



CODEBRAINERS

SPÓŁKA Z

OGRANICZONĄ

ODPOWIEDZIALNOŚĆ

CIĄ



Data Science i Deep Learning z Apache Spark i CNN (kurs praktyczny)

Numer usługi 2025/05/12/118259/2739372

zdalna w czasie rzeczywistym

Usługa szkoleniowa

52 h

20.05.2025 do 12.06.2025

4 680,00 PLN brutto

4 680,00 PLN netto

90,00 PLN brutto/h

90,00 PLN netto/h

Informacje podstawowe

Kategoria	Informatyka i telekomunikacja / Programowanie
Identyfikator projektu	Małopolski Pociąg do kariery
Sposób dofinansowania	wsparcie dla osób indywidualnych wsparcie dla pracodawców i ich pracowników
Grupa docelowa usługi	Kurs skierowany jest do osób interesujących się praktycznym zastosowaniem sztucznej inteligencji oraz modeli uczenia maszynowego w analizie danych - wiedzy, która daje obecnie olbrzymie możliwości rozwoju w praktycznie każdej branży. Zaleca się, aby Uczestnicy posiadali wiedzę z zakresu programowania w języku Python. Usługa adresowana również również dla Uczestników Projektu Kierunek – Rozwój, Małopolski Pociąg do Kariery, mBony+, Graj po Zielone (oraz innych programów w ramach 6.6 i 10.17 z terenu woj. Śląskiego), a także dla uczestników innych programów dofinansowań.
Minimalna liczba uczestników	8
Maksymalna liczba uczestników	14
Data zakończenia rekrutacji	19-05-2025
Forma prowadzenia usługi	zdalna w czasie rzeczywistym
Liczba godzin usługi	52

Cel

Cel edukacyjny

Kurs przygotowuje do samodzielnego projektowania, trenowania i wdrażania modeli uczenia maszynowego z wykorzystaniem narzędzi takich jak Apache Spark, Databricks i MLlib oraz budowy i zastosowania sieci neuronowych (CNN) w praktycznych projektach analizy danych. Uczestnicy nabędą umiejętność pracy w środowiskach przetwarzania danych na dużą skalę oraz zdobędą doświadczenie w realizacji pełnych projektów AI.

Efekty uczenia się oraz kryteria weryfikacji ich osiągnięcia i Metody walidacji

Efekty uczenia się	Kryteria weryfikacji	Metoda walidacji
Posługuje się wiedzą z zakresu programowania w języku Python oraz uczenia maszynowego	charakteryzuje składnię Pythona (typy danych oraz podstawowe struktury danych, takie jak listy, słowniki, krotki)	Test teoretyczny z wynikiem generowanym automatycznie
	charakteryzuje kluczowe zagadnienia w uczeniu maszynowym, takie jak podział danych, walidacja krzyżowa, regularyzacja oraz metryki oceny modeli	Test teoretyczny z wynikiem generowanym automatycznie
	charakteryzuje działanie algorytmów uczenia maszynowego w Pythonie	Test teoretyczny z wynikiem generowanym automatycznie
Tworzy modele uczenia maszynowego	dostosowuje wybór dostępnych bibliotek i narzędzi (Scikit-learn, TensorFlow, XGBoost) do specyfiki problemu - pod kątem ich implementacji i trenowania modeli uczenia maszynowego	Test teoretyczny z wynikiem generowanym automatycznie
	korzysta z algorytmów k-NN, regresji logistycznej, k-means dla uzyskania optymalnych wyników	Analiza dowodów i deklaracji
	wykorzystuje modele uczenia maszynowego do klasyfikacji danych, w tym danych środowiskowych	Analiza dowodów i deklaracji
Współpracuje i komunikuje się z innymi członkami zespołu	wskazuje prawidłowe sposoby komunikacji za pośrednictwem narzędzi kontroli wersji (git)	Test teoretyczny z wynikiem generowanym automatycznie

Kwalifikacje

Kompetencje

Usługa prowadzi do nabycia kompetencji.

Warunki uznania kompetencji

Pytanie 1. Czy dokument potwierdzający uzyskanie kompetencji zawiera opis efektów uczenia się?

