GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

Subject name and code	MACHINE LEARNING, PG_00068305						
Field of study	Economic Analytics						
Date of commencement of studies	October 2025		Academic year of realisation of subject		2026/2027		
Education level	second-cycle studies		Subject group		Optional subject group Specialty subject group Subject group related to scientific research in the field of study		
Mode of study	Full-time studies		Mode of delivery		at the university		
Year of study	2		Language of instruction		Polish		
Semester of study	3		ECTS credits		4.0		
Learning profile	general academic profile		Assessment form		assessment		
Conducting unit	Department Of Statistics And Econometrics -> Faculty Of Management And Economics -> Wydziały Politechniki Gdańskiej						
Name and surname of lecturer (lecturers)	Subject supervisor		dr inż. Karol Flisikowski				
	Teachers						
Lesson types and methods of instruction	Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
	Number of study hours	0.0	0.0	30.0	0.0	0.0	30
	E-learning hours included: 0.0						
Learning activity and number of study hours	Learning activity	Participation in didactic classes included in study plan		Participation in consultation hours	Self-study		SUM
	Number of study hours	30		5.0	65.0		100
Subject objectives	The objective of the course is to introduce students to the fundamental concepts, techniques, and algorithms used in machine learning for data analysis, prediction, and decision-making. Students will acquire both theoretical knowledge and practical skills in applying supervised and unsupervised learning methods, data preprocessing, model validation, and performance evaluation. The course emphasizes understanding the machine learning workflow, interpreting models, and applying them to real-world problems across various domains.						

Learning outcomes

Course outcome	Subject outcome	Method of verification
[K7_W03] demonstrates in-depth knowledge of the applications of analytical methods and techniques for formulating and solving socio-economic problems.	The student has advanced knowledge of data analysis methods and machine learning algorithms, enabling the identification and modeling of complex and unstructured processes. They are familiar with modern tools and technologies used for processing data from heterogeneous sources.	[SW3] Assessment of knowledge contained in written work and projects
[K7_U01] creates innovative solutions for complex and unstructured processes, considering unpredictable environmental conditions through the synthesis of information from various sources.	The student is able to design and implement machine learning models to analyze complex processes under changing environmental conditions. They can assess the relevance of various data sources and integrate them to obtain consistent and accurate conclusions.	[SU1] Assessment of task fulfilment [SU4] Assessment of ability to use methods and tools [SU2] Assessment of ability to analyse information

Subject contents	
	1. **Introduction to Machine Learning** • Definition and history of machine learning • Differences between artificial intelligence, machine learning, and deep learning • Application areas of machine learning (image recognition, text analysis, predictions, etc.) • Main categories of ML algorithms: supervised, unsupervised, reinforcement learning 2. **Mathematical Foundations of Machine Learning** • Introduction to linear algebra (matrices, vectors, operations) • Statistics: central moments, probability distributions, estimation • Concepts related to optimization (cost functions, gradient descent) 3. **Data Preparation** • Basics of data preprocessing: cleaning, missing value imputation, normalization, standardization • Feature transformation: encoding categorical variables, feature engineering • Data splitting: training, validation, and test sets • Challenges with large datasets (Big Data) 4. **Supervised Learning** • **Regression**: linear, polynomial, logistic regression • **Classification**: Naive Bayes, k-NN, decision trees, SVM • **Ensemble models**: Random Forest, Gradient Boosting (XGBoost, LightGBM, CatBoost) • **Neural Networks (MLP)** • Optimization: cross-validation, regularization (L1, L2), k-fold validation 5. **Unsupervised Learning** • **Clustering**: K-Means, DBSCAN, hierarchical clustering • **Dimensionality Reduction**: PCA, t-SNE, LDA • **Dependency Analysis**: principal component analysis, factor analysis

