LAB 1 – Bartłomiej Kózka

1. Zainstaluj Anaconde

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Oprogramowanie multimedialne

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

1. Zainstaluj PyCharm Pro

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Oprogramowanie multimedialne

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

1. GitHub repo link

<https://github.com/bartlomiejkozka/ML>

1. Google Colab

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Strona internetowa

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

1. Conda envs list

Obraz zawierający tekst, zrzut ekranu, Czcionka

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

Brak stworzenia środowiska DataScience\_DL wynika z braku posiadania gpu NVIDIA.

1. Conda envs packages

Obraz zawierający tekst, zrzut ekranu, menu

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

Obraz zawierający tekst, zrzut ekranu, menu

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

Obraz zawierający tekst, zrzut ekranu, menu

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

Obraz zawierający tekst, menu, zrzut ekranu

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

Obraz zawierający tekst, menu, zrzut ekranu

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

* 1. Krótki opis zainstalowanych bibliotek
* **category-encoders** – zestaw transformatorów w stylu scikit-learn służących do kodowania zmiennych kategorycznych w zmienne numeryczne np. zamiana wartości typu string w DF na odpowiednie im liczby.
* **feature-engine** – bardziej rozbudowana biblioteka niż category-encoders, zapewniająca dodatkowo takie narzędzia jak imputacja brakujących danych, wyrywanie i usuwanie outlierów, selekcję cech, wszystkie transformaty można wpiąć do pipeline.
* **Featuretools** – automatyczne tworzenie nowych cech z danych tabelarycznych, szczególnie jeśli są to relacyjne bazy danych.
* **Seaborn** – biblioteka to wizualizacji danych zbudowana na bazie matplotlib. Ułatwia tworzenie bardziej estetycznych wykresów przy małej ilości kodu.
* **Numpy** – podstawowa biblioteka do obliczeń numerycznych, zapewnia szybkie operacje na dużych tablicach (macierzach).
* **Pandas** – biblioteka do analizy i przetwarzania danych. Ułatwia szybkie i łatwe manipulowanie danymi w postaci tabelarycznej.
* **Scikit-learn**  - umożliwia łatwe budowanie, trenowanie i ocenianie modeli ML.
* **Scipy** – rozszerzenie możliwości bib. **Numpy**  w zakresie obliczeń inżynierskich.
* **Statsmodels** – biblioteka służąca do statystycznego modelowania, testowania hipotez i analizy danych, wgląd na bardziej szczegółowe statystyki.
* **pyTorch** – umożliwia łatwe tworzenie i trenowanie modeli sieci neuronowych, przeznaczona do deep learning.
* **GeoPandas** – rozszerzenie funkcjonalności biblioteki **pandas** o obsługę danych geograficznych, czyli pracę z danymi przestrzennymi.

1. Różnica pip a conda

Są to dwa różne narzędzia do zarządzania pakietami w Pythonie, ale różnią się w zakresie funkcji jakie mogą wykonywać. Pip jest menadżerem pakietów który zarządza instalacją pakietów z PyPi tylko dla pythona oraz nie ma możliwości zarządzania środowiskami. Natoamiast conda jest menadżerem pakietów który zarządza instalacją pakietów do (pythona, R, Scala) jak i środowiskami wirtualnymi.

Obraz zawierający tekst, zrzut ekranu, menu, Czcionka

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

1. source activate myenv

2. python -m ipykernel install --user --name myenv --display-name "Python py36"

Procedura ta służy do aktywacji środowiska wirtualnego stworzonego za pomocą conda a następnie rejestracji jego jako dostępnego kernela w Jupyter Notebook.

Obraz zawierający tekst, oprogramowanie, zrzut ekranu

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.Obraz zawierający tekst, oprogramowanie, Oprogramowanie multimedialne, zrzut ekranu

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.