SYLLABUS

1. Information regarding the programme

1.1 Higher education institution	**Babeş-Bolyai University of Cluj-Napoca**
1.2 Faculty	**Faculty of Environmental Science and Engineering**
1.3 Department	**Department of Environmental Analysis and Engineering**
1.4 Field of study	**Environmental Engineering**
1.5 Study cycle	**Bachelor**
1.6 Study programme / Qualification	**Environmental engineering**

2. Information regarding the discipline

2.1 Name of the discipline	**Machine Learning Applications for Environmental Science**						
2.2 Course coordinator	**CS III PhD Mereuță Alexandru**						
2.3 Seminar coordinator	**CS III PhD Mereuță Alexandru**						
2.4. Year of study	**3**	2.5 Semester	**6**	2.6. Type of evaluation	**E**	2.7 Type of discipline	**Opt.**

3. Total estimated time (hours/semester of didactic activities)

3.1 Hours per week	4	Of which: 3.2 course	2	3.3 seminar/laboratory	2
3.4 Total hours in the curriculum	56	Of which: 3.5 course	28	3.6 seminar/laboratory	28

Time allotment:	hours
Learning using manual, course support, bibliography, course notes	18
Additional documentation (in libraries, on electronic platforms, field documentation)	10
Preparation for seminars/labs, homework, papers, portfolios and essays	10
Tutorship	2
Evaluations	2
Other activities:	-

3.7 Total individual study hours	42
3.8 Total hours per semester	98
3.9 Number of ECTS credits	4

4. Prerequisites (if necessary)

4.1. curriculum	-
4.2. competencies	-

5. Conditions (if necessary)

5.1. for the course	Video projector
5.2. for the seminar /lab activities	Laboratory with computers;

6. Specific competencies acquired

Professional competencies	• Understand fundamental ML concepts and techniques. • Be able to preprocess data and build ML models. • Gain hands-on experience with Python's ML libraries. • Choose the best model for Environmental ML aplications. • Develop an understanding of ethical considerations in ML applications. • Have a foundation for exploring advanced topics in ML and deep learning.

Transversal competencies	• Creativity and innovation. • Communication attitudes and work in interdisciplinary teams. • Ability to interpret, analyze and synthesize data available in the analysis of complex situations • Computational and data literacy. • Critical thinking and decision-making.

7. Objectives of the discipline (outcome of the acquired competencies)

7.1 General objective of the discipline	• Develop foundational knowledge in Machine Learning • Training, testing, validating and deploying ML for environmental science applications
7.2 Specific objective of the discipline	• Understanding the role of machine learning in analyzing and modeling environmental systems • Developing skills in preprocessing and analyzing environmental datasets • Applying machine learning techniques to solve real-world environmental challenges • Designing and implementing ethical and sustainable machine learning models

8. Content

8.1 Course	Teaching methods	Remarks
1.Course Overview and Introduction to Machine Learning • What is Machine Learning? • Applications of ML in environmental science. • Types of ML: Supervised, Unsupervised, Semi-supervised, and Reinforcement Learning.	Interactive exposure Explanation Conversation	
2. Data Preprocessing and Exploration • Importance of data quality. • Handling missing data, outliers, and categorical variables. • Exploratory Data Analysis (EDA) techniques.	Interactive exposure Explanation Conversation	
3. Supervised Learning I - Regression • Linear regression. • Polynomial regression. • Performance metrics (MAE, MSE, R^2).	Interactive exposure Explanation Conversation	
4. Supervised Learning II – Classification • Logistic regression. • Decision trees. • Performance metrics (accuracy, precision, recall, F1-score, ROC-AUC).	Interactive exposure Explanation Conversation	
5.Feature Engineering • Feature selection vs. extraction. • Principal Component Analysis (PCA). • Feature importance and interpretability.	Interactive exposure Explanation Conversation	
6. Unsupervised Learning I – Clustering • k-Means clustering. • Hierarchical clustering. • Evaluating clustering (silhouette score, inertia).	Interactive exposure Explanation Conversation	
7.Modeling the dispersion of air pollutants • Curse of dimensionality. • Techniques: PCA, t-SNE.	Interactive exposure Explanation Conversation	

8. Model Evaluation and Tuning • Cross-validation techniques. • Hyperparameter tuning (grid search, random search). • Avoiding overfitting and underfitting.	Interactive exposure Explanation Conversation	
9.Ensemble Learning • Bagging (Random Forests). • Boosting (Gradient Boosting, AdaBoost, XGBoost). • Stacking models.	Interactive exposure Explanation Conversation	
10. Introduction to Deep Learning • Neural networks basics. • Forward and backward propagation. • Activation functions.	Interactive exposure Explanation Conversation	
11. Advanced Topics and Trends in ML • Transfer learning. • Explainable AI (XAI). • Overview of reinforcement learning.	Interactive exposure Explanation Conversation	
12. Ethical and Responsible AI • Bias and fairness in ML models. • Data privacy concerns. • Ethical considerations in deploying ML systems.	Interactive exposure Explanation Conversation	
13. Course Review and Project Work • Revisiting key concepts. • Open Q&A session. • Project guidance.	Interactive exposure Explanation Conversation	
14. Project Presentations and Course Wrap-Up • Final project presentations. • Peer reviews and feedback. • Course wrap-up.	Evaluation Conversation	