Tak, dokument potwierdzający uzyskanie kompetencji zawiera szczegółowy opis efektów uczenia się, co pozwala jasno zrozumieć osiągnięte kompetencje oraz postępy w nauce.

Pytanie 2. Czy dokument potwierdza, że walidacja została przeprowadzona w oparciu o zdefiniowane w efektach uczenia się kryteria ich weryfikacji?

Tak, dokument potwierdza, że walidacja została przeprowadzona zgodnie z zdefiniowanymi w efektach uczenia się kryteriami weryfikacji, co gwarantuje rzetelność i wiarygodność procesu oceny osiągniętych kompetencji.

Pytanie 3. Czy dokument potwierdza zastosowanie rozwiązań zapewniających rozdzielenie procesów kształcenia i szkolenia od walidacji?

Tak, dokument potwierdza zastosowanie rozwiązań zapewniających rozdzielenie procesów kształcenia i szkolenia od walidacji. Dzięki temu procesowi zapewniona jest obiektywność i uczciwość oceny osiągniętych kompetencji

Program

Kurs to intensywny, praktyczny program online, który przygotowuje uczestników do pracy z dużymi zbiorami danych oraz budowy modeli sztucznej inteligencji. Uczestnicy uczą się tworzyć modele ML w środowisku Apache Spark i Databricks, a także trenować konwolucyjne sieci neuronowe (CNN) do analizy obrazów. Kurs opiera się na pracy projektowej, łączy teorię z praktyką i kończy się realizacją indywidualnego projektu oraz walidacją zdobytych kompetencji..

W szkoleniu mogą wziąć udział zarówno osoby, które myślą o przyszłej pracy na stanowiskach Data Scientist, AI Engineer, Machine Learning Developer, jak również osoby chcące zdobyć nowe umiejętności w zajmowanych już stanowiskach pracy, kadra kierownicza.

Zaleca się, aby Uczestnicy posiadali wiedzę z zakresu programowania w języku Python.

Z racji dynamicznego rozwoju branż zielonych technologii, kurs uczy tworzenia modeli uczenia maszynowego w sposób pozytywnie wpływający na środowisko i wspierający zrównoważony rozwój oraz pozwala na zrozumienie roli technologii cyfrowych we wspieraniu zielonej gospodarki.

--

STRUKTURA KURSU:

- **kurs obejmuje 52h lekcyjne (45 min) = w przeliczeniu 39h zegarowych (60 min)) prowadzonych na żywo (on-line)**, na platformie webinarowej, w formie wirtualnej klasy, w formule live-coding - przez cały czas z trenerem
- dodatkowo planowana jest samodzielna praca własna kursantów w domu (ćwiczenia, projekty), z możliwością konsultacji na platformie Slack - praca ta pozwala utrwalić zdobyta podczas zajęć wiedzę i nie jest wliczana do czasu trwania usługi - nie jest to obowiązkowe;
- zajęcia odbywają się na żywo (online, w formie wirtualnej klasy) w formule wieczorowo-weekendowej - 2x w tygodniu (wieczorem) oraz w wybraną sobotę
- grupa liczy maksymalnie 14 osób i jest jedną z najmniejszych grup na rynku

--

MODUŁ 1: Uczenie maszynowe z Apache Spark i Databricks

- Wprowadzenie do Big Data, Apache Spark i Databricks
- Praca z danymi w Spark (DataFrames, transformacje)
- Tworzenie modeli ML w Spark MLlib (klasyfikacja, regresja, pipeline)

- Walidacja modeli, tuning hiperparametrów
- **Projekt zespołowy:** przygotowanie i wdrożenie modelu MLlib w Databricks

MODUŁ 2: Deep Learning – sieci neuronowe CNN w praktyce

- Podstawy deep learningu i architektura CNN
- Przetwarzanie i klasyfikacja obrazów (wiek, płeć)
- Praca w TensorFlow lub PyTorch
- **Projekt praktyczny:** budowa modelu CNN do klasyfikacji zdjęć

