

HARMONOGRAM KURSU DATA SCIENCE 02.09.2024

Opiekun merytoryczny - Norbert Ryciak

Dzień	Terminy	MODUŁ	Zagadnienia
1	02.09.2024	Programowanie w języku Python	* Podstawy Pythona: typy danych, mutowalność, kopiowanie a referencja, instrukcja sterujące, funkcje
2	03.09.2024		* Środowisko Jupyter
3	05.09.2024		* Wyrażenia listotwórcze * zaawansowane aspekty implementacji funkcji (wartości domyślne, arg, kwargs) * Wyrażenie lambda * Wyrażenia regularne
4	06.09.2024		* Obsługa danych czasowych * Programowanie obiektowe * Obsługa wyjątków
5	09.09.2024	Analiza danych w języku Python	* Środowisko programistyczne PyCharm
6	10.09.2024		* Modularyzacja kodu * Dokumentowanie kodu * Bash * System kontroli wersji Git
7	12.09.2024		* Obliczenia numeryczne - biblioteka numpy * Wizualizacja danych - biblioteka matplotlib
8	13.09.2024		* Przetwarzanie i analiza danych tabelarycznych - biblioteka pandas
9	16.09.2024	Uczenie maszynowe	* Przetwarzanie danych różnych formatów * Podstawy języka SQL
10	17.09.2024		* Podstawy rachunku prawdopodobieństwa * Statystyka: problem estymacji
11	19.09.2024		* Model regresji liniowej: mechanizm działania modelu, diagnostyka i ewaluacja jakości model
12	20.09.2024		* Modele regresji liniowej: transformacje zmiennych objaśniających i zmiennej objaśnianej, regresja wielomianowa, regularyzacja modelu
13	23.09.2024	Uczenie maszynowe	* Problem klasyfikacji * Model regresji logistycznej * Drzewa decyzyjne * Automatyzacja optymalizacji modeli
14	24.09.2024		* Ensembling, bagging, lasy losowe
15	26.09.2024		* Przygotowywanie danych do modelowania: feature engineering
16	27.09.2024		* Typowe problemy z rzeczywistymi danymi: obsługa braków danych, problem niezbalansowanych
17	30.09.2024	Deep learning	* Metryki jakości klasyfikatorów
18	01.10.2024		* Metodologia rozwiązywania problemów biznesowych z wykorzystaniem uczenia maszynowego * Case study
19	03.10.2024		* Implementacja złożonych procesów predykcyjnych (machine learning pipelines) cz.1
20	04.10.2024		* Implementacja złożonych procesów predykcyjnych (machine learning pipelines) cz.2
			* Wprowadzenie do sieci neuronowych * Perceptron wielowarstwowy * Proces uczenia sieci neuronowych * Problem przeuczenia sieci i techniki regularyzacji * Biblioteka Tensorflow.Keras
			* Computer vision - przetwarzania obrazów * Konwolucyjne sieci neuronowe * Augmentacja danych - powiększanie zbior danych uczących * Transfer learning - wykorzystywania gotowych sieci w praktyce
			* NLP - przetwarzanie języka naturalnego * Rekurencyjne sieci neuronowe * Sieci LSTM * Word embeddings, model word2vec
			* Różne algorytmy uczenia, strojenie procesu uczenia sieci neuronowych * Praktyczne poszukiwanie najlepszych sieci - optymalizacja Bayesowska * Obsługa danych wielomodalnych i wielozadaniowość - sieci o nieliniarnej strukturze