differentiate high-performing companies from others: strategy, talent and leadership, ways of working, technology, data and adoption.

This subject covers what managers need to know to develop a business culture to empower Artificial Intelligence and Data projects. In particular, it will cover what are the important aspects when developing projects and managing data science teams by analysing the different parts of the AI process and detailing what are the key aspects to success or to fail in each.

LEARNING OBJECTIVES

In this course the student learns to manage, develop and evaluate the performance of teams in Artificial Intelligence and Big Data.

At the end of the course, students will obtain a solid grasp of the theoretical and practical application of managing high technical teams. In particular the subject will cover areas such as:

- Project planning in data science.
- Team goals.
- Managing technical people.
- Managing stakeholders.

TEACHING METHODOLOGY

Learning Activity	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	20.0 %	15.0 hours
Discussions	20.0 %	15.0 hours
Exercises in class, Asynchronous sessions, Field Work	13.33 %	10.0 hours
Group work	33.33 %	25.0 hours
Individual studying	13.33 %	10.0 hours
TOTAL	100.0 %	75.0 hours

PROGRAM

SESSION 1 (LIVE IN-PERSON)

Introduction

This first session is an introduction to the key concepts of the course. This section will also answer why is different managing teams in Data Science than other projects?

SESSION 2 (LIVE IN-PERSON)

Roles in Data Science

Data scientists are the most known role in the Data Ecosystem but there are many more. It is important to understand the differences, the interests and the motivitations of each person of the team.

SESSION 3 (LIVE IN-PERSON)

Attract and Retain Top Technology Talent

Hiring and retaining technical people is really difficult. There is a lot of competition on this profiles. IT pros receive twice as many LinkedIn InMails versus the average LinkedIn member, and software
engineers get more than two-and-a-half times as many.

Practical Case: Southwest Airlines (A) (HBS HR1A-PDF-ENG)

SESSION 4 (LIVE IN-PERSON)

Scaling Data teams

When starting a startup or the new data department is not easy to define what are the roles you need how you are going to make the people grow and so on. We are going to discuss how to make the team grow.

SESSION 5 (LIVE IN-PERSON)

CRISP-DM

Cross-industry standard process for data mining, known as CRISP-DM, is an open standard process model that describes common approaches used by data mining experts. It is the most widely-used analytics model.

SESSION 6 (LIVE IN-PERSON)

Project requirements

The profesor will send a Request for Proposals for a specific purpose and the students will need to read it and start working to define the Project and design a solution, a team, etc. based on what we cover in class.

SESSION 7 (LIVE IN-PERSON)

Project Management Methodologies

Review of Project management methodologies and how to adapt them to Data projects.

SESSION 8 (LIVE IN-PERSON)

Project Management Methodologies

Review of Project management methodologies and how to adapt them to Data projects.

SESSION 9 (LIVE IN-PERSON)

Presentation of the first assignment.

SESSION 10 (LIVE IN-PERSON)

Business Understanding: refining requirements

Meetings with the “client” (professor) to refine the requirements of the Project.

SESSION 11 (LIVE IN-PERSON)

Data preparation

Data preparation is one of the most important steps in data projects. We are going to cover risks, problems, estimation, etc.

SESSION 12 (LIVE IN-PERSON)

Training & Evaluation

Step in which the team train the models and evaluate them.

SESSION 13 (LIVE IN-PERSON)

Group Project presentation

SESSION 14 (LIVE IN-PERSON)

MLops

When Machine Learning models are in operation is important to apply DevOps principles (Continuous Integration and development, tests, etc.)

SESSION 15 (LIVE IN-PERSON)

Exam

EVALUATION CRITERIA

criteria	percentage	Learning Objectives	Comments
Final Exam	20 %		
Individual Work	30 %		
Workgroups	40 %		
Class Participation	10 %		

RE-SIT / RE-TAKE POLICY

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.