MODUŁ 3: Projekt końcowy i walidacja efektów uczenia się

- Projekt końcowy: samodzielny wybór problemu i narzędzi (Spark lub CNN)
- Konsultacje z mentorem, przygotowanie prezentacji
- Test cyfrowy walidujący efekty uczenia się
- Tydzień kariery: przegląd CV, LinkedIn, portfolio projektów

- **całość zajęć prowadzona jest na żywo online**
- **aby osiągnąć zakładany cel** realizacji usługi, uczestnik powinien być obecny w trakcie zajęć zdalnych w czasie rzeczywistym
- usługa szkoleniowa realizowana jest w godzinach dydaktycznych (1 godzina dydaktyczna = 45 min.) - łącznie 52h dydaktyczne
- w ramach usługi przewidziane są przerwy podczas zajęć 6 godzinnych w soboty, które zostały uwzględnione w harmonogramie usługi, jednak nie wliczają się do ilości godzin samej usługi
- walidacja efektów kształcenia odbywa się w formie testu teoretycznego w formie cyfrowej, z wynikiem generowanym automatycznie. Test prowadzony jest na zewnętrznej platformie, w oparciu o indywidualne kody dostępu przypisane do każdego z uczestników, z zapewnieniem rozdzielnosci pomiędzy szkoleniem, a walidacją

Kurs uczy zaawansowanych technik analizy danych z wykorzystaniem modeli opartych na sztucznej inteligencji, które to modele są z powodzeniem wykorzystywane m.in. w celu predykcji zmian klimatycznych, optymalizacji procesów i redukcji zużycia zasobów. Big data wspiera optymalizację zużycia energii w budynkach i infrastrukturze oraz śledzenie emisji i odpadów, umożliwia automatyzację i personalizację, wspiera rozwój energii odnawialnej i monitorowanie zużycia wody, energii i pozostałych zasobów.

Dzięki przekazywaniu umiejętności ogólnych niezbędnych do pracy w sektorze zielonej gospodarki, szkolenie przyczynia się również do tworzenia tzw. "zielonych miejsc pracy" zarówno w sektorach zielonej gospodarki, jak również w sektorach tradycyjnych.

Wiedza zdobyta podczas szkolenia wykorzystywana może być m.in. w celu realizacji inwestycji opisanych m.in. w Rozp. nr 2021/1056 PEiR(UE) ustanawiającym Fundusz Sprawiedliwej Transformacji, tj. m.in. w przypadku wdrażania technologii oraz systemów i infrastruktury zapewniającej czystą energię, redukcję emisji gazów cieplarnianych, inwestycji w energię odnawialną i w efektywność energetyczną, inteligentną i zrównoważoną mobilność lokalną, poprawę efektywności energetycznej, cyfryzację i łączność cyfrową.

Harmonogram

Liczba przedmiotów/zajęć: 17

Przedmiot / temat zajęć	Prowadzący	Data realizacji zajęć	Godzina rozpoczęcia	Godzina zakończenia	Liczba godzin
1 z 17 Projekt MLlib / Databricks	Bartek Bilski	20-05-2025	18:00	21:00	03:00
2 z 17 Projekt MLlib / Databricks	Bartek Bilski	22-05-2025	18:00	21:00	03:00

Przedmiot / temat zajęć	Prowadzący	Data realizacji zajęć	Godzina rozpoczęcia	Godzina zakończenia	Liczba godzin
3 z 17 Projekt Mllib / Databricks	Mikołaj Kucharski	24-05-2025	09:00	12:00	03:00
4 z 17 Przerwa w usłudze	Mikołaj Kucharski	24-05-2025	12:00	12:30	00:30
5 z 17 Projekt Mllib / Databricks	Mikołaj Kucharski	24-05-2025	12:30	15:30	03:00
6 z 17 Projekt sieci neuronowych CNN-klasyfikacja wieku i płci na podstawie zdjęć (on-line, na żywo, wykład + ćw.)	-	27-05-2025	18:00	21:00	03:00
7 z 17 Projekt sieci neuronowych CNN-klasyfikacja wieku i płci na podstawie zdjęć (on-line, na żywo, wykład + ćw.)	-	29-05-2025	18:00	21:00	03:00
8 z 17 Projekt sieci neuronowych CNN-klasyfikacja wieku i płci na podstawie zdjęć (on-line, na żywo, wykład + ćw.)	Bartek Bilski	31-05-2025	09:00	12:00	03:00
9 z 17 Przerwa w usłudze	Bartek Bilski	31-05-2025	12:00	12:30	00:30
10 z 17 Projekt sieci neuronowych CNN-klasyfikacja wieku i płci na podstawie zdjęć (on-line, na żywo, wykład + ćw.)	Bartek Bilski	31-05-2025	12:30	15:30	03:00