6. **Deep Learning**
   - Introduction to neural networks
   - Structure and operation of perceptrons
   - Convolutional Neural Networks (CNN) for image analysis
   - Recurrent Neural Networks (RNN) and LSTMs for sequential data
   - Transfer learning and fine-tuning
   - Introduction to libraries: TensorFlow, Keras, PyTorch
7. **Model Optimization and Hyperparameter Tuning**
   - Hyperparameter selection (Grid Search, Random Search, Bayesian Optimization)
   - Regularization and techniques to prevent overfitting (dropout, early stopping)
   - Cross-validation and k-fold validation
   - Error analysis and performance metrics (RMSE, MAE, AUC-ROC, F1-score)
8. **Practical Applications of Machine Learning**
   - **Image Recognition**: image classification using CNNs
   - **Natural Language Processing (NLP)**: text analysis, word embeddings (Word2Vec, GloVe)
   - **Recommendation Systems**: collaborative filtering, content-based filtering
   - **Prediction**: time series forecasting and trend analysis
   - **Anomaly Detection and Fraud Detection**: identifying outliers in data
9. **Ethics and Responsibility in Machine Learning**
   - Bias issues in data (data bias)
   - Ethical use of algorithms in various industries
   - Model transparency and interpretability
   - Responsibility for algorithm-based decision-making (e.g., in healthcare, finance, justice)

	10. **Practical Workshops and Projects** • Implementation of selected ML algorithms in Python (scikit-learn, pandas, numpy, matplotlib) • Case study analysis • Working with real-world datasets (e.g., social media data, financial data, images) • Mini-project: model development and presentation of results based on a custom dataset
Prerequisites and co-requisites	Descriptive statistics, mathematical statistics, fundamentals of programming in R/Python.

Assessment methods and criteria

Subject passing criteria	Passing threshold	Percentage of the final grade
Tests	50.0%	10.0%
Report in Markdown	50.0%	40.0%
Final project	50.0%	50.0%

Recommended reading

Basic literature	1. **Géron, A. (2022).** *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems.* OReilly Media. 2. **Raschka, S., & Mirjalili, V. (2019).** *Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, Keras, and TensorFlow.* Packt Publishing. 3. **VanderPlas, J. (2019).** *Python Data Science Handbook: Essential Tools for Working with Data.* OReilly Media.
Supplementary literature	1. Chollet, F. (2025). *Deep learning with Python* (3rd ed.). Manning Publications. 2. Chen, S., Zhang, H., & Li, J. (2024). *Deep learning and machine learning Python data structures and mathematics fundamentals: From theory to practice*. Springer.
eResources addresses	Podstawowe https://github.com/ageron/handson-ml3 - A series of Jupyter notebooks that walk you through the fundamentals of Machine Learning and Deep Learning in Python using Scikit-Learn, Keras and TensorFlow 2. https://github.com/ageron/handson-ml3 - Python Machine Learning (3rd Ed.) Code Repository Uzupełniające Adresy na platformie eNauczanie:

Example issues/ example questions/ tasks being completed	1. **Fundamentals of Machine Learning:** • Classification vs. Regression: Whats the difference and when to use them? • Use cases for various classification algorithms (e.g., decision trees, SVM, KNN). • Basic optimization techniques in machine learning (e.g., gradient descent). 2. **Data Preparation:** • Data exploration and cleaning (missing values, outliers). • Normalization, standardization, and feature engineering. • Feature selection and dimensionality reduction techniques (PCA, LDA). 3. **Deep Learning:** • Fundamentals of neurons and neural networks. • Neural network architectures: CNN, RNN, GANs. • Overfitting and regularization in deep learning (dropout, L2 regularization). 4. **Modeling and Evaluation:** • Cross-validation and validation techniques. • Confusion matrix, Precision, Recall, F1-score. • Hyperparameter tuning and model selection. **Questions:** 1. What are the key differences between traditional machine learning algorithms and deep learning methods? 2. What techniques can be applied to prevent overfitting in deep learning models? 3. How can you evaluate the effectiveness of a classification model in the context of imbalanced classes? 4. What is gradient descent, and how does it impact the training process of models? **Tasks:**

	1. **Task 1:** • Use Scikit-learn to perform classification on a dataset using the KNN algorithm. Compare the results with other algorithms, such as SVM or decision trees. 2. **Task 2:** • Apply PCA (Principal Component Analysis) for dimensionality reduction on a dataset (e.g., Iris dataset) and evaluate the effectiveness of this method on classification results. 3. **Task 3:** • Build a neural network model using Keras/TensorFlow for image classification (e.g., CIFAR-10 dataset). Compare the results with traditional machine learning algorithms. 4. **Task 4:** • Perform data exploration and cleaning (handling missing values, outliers) on a customer dataset. Then, build a predictive model for customer churn. 5. **Task 5:** • Apply regularization methods (L2, Dropout) to a deep learning model for an image classification task. Evaluate the effectiveness in preventing overfitting.
Work placement	Not applicable