# Programowanie AI i ML w Pythonie (kod: K-PYTHON-2-ANALIZA-AI)

## Opis i cel kursu

Kurs prezentuje techniki pracy i biblioteki programistyczne języka Python, które służą szeroko pojętemu *data science*. Obejmuje to samodzielne analizowanie i przetwarzanie danych, wykonywanie obliczeń, stosowanie technik uczenia maszynowego (*machine learning*) oraz głębokiego (*deep learning*).

W czasie kursu zostaną omówione i przećwiczone takie elementy, jak eksploracja i przygotowanie danych, regresja, klasyfikacja, segmentacja, predykcje wartości i zdarzeń, uczenie nadzorowane i nienadzorowane, sieci neuronowe, modele językowe (LLM/GPT). Szkolenie łączy w sobie techniczne spojrzenie właściwe dla programistów ze spojrzeniem analitycznym, charakterystycznym dla statystyków czy analityków danych. Łącząc obie te kompetencje, będziecie Państwo w pełni przygotowani do profesjonalnego wykorzystania Pythona w świecie danych i sztucznej inteligencji.

Zajęcia stanowią naturalną kontynuację dla słuchaczy, którzy niedawno zakończyli kurs "Python Developer I" i wybierają ścieżkę "analityczną" (możliwy jest również wybór "specjalizacji" w kierunku aplikacji webowych/backendu). Przy czym, ukończenie kursu "Nauka programowania w Pythonie" nie jest wymagane formalnie, aby przystąpić do niniejszych zajęć.

## Program

1. **Środowisko pracy z Pythonem dla analityka**
   - Możliwości instalacji lokalnej
     - Czysty Python i wirtualne środowiska (venv i pip)
     - Anaconda
     - Edytory i środowiska deweloperskie
   - Możliwości pracy zdalnej – Google Colab i podobne usługi
     - Asystent AI Gemini
   - Środowiska typu Jupyter – jak się w nich pracuje i dlaczego warto?
     - Komórki i specyfika pracy interaktywnej
     - Używanie Markdown i Latex do tworzenia sformatowanych tekstów
     - Skróty klawiszowe i inne opcje ułatwiające pracę
     - Podpowiedzi i dokumentacja
   - Rzut oka na „ekosystem” bibliotek i narzędzi służących do *data science* w Pythonie
2. **Numpy - tablice i obliczenia w Pythonie**
   - Przyczyny używania dodatkowych bibliotek analityczno-obliczeniowych
   - Wektory, macierze, tablice wielowymiarowe
     - Różne sposoby tworzenia
     - Nawigacja i wybieranie fragmentów
     - Zmiana kształtu (`reshape`) i transpozycja
   - Typy liczbowe i konsekwencje wyboru typu
   - Operacje na tablicach Numpy
     - Wektoryzacja operatorów i funkcji
     - Zasada rozgłaszania (*broadcasting*)
     - Mnożenie macierzy i przykładowe zastosowania
     - Funkcje agregujące i osie
     - Rozkłady, histogramy, statystyki opisowe
   - Generator liczb pseudolosowych
3. **Pandas i „biznesowa” analiza danych**

**Zapytaj o szczegóły**

tel. 22 63 64 164
akademia@alx.pl

**Najbliższe terminy**

2026-10-15 (Zdalnie)
2026-10-15 (Warszawa)
2026-12-15 (Zdalnie)
2026-12-15 (Warszawa)
2026-12-19 (Zdalnie)
2026-12-19 (Warszawa)
2027-01-23 (Warszawa)
2027-01-23 (Zdalnie)
2027-03-13 (Zdalnie)
2027-03-13 (Warszawa)
2027-03-25 (Zdalnie)
2027-03-25 (Warszawa)