Przedmiot / temat zajęć	Prowadzący	Data realizacji zajęć	Godzina rozpoczęcia	Godzina zakończenia	Liczba godzin
11 z 17 Projekt sieci neuronowych CNN-klasyfikacja wieku i płci na podstawie zdjęć (on-line, na żywo, wykład + ćw.)	Bartek Bilski	05-06-2025	18:00	21:00	03:00
12 z 17 projekt finalny (on-line, na żywo, wykład + ćw.)	Bartek Bilski	07-06-2025	09:00	12:00	03:00
13 z 17 Przerwa w usłudze	Bartek Bilski	07-06-2025	12:00	12:30	00:30
14 z 17 projekt finalny (on-line, na żywo, wykład + ćw.)	Bartek Bilski	07-06-2025	12:30	15:30	03:00
15 z 17 projekt finalny (on-line, na żywo, wykład + ćw.)	Bartek Bilski	10-06-2025	18:00	21:00	03:00
16 z 17 career week	Bartek Bilski	12-06-2025	18:00	20:00	02:00
17 z 17 walidacja efektów uczenia się (test w formie cyfrowej)	Szymon Budziak	12-06-2025	20:00	21:00	01:00

Cennik

Cennik

Rodzaj ceny	Cena
Koszt przypadający na 1 uczestnika brutto	4 680,00 PLN
Koszt przypadający na 1 uczestnika netto	4 680,00 PLN
Koszt osobogodziny brutto	90,00 PLN

Prowadzący

Liczba prowadzących: 3



1 z 3

Szymon Budziak

Data Scientist w Transmission Dynamics Poland. Aktywny uczestnik wielu Hackatonów i wykładowca w Kole Naukowym BIT AI. Programista i inżynier systemów ACK Cyfronet przy komputerach HPC. Pasjonat programowania, sztucznej inteligencji, uczenia maszynowego oraz inwestowania. Prywatnie również trener tenisa ziemnego.

Wykształcenie: Computer Science, Engineer's degree, Università degli Studi di Napoli Federico II (2023), inż., Informatyka, AGH w Krakowie (2024). Dodatkowe szkolenia, m.in.: Python for Data Science and Machine Learning, Machine Learning Specialization by Andrew Ng.

Posiada dośw. w zakresie ziel. komp. W okresie ostatnich 5 lat: m.in.: twórca aplikacji przewidującej efektywność energetyczną w gospodarstwach domowych USA na podstawie danych energetycznych (z wykorzystaniem AI), w oparciu o technologie Python oraz TensorFlow do analizy danych energetycznych, a także API do zbierania informacji o zużyciu energii. Aplikacja miała na celu wspieranie decyzji o optymalizacji zużycia energii w gospodarstwach domowych i promowanie bardziej zrównoważonych praktyk w gospodarce.



2 z 3

Mikołaj Kucharski

Obecnie Big Data Engineer w Allegro, wcześniej Data Scientist w Capgemini oraz Data Engineer w Ernst & Young. Zawodowo zajmuje się zagadnieniami z pogranicza data engineeringu, data science i sztucznej inteligencji w branży konsultingowej. Głównymi narzędziami jego pracy są Python, SQL, pyspark i technologie chmurowe, przede wszystkim Microsoft Azure. Jego obszarem zainteresowań są w głównej mierze konwolucyjne sieci neuronowe i ich zastosowanie w analizie obrazów. Poza rozwojem zawodowym stawia również na rozwój fizyczny trenując brazylijskie jiu jitsu. Wolne chwile lubi spędzać przy dobrym filmie lub książce.