REAL TIME DATA AND AI

Bachelor in Data and Business Analytics BDBA SEP-2023
RTDAI-DBA.1C.4.M.A

Area Others
Number of sessions: 25
Academic year: 23-24
Degree course: FOURTH
Number of credits: 5.0
Semester: 1º
Category: OPTIONAL
Language: English

Professor: **MANOEL FERNANDO GADI ALONSO**

E-mail: mfalonso@faculty.ie.edu

MANOEL FERNANDO GADI ALONSO

Academic Background:

- Executive MBA, Business Administration and Management, IE Business School, Spain
- Master in Computer Science and Statistics, University of Sao Paulo, Brazil
- Computer Science Degree, University of Sao Paulo, Brazil

Academic Experience:

- Risk & Fraud Analytics - Master in Business Analytics & Big Data at IE Business School
- Statistical Programming Python - Master in Business Analytics & Big Data at IE Business School
- Building a Fintech - Master in Business Analytics & Big Data at Universidad de Alcalá de Henares
- Math and Stats for Big Data - Master in Business Analytics & Big Data at U-TAD

Corporate Experience:

- Founder and CEO of Suncaged Analytics Consulting Europe (Madrid: 2019-curr)
- Director of Analysis & Reporting, Altamira Real Estate (Madrid: 2018-2019)
- Director of Credit Rating, Big Data and Business Analytics, Bravo Capital (Madrid: 2015-2018)
- Head of R&D for Risk Analytics Area, Santander Bank Headquarters Spain (Madrid: 2012-2015)
- R&D Risk Analytics Area Manager, Santander Bank United Kingdom (Milton Keynes: 2008-2012)
- Credit Risk Modelling Manager, Santander Bank Brazil (Sao Paulo: 2007-2008)
- Credit Risk Supervisor, ibi bank (C&A group) (Sao Paulo: 2006-2007)
- Credit Card Risk analyst, Citibank Brazil (Sao Paulo: 2003-2005)

mfalonso@faculty.ie.edu

PREREQUISITES

- Knowledge acquired in previous or concurrent courses, in particular:
  - Python
  - Machine Learning
  - Stream Processing

SUBJECT DESCRIPTION

In this course, you will learn about various big data technologies and how they are used to solve real-world problems at scale. We will start by recaping streaming in Spark and distributed data processing for Machine Learning and AI with NLP and Large Language Models. Then, we will learn to development several Machine Learning and AI algorithms in a distributed enviroment and also deploy those models into Real Time applications. We then explore several use cases of real time technology in real company like LinkedIn, Netflic and ING Bank.

LEARNING OBJECTIVES

1. Understand how big data technologies are being used in the entertainment industry, with a focus on Netflix's "Stranger Things" production.
2. Learn about the role of Kafka in real-time data processing and analytics
3. Explore the use of real-time fraud detection system in preventing fraudulent activities.
4. Understand how Anti Money Laundering Revolution helps financial institutions stay compliant with regulatory requirements.
5. Develop a working knowledge of ChatGPT and large language models, and learn how they can be used to solve complex problems.

TEACHING METHODOLOGY

Learning Activity	**Weighting**	**Estimated time a student should dedicate to prepare for and participate in**
Lectures	28.0 %	35.0 hours
Discussions	4.0 %	5.0 hours
Exercises in class, Asynchronous sessions, Field Work	8.0 %	10.0 hours
Group work	40.0 %	50.0 hours
Individual studying	20.0 %	25.0 hours
TOTAL	100.0 %	125.0 hours

PROGRAM

SESSION 1 (LIVE IN-PERSON)

Introduction to the course and recap of basic Streaming Processing technologies and terminology.

SESSION 2 (LIVE IN-PERSON)

Data Processing (recap) for Machine Learning in Spark

SESSION 3 (LIVE IN-PERSON)

Distributed Machine Learning in Spark - Development of **Clustering** in Spark and Deployment in Real Time solution

SESSION 4 (LIVE IN-PERSON)

Distributed Machine Learning in Spark - Development of **Recommender System** in Spark and Deployment in Real Time solution

SESSION 5 (LIVE IN-PERSON)

Distributed Machine Learning in Spark - Development of **Linear Regression** in Spark and Deployment in Real Time solution

SESSION 6 (LIVE IN-PERSON)

Distributed Machine Learning in Spark - Development of **Logistic Regression** in Spark and Deployment in Real Time solution

SESSION 7 (LIVE IN-PERSON)

Distributed Machine Learning in Spark - Development of **Random Forest** in Spark and Deployment in Real Time solution

SESSION 8 (LIVE IN-PERSON)

Distributed Machine Learning in Spark - Development of **Natural Language Processing** in Spark and Deployment in Real Time solution

SESSION 9 (LIVE IN-PERSON)

Graph Technology for Real Time processing

SESSION 10 (LIVE IN-PERSON)

ChatGPT - scaling Large **Language Models** for Real Time use

SESSION 11 (LIVE IN-PERSON)

ChatGPT - scaling Large **Language Models** for Real Time use

SESSION 12 (LIVE IN-PERSON)

Recap for Mid Term Exam

SESSION 13 (LIVE IN-PERSON)

MID-TERM EXAM (Quiz)

SESSION 14 (LIVE IN-PERSON)

ChatGPT - connecting Large Language Models to Real Time user applications

SESSION 15 (LIVE IN-PERSON)

Use case 1: The creation and use of Kafka at **LinkedIn**.

