esade

TEACHING GUIDE - 2024-2025

Machine Learning

UGRA_015032

Departments	Dept. of Operations, Innovation & Data Sciences
Teaching Languages	English
ECTS	6
Teacher responsible	Bagherzade Ghazvini Mina - mina.bagherzade@esade.edu

Course Goals

By the end of the course, students will have acquired the following knowledge and skills:

· **Comprehensive understanding of advanced machine learning concepts**: Students will gain in-depth knowledge of ensemble learning techniques, such as bagging and boosting, as well as unsupervised learning methods like clustering and dimensionality reduction.

· **Practical skills in implementing and optimizing algorithms**: Students will develop hands-on experience in implementing and fine-tuning machine learning algorithms, including Random Forest, AdaBoost, K-means, or PCA.

· **Competence in applying machine learning techniques to real-world problems**: Students will learn to create robust predictive models and improve decision-making and strategic planning in business environments.

· **Ability to handle complex coding tasks and understand practical implications**: Students will be prepared to tackle advanced coding challenges and apply machine learning solutions effectively in industrial and financial settings.

Previous knowledge

Students should have a foundational understanding of machine learning concepts, algorithms, and techniques. This includes familiarity with key concepts such as classification and regression, as well as model training and evaluation, error measures, and performance metrics. A strong background in statistics and probability theory is also expected, as well as proficiency in linear algebra and calculus, especially in differentiation and optimization techniques; and familiarity with discrete mathematics. Additionally, since this will be a hands-on course, proficiency in at least one programming language is essential, with a preference for Python. Students should be comfortable writing and debugging code, and be familiar with data preprocessing, cleaning, and exploratory data analysis (EDA) techniques.

Prerequisits

This course will be managed through a dedicated eCampus website. Students will find there all the necessary class materials, including assigned readings and pre-class work, deliverables and exercises, and further references. Students should familiarize themselves with this environment before the start of the course and check for updates regularly.

Teaching methodology

This course will combine a theoretical description of topics and a practical implementation. It will emphasize hands-on experience through practical assignments and projects. Students will apply the concepts learned in the course to real-world problems. This will help them understand the practical implications and challenges of implementing machine learning solutions in the industry.

Description

Course contribution to program

This course is a follow-up of the Introduction to Artificial Intelligence course. It builds upong students' previous knowledge on data structures and algorithmic techniques and computational problem solving to explore additional topics on machine learning, including ensemble and unsupervised learning. Understanding ensemble learning techniques like boosting and bagging will help students grasp advanced topics in deep learning, such as how ensemble methods can be used to improve neural network performance and reduce overfitting. Knowledge of clustering methods and dimensionality reduction techniques will also aid them in understanding unsupervised learning approaches in deep learning, such as autoencoders. Additionally, gaining practical experience with machine learning algorithms will be beneficial for transitioing to more complex coding tasks and frameworks. Finally, familiarity with clustering and dimensionality reduction techniques can be directly applied to market segmentation, customer analysis, and other business-oriented tasks, therefore providing students with a wider perspective on the uses of AI for business.

Short description

This course is designed to deepen students' understanding of machine learning techniques, building on foundational knowledge acquired in previous courses. It focuses on two major areas: **ensemble learning and unsupervised learning**. Students will explore bagging, boosting, clustering, and dimensionality reduction, among others. These techniques are crucial for developing robust predictive models that can improve decision-making and strategic planning in business environments. Practical exercises will enable students to implement and optimize these algorithms, reinforcing their coding skills, understanding of model performance, and ability to interpret results in real-world applications. By the end of the course, students will have a comprehensive understanding of advanced machine learning concepts, preparing them for deep learning topics and equipping them to tackle complex problems in industrial and financial settings using state-of-the-art techniques.

Bibliography

HASTIE, Trevor, TIBSHIRANI, Robert & FRIEDMAN, Jerome, Unsupervised Learning, Springer, https://doi.org/10.1007/978-0-387-84858-7_14 (Book Chapter)
GREENE, Derek Greene, CUNNINGHAM, Pádraig & MAYER, Rudolf, Unsupervised Learning and Clustering, Springer, https://doi.org/10.1007/978-3-540-75171-7_3 (Book Chapter)
PROVOST, Foster & FAWCETT, Tom, Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking, O'Reilly Media, 9781449374266 (Book)
PATEL, Ankur, Hands-On Unsupervised Learning Using Python How to Build Applied Machine Learning Solutions from Unlabeled Data, O'Reilly Media, 978-1492035640 (Book)

Content

#	Topic
1	Introduction to Ensemble Learning. Students will be introduced to the fundamental concepts of ensemble learning, explaining its importance and the basic idea behind combining multiple models to improve overall performance. Key concepts such as bias-variance trade-off, model diversity, and voting mechanisms will be covered.
2	Bagging Techniques. Students will delve into bagging (Bootstrap Aggregating) methods, with a focus on Random Forest Classifier (RFC). They will learn how bagging reduces variance and improves model stability, and complete practical exercises on implementing and tuning bagging algorithms.
3	Boosting Methods. The course will also cover boosting techniques, emphasizing algorithms such as AdaBoost, Gradient Boosting, and XGBoost.. Students will understand how boosting sequentially combines weak learners to form a strong learner, and will engage in practical exercises to implement and optimize boosting models.

#	Topic
4	Introduction to Clustering. Students will be introduced to the basic principles of clustering, explaining the differences between various clustering methods. Students will learn about the importance of clustering in unsupervised learning and its applications.
5	K-means and Hierarchical Clustering. Students will explore K-means and hierarchical clustering techniques in detail. They will learn about the algorithms, their implementation, and methods to choose the number of clusters. Practical exercises will help students apply these methods to real datasets.
6	Density-Based Clustering. The course will also cover Density-Based Spatial Clustering of Applications with Noise (DBSCAN). Students will understand the algorithm's working principles, advantages, and limitations. Practical exercises will include implementing DBSCAN and tuning its parameters.
7	Dimensionality Reduction Techniques. Students will learn about dimensionality reduction methods, including Principal Component Analysis (PCA) and/or t-Distributed Stochastic Neighbor Embedding (t-SNE). They will cover the theory behind these techniques, their applications, and practical implementation exercises to handle high-dimensional data.

Assessment

Tool	Assessment tool	Category	Weight %
Written and/or oral exams	Final exam. Students will sit a final exam, covering all the topics of the course.	Ordinary round	30.00%
Written and/or oral exams	Mid-term exam. Students will sit a mid-term exam to evaluate their understanding of the topics discussed during the first half of the course.	Ordinary round	20.00%
Attendance and punctuality	Attendance. In accordance with ESADE regulations, attendance is mandatory for this course. Students who fail to attend 80% of the course will not be allowed to pass and will be required to sit the retake exam.	Ordinary round	0.00%
Individual or team exercises	Assignments. Students will complete submit several deliverables, including invidual class exercises and/or group projects.	Retake and ordinary round	35.00%
In-class analysis and discussion of issues	Active contribution. Students will be expected to actvely participate in class and to engage in the different course activities and debates.	Ordinary round	15.00%
Written and/or oral exams	Retake exam. Students who fail to pass the course in the ordinary sitting will be required to sit a retake exam that will contribute a 65% of their final course grade.	Retake	65.00%

PROGRAMS

DBAI21-Double Degree in Business Administration and Artificial Intelligence for Business (Undergraduates: Business)
DBAI21 Year 3 (Mandatory)

DBAI23-Double Degree in Business Administration and Artificial Intelligence for Business (Undergraduates: Business)
DBAI23 Year 3 (Mandatory)