- Serie (`Series`) i tabele / „ramki danych” (`DataFrame`)
  - Indeksy i nazwy kolumn
  - Nawigacja w strukturach danych - wybieranie komórek i fragmentów, iteracja
- Pozyskiwanie danych z różnych źródeł
  - Pliki CSV i Excel (odczyt i zapis)
  - Zasoby w internecie (JSON, XML, HTML)
  - Bazy danych (SQL)
- Wyszukiwanie i filtrowanie danych
  - Warunki logiczne w Pandas oraz Numpy: maski, spójniki logiczne
- Przygotowywanie i czyszczenie danych
  - Usuwanie kolumn i wierszy
  - Usuwanie duplikatów
  - Zamiana i normalizacja wartości
  - Strategie postępowania z wartościami nieokreślonymi (NaN)
  - * Usuwanie wartości nieokreślonych
  - * Zastępowanie domyślną lub dominującą wartością
  - * Wypełnianie wartościami sąsiednimi
  - * Interpolacja
  - Typy kolumn oraz ich wpływ na wydajność i funkcjonalność
  - * Typy liczbowe; zakres i precyzja wartości
  - * Typ daty i czasu; akcesor `.dt`
  - * Typ kategoryczny - kiedy warto stosować
  - Praca z danymi tekstowymi, w tym wykorzystanie wyrażeń regularnych (*regex*)
- Pionowe oraz poziome łączenie tabel (`concat`, `merge`, `join`)
- Sortowanie serii i tabel
  - Ranking i poszukiwanie najmniejszych/największych wartości
- Opisowa analiza danych
  - Gotowe funkcje do generowania podstawowych statystyk
  - Grupowanie
  - * Agregacja i obliczanie statystyk dla grup
  - * Iteracja po grupach i wykorzystanie grupowania do podziału danych
  - Tabele przestawne (*pivot table*)
  - Funkcje kroczące („okienkowe”) i narastające („skumulowane”)
- Oś czasu i szeregi czasowe
  - Generowanie szeregów czasowych
  - Specyfikowanie okresów czasu
  - Agregacja okresów czasu – operacja `resample`
- Elementy analizy statystycznej
  - Seria danych jako zmienna statystyczna; normalizacja
  - Korelacja i istotność statystyczna; mapy ciepła
  - Wstęp do regresji liniowej
  - Wnioskowanie statystyczne

4. **Wykresy i wizualizacja danych**
   - Matplotlib i wykresy na podstawie danych z Pandas i Numpy
     - Różne typy wykresów
     - Opcje i ustawienia
   - Generowanie wykresów bezpośrednio z Pandas
   - Przegląd dodatkowych bibliotek wizualizacji danych, m.in. Seaborn, Bokeh, Plotly
5. **Wprowadzenie do uczenia maszynowego**
   - Czym ogólnie jest uczenie maszynowe (*machine learning*, ML)?
   - Podział metod uczenia maszynowego

**Zapytaj o szczegóły**

tel. 22 63 64 164
akademia@alx.pl

**Najbliższe terminy**

2026-10-15 (Zdalnie)
2026-10-15 (Warszawa)
2026-12-15 (Zdalnie)
2026-12-15 (Warszawa)
2026-12-19 (Zdalnie)
2026-12-19 (Warszawa)
2027-01-23 (Warszawa)
2027-01-23 (Zdalnie)
2027-03-13 (Zdalnie)
2027-03-13 (Warszawa)
2027-03-25 (Zdalnie)
2027-03-25 (Warszawa)

- Uczenie nadzorowane (*supervised*), nienadzorowane (*unsupervised*) i ze wzmocnieniem (*reinforcement*)
- Czym jest zagadnienie regresji, a czym problem klasyfikacji?

— *Exploratory Data Analysis* (EDA), czyli eksploracja i przygotowanie danych do uczenia maszynowego
- Statystyki opisowe (za pomocą wcześniej poznanych narzędzi)
- * Rozkłady, histogramy, identyfikacja wartości odstających (*outliers*) i błędnych
- Oczyszczanie zbioru danych (wartości błędne, nieokreślone, duplikaty)
- Identyfikacja korelacji i zależności logicznych, wstępne hipotezy
- Standaryzacja i normalizacja danych
- Czym jest inżynieria cech (*feature engineering*)?

— Proces uczenia maszynowego – metodologia CRISP-ML(Q)
- Zrozumienie dziedziny i określenie wymagań wobec projektu ML
- Przygotowanie zbioru uczącego i zbioru testowego z zastosowaniem EDA
- Zbudowanie i szkolenie modelu
- Walidacja i tuning modelu
- Wdrożenie i utrzymywanie rozwiązań

— Na czym polega „wyjaśnialność" (*explainability*) modelu?
— Wybór najwłaściwszej techniki ML w zależności od zadania

6. **Praktyka uczenia maszynowego w Pythonie**

— Biblioteka Scikit-learn
- Transformatory, estymatory, pipeline'y
- Przegląd dostępnych metod i operacji (w miarę poznawania ☐ kolejnych technik ML)
- Serializacja i ponowne wykorzystywanie modeli