SESSION 16 (LIVE IN-PERSON)

Use case 2: How **Netflix** Does Stranger Things with Big Data at Scale

SESSION 17 (LIVE IN-PERSON)

Use case 3: **ING Bank** - The Evolution with Kafka

SESSION 18 (LIVE IN-PERSON)

Use Case 4: Modern Risk Managment at scale with Spark and **Delta Lake**

SESSION 19 (LIVE IN-PERSON)

Use Case 4: Modern Risk Managment at scale with Spark and **Delta Lake**

SESSION 20 (LIVE IN-PERSON)

Use Case 5: Modern Model Managment at scale with Spark and **MLFlow**

SESSION 21 (LIVE IN-PERSON)

Use Case 5: Modern Model Managment at scale with Spark and **MLFlow**

SESSION 22 (LIVE IN-PERSON)

Use case 6: Anti Money Laundering Revolution with **Databricks**

SESSION 23 (LIVE IN-PERSON)

Recap for Mid Term Exam and Q & A for the group project

SESSION 24 (LIVE IN-PERSON)

FINAL EXAM - QUIZ - content of session 14 to 22.

SESSION 25 (LIVE IN-PERSON)

Final Group Presentation

EVALUATION CRITERIA

Throughout this course, you will be asked to read material related to the sessions, participate in live and online discussions, complete individual assignments and labs, and participate in a group project.

Specifically, grading will be based on the following criteria.

criteria	percentage	Learning Objectives	Comments
Final Exam	30 %		
Midterm Exam	20 %		
Group Assignment	30 %		
Class Participation	20 %		

RE-SIT / RE-TAKE POLICY

Midterm Exam

- Content from sessions 1-13

Final Exam

- Content from the whole course

Group Assignment

Group assignments will be developed by teams, who will have to work on a real life Use Case involving Stream Analytics. The team will play the role of a Chief Technology Officer who puts together the best solution taking into account technical, business and environmental constraints.

A list of topics will be proposed by the professor, but students are free to propose their own subjects.

Class Participation

You are expected to attend every class and participate in the discussions and class activities. This includes optional exercises, voluntary participation on the whiteboard, discussion board (forum) activity, class attendance, and active participation in in-class discussions, with the goal of ensuring a continued learning process, good teamworking, and ability to apply class concepts in real-world problems. Participation is based on the quality, rather on the quantity, of your contributions.

Late Assignments/Presentation

If you should be late in submitting an assignment, be sure to inform me in advance of the reason to try and accommodate you. Any unjustified late submission will not be graded.

Minimum passing grade

A minimum passing grade applies in final exams (3.5). If your score is lower than this minimum you will have to retake, irrespective of your overall course grade. Also keep in mind that the overall passing course grade is 5.0.

BIBLIOGRAPHY

Recommended

- Tyler Akidau, Slava Chernyak, Reuven Lax. *Streaming Systems.* 2018. O'Reilly Media, Inc.. ISBN 9781491983874 (Digital)

Streaming data is a big deal in big data these days. As more and more businesses seek to tame the massive unbounded data sets that pervade our world, streaming systems have finally reached a level of maturity sufficient for mainstream adoption. With this practical guide, data engineers, data scientists, and developers will learn how to work with streaming data in a conceptual and platform-agnostic way.

- Jay Kreps. (2014). *I Heart Logs.* O'Reilly Media, Inc.. ISBN 9781491909386 (Digital)

Why a book about logs? That’s easy: the humble log is an abstraction that lies at the heart of many systems, from NoSQL databases to cryptocurrencies. Even though most engineers don’t think much about them, this short book shows you why logs are worthy of your attention.

- Piethein Strengholt. (2020). *Data Management at Scale.* O'Reilly Media, Inc.. ISBN 9781492054788 (Digital)

As data management and integration continue to evolve rapidly, storing all your data in one place, such as a data warehouse, is no longer scalable. In the very near future, data will need to be distributed and available for several technological solutions. With this practical book, you’ll learn how to migrate your enterprise from a complex and tightly coupled data landscape to a more flexible architecture ready for the modern world of data consumption.

