BOTTEGA IT minds

Program szkolenia:

Machine Learning and Local GenAI: Od Scikit-Learn i TensorFlow do LM Studio

Informacje:

Nazwa:	**Machine Learning and Local GenAI: Od Scikit-Learn i TensorFlow do LM Studio**
Kod:	**ml-gen**
Kategoria:	ML
Czas trwania:	1 dzień
Forma:	50% wykłady / 50% warsztaty

W świecie AI narzędzia zmieniają się błyskawicznie. To szkolenie porządkuje wiedzę, łącząc sprawdzone metodyki Data Science z najnowszymi trendami Local GenAI.

Uczymy budować kompletne pipeline'y przetwarzania danych, świadomie dobierać narzędzia (kiedy klasyczny ML, a kiedy Deep Learning?) oraz uruchamiać modele LLM lokalnie na własnym sprzęcie. Program inspirowany projektami badawczymi, gdzie kluczowa była prywatność danych (Off-line AI) oraz precyzyjna walidacja wyników.

Dla kogo jest to szkolenie?

- Szkolenie dla zespołów technicznych, które chcą wyjść poza proste skrypty i zacząć budować skalowalne, bezpieczne rozwiązania ML, rozumiejąc różnice między dostępnymi technologiami.

Zalety szkolenia:

- Architektury rozwiązań: Kiedy stosować Scikit-learn (klasyczne algorytmy), kiedy TensorFlow (sieci neuronowe), a kiedy gotowe modele LLM.
- TensorFlow od podstaw: Zrozumienie tensorów, warstw i budowa prostego modelu, z naciskiem na różnice względem podejścia klasycznego.
- LM Studio and Local LLMs: Jak uruchamiać potężne modele językowe (Llama 3, Mistral) lokalnie, bez wysyłania danych do chmury – idealne dla danych wrażliwych.
- Unikania pułapek: Data Leakage, Overfitting i Bias w modelach.

BOTTEGA IT minds

Szczegółowy program:

1. Fundamenty Pipeline'u i Feature Engineering

1.1. Data Science Methodology: Dlaczego model to tylko 20% sukcesu? Cykl życia projektu.

1.2. Przygotowanie danych w Pandas: Czyszczenie, normalizacja i walidacja jakości.

1.3. Feature Engineering: Jak zamienić surowe dane w wartościowe cechy? (Encoding, Scaling).

1.4. Case study: Wykrywanie Data Leakage – jak nieświadomie "oszukujemy" sami siebie podczas treningu.

2. Klasyczne ML vs. Deep Learning (Scikit-learn vs. TensorFlow)

2.1. Scikit-learn: Szybkie prototypowanie modeli klasyfikacyjnych i regresyjnych (Decision Trees, Random Forest).

2.2. Wprowadzenie do TensorFlow:

- 2.2.1. Czym różni się tensor od zwykłej macierzy?
- 2.2.2. Kiedy klasyczne ML przestaje wystarczać? Granice Scikit-learn.
- 2.2.3. Budowa prostej sieci neuronowej – warstwy, aktywacje, optymalizatory.

2.3. Warsztat porównawczy: Rozwiązanie tego samego problemu (np. klasyfikacja ryzyka /anomalii) za pomocą Scikit-learn oraz TensorFlow. Analiza wyników, czasu treningu i złożoności.

3. Walidacja i Interpretowalność

3.1. Jak ocenić, czy model działa? Metryki (Accuracy vs. Recall/Precision) w zależności od problemu biznesowego.

3.2. Anomaly Detection: Wykorzystanie prostych metod statystycznych vs. Autoenkodery w TensorFlow.

3.3. Explainable AI: Interpretacja „czarnej skrzynki". Użycie bibliotek (np. SHAP) do wyjaśnienia, dlaczego model podjął daną decyzję (kluczowe w medycynie/finansach).

4. Local GenAI z LM Studio

4.1. Dlaczego warto uruchamiać LLM lokalnie? Prywatność, koszty, brak zależności od API.

4.2. LM Studio w praktyce:

- 4.2.1. Instalacja i konfiguracja środowiska.

4.2.2. Przegląd otwartych modeli (Hugging Face) – jak dobrać model do sprzętu (kwantyzacja).

4.3. Use-case: Automatyczna analiza/streszczanie dokumentów z zachowaniem pełnej poufności (Offline).

5. Podsumowanie i Best Practices

5.1. MLOps w pigułce: Jak zarządzać wersjami modeli i danych.

5.2. Etyka i Bias: Jak zapobiegać stronniczości modeli w zastosowaniach badawczych i komercyjnych.

5.3. Q and A i roadmapa dalszego rozwoju dla uczestników.