— Rola bibliotek pomocniczych – Numpy, Pandas, Seaborn
— Praca interaktywna w Jupyter/Colab vs wykorzystanie ML we własnych aplikacjach

7. **Techniki uczenia nadzorowanego**

— Metody regresji
- Regresja liniowa
- Regresja wielomianowa
- Obsługa danych o skali nieliniowej
- Reprezentacja danych nieliczbowych

— Ewaluacja modeli regresji
- Współczynnik determinacji $R^2$ oraz błędy MAE/RMSE
- Przeuczenie modelu i metody regularyzacji

— Zastosowania regresji
- Predykcja wartości liczbowych / ciągłych
- Wizualizacja trendów i zależności

— Klasyfikacja
- Decyzje binarne (prawda/fałsz) oraz klasyfikacja wieloklasowa
- Regresja logistyczna
- Drzewa decyzyjne – struktura, sposób tworzenia i wykorzystania
- Metryki klasyfikacji
- Łączenie klasyfikatorów

— Zastosowania klasyfikacji
- Predykcja zdarzeń (wydarzy się / nie wydarzy się) na podstawie uwarunkowań
- Przyporządkowanie rekordów do grup ze względu na ich cechy i wartości pól

8. **Techniki uczenia nienadzorowanego**

— Grupowanie danych (*clustering*) i algorytm K-Means

**Zapytaj o szczegóły**
tel. 22 63 64 164
akademia@alx.pl

**Najbliższe terminy**
2026-10-15 (Zdalnie)
2026-10-15 (Warszawa)
2026-12-15 (Zdalnie)
2026-12-15 (Warszawa)
2026-12-19 (Zdalnie)
2026-12-19 (Warszawa)
2027-01-23 (Warszawa)
2027-01-23 (Zdalnie)
2027-03-13 (Zdalnie)
2027-03-13 (Warszawa)
2027-03-25 (Zdalnie)
2027-03-25 (Warszawa)

- Redukcja wymiarowości i Analiza Głównych Składowych (PCA)
- Zastosowania technik nienadzorowanych (w czasie zajęć wybieramy po jednym małym przykładzie, a dodatkowe projekty są propozycjami prac domowych)
  - Ustalanie istotnych zmiennych, w tym na potrzeby uczenia nadzorowanego
  - Automatyczna detekcja anomalii, np. fraudów, awarii, ataków cybernetycznych
  - Automatyczne grupowanie, np. segmentacja klientów na podstawie historii zachowań czy pacjentów na podstawie objawów
- Czym jest uczenie ze wzmocnieniem (*reinforcement learning*, RL)?

9. **Wstęp do uczenia głębokiego i sieci neuronowych**
   - Czym uczenie głębokie (*deep learning*) różni się od „zwykłego" uczenia maszynowego?
   - Komponenty sieci neuronowych
     - Budowa i mechanizm działania sztucznego neuronu
     - Perceptory (*perceptron*)
     - „Warstwa ukryta" i perceptory wielowarstwowe (MLP)
   - Mechanizm działania uczenia głębokiego
     - Przetwarzanie informacji
     - Algorytmy uczenia z nadzorem
     - Algorytmy uczenia bez nadzoru
     - Funkcje aktywacyjne
     - Funkcje błędu
     - Tensory i transformacje
10. **Biblioteki uczenia głębokiego w Pythonie**
    - Fundamentalne biblioteki napędzające *deep learning*
      - TensorFlow – *open source* od Google
      - PyTorch – *open source* od Meta
      - JAX – zoptymalizowany silnik obliczeniowy
      - Znaczenie sprzętu w pracy z sieciami neuronowymi; procesory typu CPU, GPU, TPU
    - Keras – wysokopoziomowa praca z modelami
      - Wybór biblioteki podstawowej (*backend*)
      - Budowa modeli w Keras
      - Kompilacja modelu
      - Wybór optymalizatora i metryk
      - Trenowanie modeli
      - Sequential API vs Functional API
      - Wykorzystanie wytrenowanych modeli do predykcji
      - Klasyfikacja i podejmowanie decyzji na podstawie nowych danych
    - Hugging Face – platforma *open source* dla modeli uczenia maszynowego
      - Strojenie modeli dopasowanych do potrzeb użytkownika
      - Zastosowanie do generowania obrazów
11. **Sieci neuronowe – rodzaje, techniki i przykłady zastosowań**
    - Sztuczne sieci neuronowe (ANN)
      - Zastosowania w zadaniach klasyfikacji i regresji
    - Konwolucyjne sieci neuronowe (CNN)
      - Rozpoznawanie obrazów
      - Przetwarzanie i analiza obrazów
    - Rekurencyjne sieci neuronowe (RNN, LSTM)
      - Predykcja szeregów czasowych
      - Inne zagadnienia analityczne
12. **Duże modele językowe (LLM)**