- Martin Kleppmann. (2017). *Designing Data-Intensive Applications.* O'Reilly Media, Inc.. ISBN 9781449373320 (Digital)

Data is at the center of many challenges in system design today. Difficult issues need to be figured out, such as scalability, consistency, reliability, efficiency, and maintainability. In addition, we have an overwhelming variety of tools, including relational databases, NoSQL datastores, stream or batch processors, and message brokers. What are the right choices for your application? How do you make sense of all these buzzwords?

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.

STREAM PROCESSING

Bachelor in Data and Business Analytics BDBA SEP-2023
SP-DBA.1C.4.M.A

Area Others

Number of sessions: 25

Academic year: 23-24

Degree course: FOURTH

Number of credits: 5.0

Semester: 1º

Category: OPTIONAL

Language: English

Professor: **EDUARDO RODRÍGUEZ LORENZO**

E-mail: erodriguezl@faculty.ie.edu

EDUARDO RODRÍGUEZ LORENZO

Eduardo Rodriguez Lorenzo is Senior Manager at NETSCOUT and Adjunct Professor at IE School of Science and Technology. He is a technologist specializing in Telecommunication Networks, Cybersecurity, Software Architecture, Data Engineering and Analytics.

He studied at UPM (Universidad Politécnica de Madrid), King's College London and London University.

At NETSCOUT, he leads a global team of Data and Network Engineers with a strong focus on Network Service Assurance, Cybersecurity, Data Engineering and Analytics.

He has gained broad international experience delivering high-value Consulting Services (Customer Experience & Customer Journeys, Business Intelligence, Service Assurance, Data Monetization, Process Engineering...) and Data-driven Solutions (Cloud & Backend Architecture, Data Feeds, Database, Dashboard, Interaction & Visualisation Design) to global Enterprises and Communication Service Providers. He has played an active role in the launch, measurement and optimisation of Mobile Networks for various top international Telcos.

He is a member of the Spanish Charter of Telecommunications Engineers (COIT) where he is an active member of the Telecommunications Policy and Regulation Group and the Digital Transformation Group.

His main interests include Disruptive Technologies, Data Engineering Architectures, Networks, Distributed Systems and Graph technology.

He joined IE University in 2020.

IE | LinkedIn | Twitter

erodriguezl@faculty.ie.edu

PREREQUISITES

- Algorithms and Data Structures
- Basic SQL skills
- Intermediate Python skills
- Basic Linux command line

SUBJECT DESCRIPTION

Welcome to the "Stream Processing" elective course! In this 25-session course, you will gain the skills and knowledge required to design and develop streaming data processing systems. This course is perfect for you if you are interested in learning about the latest trends in data engineering and stream processing, and you have introductory knowledge in programming (Python), databases (SQL) and data structures.

The course covers a wide range of topics, starting from the basics of data streams, their applications, and key use cases. You will then learn about messaging paradigms, message serialization, and probabilistic data structures for data streams and massive datasets. You will also understand the tradeoffs and key abstraction models for streaming data processing systems.

The course then dives deeper into the design of data engineering pipelines based on data streams, using windowing techniques, watermarks, and streaming SQL. You will learn how to use tools and frameworks such as Kafka and Spark Streaming to process streaming data and how to harness the power of Cloud Computing to implement such solutions. Additionally, you will understand the relevance of other NoSQL technologies in stream processing.

To provide a comprehensive understanding of stream processing, the course also covers the application of stream processing for machine learning and the visualization of data streams using tools like Grafana and Streamlit. Finally, you will learn about the basics of modern design and architecture of distributed systems.

The course emphasizes hands-on experience with a group project in topics ranging from real-time anomaly detection in IoT sensor data to stream processing for traffic congestion management.

By the end of this course, you will have a strong understanding of stream processing and the tools and frameworks used in the industry. You will also have the skills to design and develop your own streaming data processing systems.

