

# AI & ML development with Databricks

Szkolenie obejmuje kluczowe zagadnienia związane z budową i wdrażaniem rozwiązań AI oraz uczenia maszynowego w środowisku Databricks. Uczestnicy poznają praktyczne aspekty przygotowania danych, tworzenia pipeline'ów ML, trenowania i wdrażania modeli oraz monitorowania ich działania. Program oparty jest na najnowszych trendach i narzędziach wykorzystywanych w branży data science i MLOps.



## Odbiorcy szkolenia

Szkolenie przeznaczone jest dla:

- Data scientistów i inżynierów ML, którzy chcą nauczyć się efektywnego przygotowania danych i budowy modeli w Databricks.
- Inżynierów danych i DataOps odpowiedzialnych za przygotowanie i zarządzanie zestawami danych do uczenia maszynowego.
- Specjalistów MLOps wdrażających i monitorujących modele ML w środowisku chmurowym.
- Zespołów projektujących pipeline'y ML oraz osób przygotowujących się do certyfikacji Databricks Certified Machine Learning Professional.



## Korzyści

- Kompleksowa wiedza o pipeline'ach ML – Nauczysz się budować modułarne pipeline'y do przygotowania danych i trenowania modeli.
- Praktyczne umiejętności pracy z Databricks – Poznasz narzędzia takie jak PySpark, MLflow, Feature Store i Unity Catalog.
- Integracja i automatyzacja procesów ML – Dowiesz się, jak logować metryki, wersjonować modele i zarządzać cechami w Feature Store.

- Zaawansowane przygotowanie danych – Opanujesz techniki EDA, imputacji, kodowania, standaryzacji i inżynierii cech.
- Wdrażanie i monitorowanie modeli – Nauczysz się wdrażać modele do środowiska produkcyjnego i monitorować ich wydajność.



## Program szkolenia

### 1. Tworzenie modeli ML

- Wprowadzenie do procesu budowy modeli uczenia maszynowego w środowisku Databricks.

### 2. Zasady uczenia maszynowego

- Omówienie kluczowych pojęć i zasad uczenia maszynowego, typy algorytmów, dobór modeli.

### 3. Trenowanie modeli uczenia maszynowego w Databricks

- Praktyczne ćwiczenia z trenowania modeli ML z wykorzystaniem Databricks i PySpark.

### 4. Cykl życia modeli ML

- Etapy pracy nad modelem: przygotowanie danych, inżynieria cech, trenowanie, walidacja, wdrożenie.

### 5. Monitorowanie i logowanie eksperymentów i modeli przy użyciu MLflow

- Zarządzanie cyklem życia modeli ML z użyciem MLflow: rejestrowanie eksperymentów, wersjonowanie modeli, śledzenie metryk.

### 6. Ewaluacja wydajności modeli

- Metody oceny jakości modeli: metryki, walidacja krzyżowa, interpretacja wyników.

### 7. Strojenie hiperparametrów

- Optymalizacja hiperparametrów modeli z wykorzystaniem narzędzi Databricks.

### 8. Wykorzystanie AutoML w Databricks

- Automatyzacja procesu budowy modeli ML z użyciem AutoML dostępnego w Databricks.

### 9. Trenowanie modeli deep learning w Databricks

- Wprowadzenie do trenowania modeli głębokiego uczenia (deep learning) w środowisku Databricks.



## Oczekiwane przygotowanie uczestnika

- Znajomość podstawowych pojęć i usług związanych z AI oraz ML.
- Znajomość SQL i/lub Pythona.
- Doświadczenie w pracy z danymi (np. ukończenie modułu data engineering lub równoważna wiedza).
- Zalecane ukończenie szkolenia „Fundamentals of Databricks”.



## Szkolenie obejmuje

\* dostęp do portalu słuchacza Altkom Akademii

Metoda szkolenia:

- Wykład (60%)
- Warsztaty (30%)
- Ćwiczenia (30%)

Główne narzędzia dydaktyczne: Databricks, PySpark, MLflow, Feature Store, Unity Catalog, prezentacje PowerPoint, środowiska laboratoryjne, zasoby Microsoft Learn.

Uwaga: W trakcie wybranych szkoleń Akademia zapewnia Uczestnikom dostęp do płatnych narzędzi i usług niezbędnych do realizacji programu szkolenia (w szczególności usług Azure, Fabric oraz innych zasobów chmurowych). Uczestnik zobowiązany jest do korzystania z udostępnionych zasobów wyłącznie w zakresie wynikającym z agendy szkolenia oraz zgodnie z poleceniami trenera. Wykorzystywanie tych zasobów w sposób wykraczający poza zakres szkoleniowy, w szczególności do realizacji projektów prywatnych, testowania dodatkowych usług, modyfikowania konfiguracji środowiska lub podejmowania działań generujących koszty niezwiązane z realizacją ćwiczeń – skutkować będzie obciążeniem Uczestnika kosztami przekraczającymi przewidziany limit, na podstawie wydatków poniesionych przez Akademię (<https://www.altkomakademia.pl/ogolne-warunki-uczestnictwa-w-szkoleniach/>)



## Język

Wykład: polski

Materiały: angielski

## Metoda egzaminacyjna

Szkolenie nie kończy się egzaminem. Przygotowuje jednak do egzaminu Databricks Certified Machine Learning Professional.

## Czas trwania

1 dni / 7 godzin

## Opis egzaminu

Szkolenie nie kończy się egzaminem .Przygotowuje jednak do egzaminu Databricks Certified Machine Learning Professional.

Opis egzaminu: <https://www.databricks.com/learn/certification/machine-learning-professional>