**Zapytaj o szczegóły**

tel. 22 63 64 164
akademia@alx.pl

**Najbliższe terminy**

2026-10-15 (Zdalnie)
2026-10-15 (Warszawa)
2026-12-15 (Zdalnie)
2026-12-15 (Warszawa)
2026-12-19 (Zdalnie)
2026-12-19 (Warszawa)
2027-01-23 (Warszawa)
2027-01-23 (Zdalnie)
2027-03-13 (Zdalnie)
2027-03-13 (Warszawa)
2027-03-25 (Zdalnie)
2027-03-25 (Warszawa)

- Czym jest model językowy i jak działają narzędzia typu GPT?
  - Ewolucja AI w przetwarzaniu języka naturalnego (NLP)
  - Od sieci neuronowych RNN do architektury Transformer
  - Architektura popularnych modeli
  - Podstawowe pojęcia: etapy trenowania, model wytrenowany i przetrenowany, token, prompt, konwersacja, pamięć, ...
- Zagadnienia bezpieczeństwa i prywatności w kontekście LLM i AI
  - Etyczne wykorzystanie AI: prawa autorskie, uczciwość i rzetelność narzędzi AI
  - Aspekty prawne i istniejące regulacje (pobieżnie)
  - Zagrożenia ewentualnego wycieku danych i zarządzanie prywatnością
  - Korzystanie ze zdalnych usług AI/LLM czy lokalna instalacja modelu? – dyskusja
- Dostęp programistyczny (API) do zdalnych modeli LLM na przykładzie OpenAI API (ChatGPT) lub analogicznej usługi
  - Generyczne API sieciowe (REST) i dedykowana biblioteka w Pythonie
  - Zarządzanie kluczem dostępowym (API_KEY)
  - Parametry i opcje zapytań
  - * Role promptów: user, developer, ...
  - * Dostrajanie procesu myślenia: zaangażowane zasoby, temperatura, filtrowanie treści
  - Polecenia jednorazowe i konwersacje
- Praca z lokalnym modelem LLM
  - Wady i zalety własnej instalacji względem zewnętrznej usługi
  - Konfiguracja wybranego open-source'owego narzędzia LLM
  - Utworzenie prostego czat-bota jako aplikacji Python na bazie lokalnej instalacji

13. **Analiza możliwości wdrażania projektów AI opartych o Pythona**
- Wdrożenia w oparciu o własną infrastrukturę
- Wdrożenia „w chmurze"
- Wyzwania prawdziwego świata: skalowanie, bezpieczeństwo, monitoring
- Wirtualizacja, konteneryzacja i orkiestracja
  - Rola narzędzi Docker i Kubernetes
  - Wyspecjalizowane narzędzia w kontekście machine learning i deep learning
- Specyfika pracy MLOps

## Przeznaczenie i wymagania

Kurs dla osób, które chcą poszerzyć swoją znajomość Pythona o zagadnienia związane z analizowaniem danych i uczeniem maszynowym.

Oczekiwana jest znajomość Pythona na poziomie porównywalnym z absolwentami kursu Python Developer (wariant standardowy).

Przy czym, uprzednie ukończenie tego kursu u nas nie jest formalnie wymagane - najważniejsze jest znanie podstaw programowania; co to jest zmienna, jak się ją definiuje, jak działa pętla if (i tym podobne).

## Certyfikaty

Uczestnicy szkolenia otrzymują imienne certyfikaty sygnowane przez ALX.

**Zapytaj o szczegóły**
tel. 22 63 64 164
akademia@alx.pl

**Najbliższe terminy**
2026-10-15 (Zdalnie)
2026-10-15 (Warszawa)
2026-12-15 (Zdalnie)
2026-12-15 (Warszawa)
2026-12-19 (Zdalnie)
2026-12-19 (Warszawa)
2027-01-23 (Warszawa)
2027-01-23 (Zdalnie)
2027-03-13 (Zdalnie)
2027-03-13 (Warszawa)
2027-03-25 (Zdalnie)
2027-03-25 (Warszawa)