LEARNING OBJECTIVES

1. Understand the place of Streams in the Data Engineering Landscapce, its applications and key Use Cases
2. Design tradeoffs and key abstraction models for Streaming Data Processing systems
3. Understand message serialisation and most common formats used in industry: CSV, JSON, Parquet, AVRO
4. Understand probabilistic data structures for data streams and massive datasets

5. Learn messaging paradigms for Stream Data Processing, including RPC, Queues, Message Brokers
6. Design Data Engineering Pipelines based on data streams, using Windowing Techniques, Watermarks & Streaming SQL
7. Learn to use tools and frameworks for stream processing, such as Kafka and Spark Streaming
8. Understand NoSQL technologies and their relevance in stream processing
9. Learn to apply stream processing for machine learning
10. Learn to visualize data streams
11. Provide the basics of modern Design and Architecture of Distributed 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:

Learning Activity	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	24.0 %	30.0 hours
Discussions	8.0 %	10.0 hours
Exercises in class, Asynchronous sessions, Field Work	16.0 %	20.0 hours
Group work	28.0 %	35.0 hours
Individual studying	24.0 %	30.0 hours
TOTAL	100.0 %	125.0 hours

PROGRAM

WELCOME TO STREAM PROCESSING!

SESSION 1 (LIVE IN-PERSON)

Introductions.

Expectations and what to expect.

Understand the place of Streams in the Data Engineering Landscapce, its applications and key Use Cases.

Stream-related units: Throughput and data transmission speed. Bits and bytes and data volume units

Basics of TCP/IP and the OSI model

Student Environment Setup.

SESSION 2 (LIVE IN-PERSON)

STREAMING FUNDAMENTALS

Event Time Domain vs Processing Time Domain. Triggers. Watermarks. Accumulation.

Windowing techniques: Tumbling, Hopping, Sliding, Session. Advanced Windowing.

SESSION 3 (LIVE IN-PERSON)

STREAMING FUNDAMENTALS

Event Time Domain vs Processing Time Domain. Triggers. Watermarks. Accumulation.

Windowing techniques: Tumbling, Hopping, Sliding, Session. Advanced Windowing.

DATA SERIALIZATION AND ENCODING

SESSION 4 (LIVE IN-PERSON)

Encoding and transmission of messages. Schema Evolution. Data Encoding formats. Dataflow modes.

Data Streaming Encoding Lab. Overview of general-purpose and industry specific encoding formats, e.g, JSON, Parquet, AVRO, HL7, ISO8583...

STREAM PROCESSING

SESSION 5 (LIVE IN-PERSON)

STREAMING TECHNOLOGY

Streaming Tools and Frameworks. Big Data Technology timeline. Spark Streaming Introduction and Setup.

SESSION 6 (LIVE IN-PERSON)

Spark Streaming Lab

Article: Discretized Streams: Fault-Tolerant Streaming Computation at Scale (SOSP'13, Nov. 3–6, 2013, Farmington, Pennsylvania, USA) (CED)

A Jupyter-led assignment to demonstrate *pyspark.streaming* module functionality

SESSION 7 (LIVE IN-PERSON)

Spark Streaming Lab

Article: Spark SQL: Relational Data Processing in Spark (SIGMOD'15, May 31–June 4, 2015, Melbourne, Victoria, Australia) (CED)

Multimedia Documentation: Spark Structured Streaming Programming Guide (spark.apache)

A Jupyter-led assignment to demonstrate pyspark.streaming module functionality

SESSION 8 (LIVE IN-PERSON)

Spark Structured Streaming Lab

Sample Spark Notebooks to demonstrate Spark Structured Streaming Programming model.

STREAMING ALGORITHMS AND DATA STRUCTURES

SESSION 9 (LIVE IN-PERSON)

Probabilistic Data Structures for Data Streams and Massive Datasets

Understanding probabilistic data structures for data streams and massive datasets

- Sketch Algorithms: Bloom filters, Count-Min Sketch, Approximate Quantiles on Data Streams
- Implementing probabilistic data structures with Python
- Apache DataSketches

Discussion: Use cases for probabilistic data structures

SESSION 10 (LIVE IN-PERSON)

Probabilistic Data Structures for Data Streams and Massive Datasets

Understanding probabilistic data structures for data streams and massive datasets

- Sketch Algorithms: Bloom filters, Count-Min Sketch, Approximate Quantiles on Data Streams
- Implementing probabilistic data structures with Python
- Apache DataSketches

Discussion: Use cases for probabilistic data structures

MESSAGING

SESSION 11 (LIVE IN-PERSON)

Messaging Paradigms:

- Client/Server
- Queues
- Publish/Subscribe paradigm
- RPC
- APIs

SESSION 12 (LIVE IN-PERSON)

Messaging Paradigms Lab

Using ZeroMQ and gRPC to demonstrate all messaging paradigms we learnt in the previous session.