Bibliography

1. Cambridge University Press , 2023, Dive into Deep Learning; https://d2l.ai/index.html
2. O'Reilly Media, 2019, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd Edition;
3. Packt Publishing, 2022, Machine Learning with PyTorch and Scikit-Learn, Develop machine learning and deep learning models with Python;
4. Elsevier Science, 2023, Artificial Intelligence in Earth Science, Best Practices and Fundamental Challenges;
5. Wiley, 2021, Deep Learning for the Earth Sciences: A Comprehensive Approach to Remote Sensing, Climate Science and Geosciences;
6. Springer, 2022, Introduction to Python in Earth Science Data Analysis From Descriptive Statistics to Machine Learning;
7. Springer, 2021, Python Recipes for Earth Sciences;
8. Springer, 2023, Machine Learning for Earth Sciences, Using Python to Solve Geological Problems;
9. The MIT Press, 2020, Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies, 2nd Edition;

8.2 Seminar / laboratory	Teaching methods	Remarks
1. Setting up the environment (Anaconda, Jupyter Notebook, essential Python libraries). First hands-on: Linear regression example using scikit-learn.	Hands-on coding Explanation Brainstorming	
2. Data preprocessing with pandas and NumPy. EDA with Matplotlib and Seaborn.	Hands-on coding Explanation Brainstorming	
3. Implementing regression models on real-world environmental datasets. Feature scaling and normalization.	Hands-on coding Explanation Brainstorming	
4. Classification on datasets (e.g., land cover classification, species identification). Evaluating models using confusion matrix and ROC curves.	Hands-on coding Explanation Brainstorming	
5. Applying PCA and feature selection techniques. Using scikit-learn and feature importance visualizations.	Hands-on coding Explanation Brainstorming	
6. Implementing clustering techniques on environmental datasets (e.g., climate zones, pollution data).	Hands-on coding Explanation Brainstorming	
7. Visualizing high-dimensional data using t-SNE. Applying dimensionality reduction to remote sensing data.	Hands-on coding Explanation Brainstorming	
8. Using scikit-learn's GridSearchCV. Regularization techniques: L1 (Lasso), L2 (Ridge).	Hands-on coding Explanation Brainstorming	
9. Implementing ensemble methods on environmental datasets (e.g., deforestation analysis, flood prediction).	Hands-on coding Explanation Brainstorming	
10. Building a simple neural network using TensorFlow/Keras. Training and evaluating the network on environmental datasets (e.g., rainfall prediction, air quality indices).	Hands-on coding Explanation Brainstorming	
11. Fine-tuning a pre-trained model (e.g., for land use classification).	Hands-on coding Explanation Brainstorming	
12. Auditing a model for bias. Discussing case studies of ethical dilemmas in environmental AI applications.	Hands-on coding Explanation Brainstorming	
13. Group/individual project implementation.	Hands-on coding Explanation Brainstorming	
14. Group/individual project presentation.	Hands-on coding Explanation Presentation	

Bibliography:

1. (2023). *ARSET - Fundamentals of Machine Learning for Earth Science*. NASA Applied Remote Sensing https://appliedsciences.nasa.gov/get-involved/training/english/arset-fundamentals-machine-learning-earth-science
2. (2024). *ARSET - Large Scale Applications of Machine Learning using Remote Sensing for Building Agriculture Solutions*. NASA Applied Remote Sensing Training Program (ARSET). https://appliedsciences.nasa.gov/get-involved/training/english/arset-large-scale-applications-machine-learning-using-remote-sensing
3. Github repository, 2022, Awesome-Earth-Artificial-Intelligence https://github.com/ESIPFed/Awesome-Earth-Artificial-Intelligence
4. Github repository, 2024, Awesome Open Geoscience https://github.com/softwareunderground/awesome-open-geoscience

9. Corroborating the content of the discipline with the expectations of the epistemic community, professional associations and representative employers within the field of the program

The content structure of the discipline was structured by studying the recent monographs in the field and by consulting the programs and the available notes from some recognized institutions in the field.

10. Evaluation

Type of activity	10.1 Evaluation criteria	10.2 Evaluation methods	10.3 Share in the grade (%)
10.4 Course	• The correctness and completeness of the accumulated knowledge.	Oral exam	30%
10.5 Seminar/lab activities	• Project/ essay	Evaluation of the project (documentation and demonstration)	60%*
*1 point will be added to the final grade ex officio.			

10.6 Minimum performance standards
Each student has to prove that (s)he acquired an acceptable level of knowledge and understanding of the:
- Understanding the types of ML models used in environmental applications;
- Running a tailored made ML model for any given environmental dataset;
- Running exploratory data analysis and understanding the importance of data quality.

The student will prepare and sustain a project / essay according to the content of the framework.
Obtaining the minimum mark of 5 is an entry condition for the Exam.

Date	Signature of course coordinator	Signature of seminar coordinator
15.11.2024	CS III PhD Mereuță Alexandru	CS III PhD Mereuță Alexandru

Date of approval

Signature of the head of department