SESSION 13 (LIVE IN-PERSON)

Azure Queues Lab

Learn about Azure Storage Queues service, its API, SDK and Use Cases.

Implement a basic streaming solution using Azure Storage Queues.

SESSION 14 (LIVE IN-PERSON)

Kafka Concepts

Terminology, Use Cases, Core APIs, Zookeeper, Topics, Brokers, Partitions, Producers, Consumers. Local Setup

SESSION 15 (LIVE IN-PERSON)

Kafka Lab

Hands-on Kafka Lab on a relevant topic of choice involving streaming data sources.

SESSION 16 (LIVE IN-PERSON)

ksqlDB: SQL on Kafka

- Fundamental concepts: streams, tables, queries, and materialized views
- Architecture and components: server, CLI, and REST API
- Use cases and benefits of ksqlDB, such as materialized caches, streaming ETL pipelines, and event-driven microservices
- Differences between streams and tables
- Query types: push queries and pull queries

CLOUD STREAM PROCESSING

SESSION 17 (LIVE IN-PERSON)

Azure Event Hubs and Azure Stream Analytics Concepts

Conceptual mapping with Kafka and key differences, API usage, integration with other Azure components.

SESSION 18 (LIVE IN-PERSON)

Azure Event Hubs and Azure Stream Analytics Lab

Setting up real world Data Stream processing pipeline in Azure Cloud.

Technologies covered: EventHub, Stream Analytics, Blob Storage and PowerBI.

REAL TIME DASHBOARDS

SESSION 19 (LIVE IN-PERSON)

Visualizing Streaming Data

Article: How to build a real-time live dashboard with Streamlit (blog.streamlit.io)

Understanding Streamlit and its relevance in data visualization

Implementing Streamlit with Python

Other visualization stacks: Azure and PowerBI, ELK stack, Grafana
Discussion: Best practices for data visualization

EMERGING TOPICS AND TRENDS

SESSION 20 (LIVE IN-PERSON)

Other NoSQL Technologies for Stream Processing

Spark Alternatives: Apache Flink
Kafka Alternatives: Apache Pulsar, Redpanda
Redis and its relevance in stream processing.
Understanding Stream Processing of graph data: Memgraph and DGraph.

SESSION 21 (LIVE IN-PERSON)

Other NoSQL Technologies for Stream Processing

Spark Alternatives: Apache Flink
Kafka Alternatives: Apache Pulsar, Redpanda
Redis and its relevance in stream processing.
Understanding Stream Processing of graph data: Memgraph and DGraph.

ARCHITECTING REAL TIME SYSTEMS

SESSION 22 (LIVE IN-PERSON)

Lambda and Kappa architectures.
Distributed Architecture Patterns.
Job Scheduling. Synchronization. Cron Jobs. Apache Airflow & Azure Batch

GROUP PROJECT PRESENTATIONS

SESSIONS 23 - 24 (LIVE IN-PERSON)

Team presentations of Group Project
List of sample topics, though students can propose their own:

- Contact Tracing
- Fraud Detection
- Predictive Maintenance
- Anomaly detection
- Smart City topics: car telemetry, environmental sensors, etc
- Sports Analytics
- Health Monitoring

Student group will conduct a live presentation and upload accompanying materials to Campus (code, Notebooks, Slides...).

SESSION 25 (LIVE IN-PERSON)

Final Exam

EVALUATION CRITERIA

Knowledge Quizzes

There will be several in-class quizzes to support ongoing learning and understanding of core concepts.

Final Exam

At the end of the course you will have to pass an individual exam which may contain questions about any of the topics covered during this course.

Labs and Individual Practice

During the course I will propose labs and practical exercises to be solved individually

Group Assignment

A Group Project will be developed by teams of 4 to 6 students, who will have to work on a real life Use Case involving Stream Processing. The team will play the role of a Chief Technology Officer who puts together the best solution taking into account technical, business and environmental constraints.

A list of topics will be proposed by the professor, but students are free to propose their own subjects.

criteria	percentage	Learning Objectives	Comments
Final Exam	30 %		
Knowledge Quizzes	25 %		
Individual Work	15 %		
Group Assignment	30 %		

RE-SIT / RE-TAKE POLICY

A minimum passing grade of 4 (40%) in the Final exam is required to pass the subject. If a student scores lower than this minimum, he will have to go to June retake, irrespective of their overall course grade. The overall passing course grade is 5.0.

BIBLIOGRAPHY

Recommended

- Martin Kleppmann. (2017). *Designing Data-Intensive Applications.* O'Reilly Media, Inc.. ISBN 9781449373320 (Digital)

Data is at the center of many challenges in system design today. Difficult issues need to be figured out, such as scalability, consistency, reliability, efficiency, and maintainability. In addition, we have an overwhelming variety of tools, including relational databases, NoSQL datastores, stream or batch processors, and message brokers. What are the right choices for your application? How do you make sense

of all these buzzwords?

- Tyler Akidau, Slava Chernyak, Reuven Lax. *Streaming Systems.* 2018. O'Reilly Media, Inc.. ISBN 9781491983874 (Digital)

Streaming data is a big deal in big data these days. As more and more businesses seek to tame the massive unbounded data sets that pervade our world, streaming systems have finally reached a level of maturity sufficient for mainstream adoption. With this practical guide, data engineers, data scientists, and developers will learn how to work with streaming data in a conceptual and platform-agnostic way.

- Emin Tahirovic, Dzejla Medjedovic, Ines Dedovic. (2022). *Algorithms and Data Structures for Massive Datasets.* Manning Publications. ISBN 9781617298035 (Digital)

- Joe Reis, Matt Housley. (2022). *Fundamentals of Data Engineering.* O'Reilly Media, Inc.. ISBN 9781098108304 (Digital)

- Jay Kreps. (2014). *I Heart Logs.* O'Reilly Media, Inc.. ISBN 9781491909386 (Digital)

Why a book about logs? That’s easy: the humble log is an abstraction that lies at the heart of many systems, from NoSQL databases to cryptocurrencies. Even though most engineers don’t think much about them, this short book shows you why logs are worthy of your attention.

- Piethein Strengholt. (2020). *Data Management at Scale.* O'Reilly Media, Inc.. ISBN 9781492054788 (Digital)

As data management and integration continue to evolve rapidly, storing all your data in one place, such as a data warehouse, is no longer scalable. In the very near future, data will need to be distributed and available for several technological solutions. With this practical book, you’ll learn how to migrate your enterprise from a complex and tightly coupled data landscape to a more flexible architecture ready for the modern world of data consumption.

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.

UNIVERSITY

LAW SCHOOL

LAW

ELECTIVES

INTERNATIONAL BUSINESS LAW

- CAPITAL MARKETS LAW
- CROSS BORDER M&A
- MOOT MADRID I
- INTERNATIONAL ARBITRATION
- PRACTIVAL AI&LAW
- RESTRUCTURING AND INSOLVENCY LAW
- MOOT MADRID II

LAW SCHOOL

TECHNOLOGY & DIGITAL LAW

- INTERNATIONAL LAW AND POLITICS OF CYBER OPERATIONS
- INTRO TO IP LAW AND TECHNOLOGY ISSUES
- INNOVATION IN THE LEGAL MARKET- LAW WITHOUT WALLS COMPETITION
- EMERGING TECHNOLOGIES & THE QUESTION OF ETHICS
- CYBERSECURITY
- PRACTICAL DATA PRIVACY & TECHNOLOGY

ie LAW SCHOOL

HUMAN RIGHTS & SOCIAL JUSTICE

- **EUROPEAN HUMAN RIGHTS MOOT ****
- **CLIMATE CHANGE AND THE POLITICS OF INTERNATIONAL LAW**
- **SEX, GENDER AND THE LAW**
- **MIGRATION AND INTERNATIONAL JUSTICE**
- **ENTERPRISES AND HUMAN RIGHTS**
- **GENDER DISCRIMINATION AND HUMAN RIGHTS**
- **MONEY LAUNDERING & TERRORISM FINANCING**

ie LAW SCHOOL

UNIVERSITY

LAW SCHOOL

MOOTS GUIDELINES

European Human Rights Moot Court Competetion

Professors: Amaya Ubeda and Alice Thomas

3 ECTS Fall Elective

This course will serve as preparation for the European Law Students Association (ELSA) Moot Court Competition (http://ehrmcc.elsa.org).

Students will be asked to solve a hypothetical, complex case on European Human Rights Law and argue it as if they were pleading before the European Court of Human Rights.

If IE University is selected to go to the final (oral) round, the top three or four students in the elective course will have the opportunity to represent IE University in the 2024 ELSA Moot Court Competition in Strasbourg.

The Final is organised annually in Strasbourg with the pleadings taking place at the Palais de l'Europe and the European Court of Human Rights

This year´s IE Students competing at the finals in Strasbourg where they were awarded the prize for the best respondent's written submission

EUROPEAN HUMAN RIGHTS MOOT COURT COMPETITION

Professor: Antonios Kouroutakis

3 ECTS Fall Elective

Through the Jessup Moot elective,
you will prepare for a Moot Court Competition on International Law. This is the largest Moot Court Competition with participants from 700 different law schools located in 100 different countries and jurisdictions.

The competition is a simulation of a fictional
dispute between countries before the International Court of Justice, the juridical organ of the United Nations. Teams prepare oral and written pleadings arguing both the applicant and respondent positions of the case against competing teams and before a panel of judges, simulating a proceeding before the International Court of Justice.

Click HERE for more information.

Jessup Moot Court Competition

Student Perspective

Fiona Wu
2022 Finalist

VIDEO LINK

JESSUP COMPETITION

Professor: Charlotte Leskinen

3 ECTS Fall Elective

In this course, students will be asked to solve a hypothetical, complex, topical EU law case and argue them as if they were pleading before the Court of justice of the European Union.

In addition, the four best students will have the opportunity to represent IE University in the 2024 Europa Moot Court Competetion (location to be confirmed).

Student Perspective

Patricia Muñoz

2021 Participant

VIDEO LINK

Europa Moot Court Competetion

IEU Team 2018
2nd place

IEU Team 2019
WINNERS

EMCC
EUROPA MOOT COURT COMPETITION

Moot Madrid Moot Court Competition

Professor: Oliver Cojo

3 ECTS Fall Elective

A hypothetical case on international arbitration. The competition takes place in Madrid and is designed for Spanish speakers.

La Competición Internacional de Arbitraje y Derecho mercantil tiene diversos objetivos que giran en torno a la formación de los estudiantes de Derecho en cuestiones relativas al Derecho Uniforme del Comercio Internacional y su resolución mediante el Arbitraje Mercantil Internacional.

For more information click HERE

Student Perspective

Teresa Olombrada Rodríguez

VIDEO LINK

MOOTmadrid

Professor: Erika Pagano

3 ECTS Fall Elective

A part-virtual, experiential learning initiative designed for practicing and aspiring legal and business professionals that brings a human-centred design perspective to law.

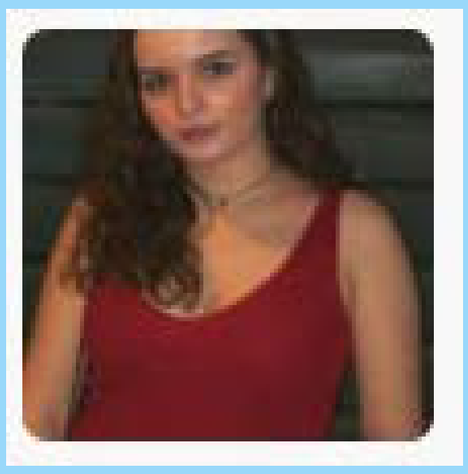
LWOW unites law and business professionals with students from 30 law and business schools around the world to co-create innovative solutions to problems at the intersection of law, business and technology while developing skills essential to all professionals today.

Leveraging intergenerational, cross-cultural, multidisciplanary exchange, LWOW changes mindsets, skillsets and behaviours.

For more information click HERE

Law Without Walls Moot Court Competition

Student Perspective



Alessia Galeotalanza
2022 Finalist

VIDEO LINK

LWOW
LAWWITHOUTWALLS®
THE FUTURE OF LAW, TODAY

Comparative Law in Action Moot Court Competition

3 ECTS Spring Elective

Students from top-tier universities across the globe will explore the intersections of Comparative Law, Technology & Policy by collaborating on a cutting-edge multimedia case.

Patricipants will work together virtually to gain experience and insight on how to manage legal cases in a technologically-driven world.

The competition and course will culminate in a 3-day experiential weekend in Madrid.

Check out the links below to see the previous cases and the experiential weekend from 2022.

2023 Case

2022 Case

2022 Experiential Weekend

Previous Participating Universities

Università Commerciale Luigi Bocconi

BUCERIUS LAW SCHOOL
HOCHSCHULE FÜR RECHTSWISSENSCHAFT

THE UNIVERSITY of EDINBURGH
Edinburgh Law School

FGV DIREITO RIO

ie LAW SCHOOL

Maastricht University

TILBURG UNIVERSITY

ie COMPARATIVE LAW IN ACTION