Doświadczenie, m.in.: obecnie, Big Data Engineer, Allegro, 2023 - 2024, Data Science, Capgemini, 2023, Data Engineer, EY, 2021 - 2023, Assistant Geotechnical Designer, Stump Franki, 2020 - 2021, Geotechnical Engineer, Jacobs, 2019 - 2020, Geotechnical Specialist. Wykształcenie: mgr, Informatyka, Polsko-Japońska Akademia Technik Komputerowych w Warszawie (2024), mgr inż., Civil Engineering - Geotechnical Engineering, SGGW w Warszawie (2021), inż., Civil Engineering - Geotechnical Engineering, SGGW w Warszawie (2022).

Posiada dośw. w zakresie ziel. komp. W okresie ostatnich 5 lat: m.in. praca nad infr. do analizy KPI z zakresu ESG, praca nad optymalizacją proc. biznesowych w celu redukcji zuż. zasobów, optymalizacja przetw. dużych zbiorów danych w celu red. kosztów i śladu węglowego.



3 z 3

Bartek Bilski

Programista języka Python oraz Data Scientist. Obecnie Data Scientist w Allegro, wcześniej Data Scientist w Kokoro Global, Data Analyst w cord. Absolwent studiów magisterskich Data Science na Uniwersytecie w Bath, z zapleczem w matematyce i statystyce. Entuzjasta sztucznej inteligencji, nauki i brazylijskiego jiu jitsu. Autor publikacji z zakresu Data Science, jak również kilku książek edukacyjnych.

Doświadczenie, m.in.: obecnie, Data Scientist, Allegro, 2023 - 2024, Data Scientist, Kokoro Global, 2022, Data Analyst, Cord, 2020 - 2021, Mortgage Advisor. Wykształcenie: University of Bath (Data Science)(2022) oraz University of West of England (Matematyka i Statystyka)(2017). Dodatkowe szkolenia, m.in.: Azure Databricks, Databricks Data + AI World Tour London, IRX @ DTX + UCX.

Posiada dośw. w zakresie ziel. komp. W okresie ostatnich 5 lat: m.in.: analizy i projekt segmentacji(GivEnergy Ltd., magazyny energii), skupiający się na odnawialnych źródłach energii i działaniach proekologicznych, z wykorzystaniem ML oraz modeli LLM (GPT).

Informacje dodatkowe

Informacje o materiałach dla uczestników usługi

W ramach szkolenia uczestnik otrzymuje:

- dostęp do materiałów oraz ćwiczeń podsumowujących zdobytą wiedzę (materiały on-line formie pdf,html, jupyter notebook)
- zbiory danych wykorzystywane podczas ćwiczeń
- bezpłatną licencję edukacyjną na wybrane IDE JetBrains
- dostęp do kanałów Slack dedykowanych szkoleniu
- dostęp do nagrań z odbytych zajęć

Warunki uczestnictwa

- **zaleca się, aby Uczestnicy posiadali wiedzę z zakresu programowania w języku Python**
- w szkoleniu mogą wziąć udział zarówno osoby, które myślą o przyszłej pracy w roli data scientist, jak również kadra kierownicza i pracownicy sektora zielonej gospodarki, czy właściciele firm, dla których dostęp do analiz jest kluczem do podejmowania trafnych decyzji w bieżącej działalności
- w przypadku korzystania z dofinansowania, warunkiem uczestnictwa jest zapisanie się przez BUR wraz z podaniem aktualnego ID wsparcia

Informacje dodatkowe

- zakres zg. z RSI Woj. Śl. 2030: Techn. Inf. i kom., (i) techn. szt. int. i uczenia masz., (ii) techn. data mining, (iii) techn. zaaw. baz danych i hurtowni danych oraz z RSI Woj. Mał.. Met. i urz. służące do poz. dan.
- zapisanie się w BUR nie jest jednoznaczne z zarezerwowaniem miejsca. W celu potwierdzenia miejsca prosimy o dodatkowy kontakt telefoniczny, mailowy, lub za pośrednictwem messenger'a albo www
- zawarto umowę z WUP w Krakowie w ramach projektu Małopolski Pociąg do Kariery
- zawarto umowę z WUP w Toruniu w ramach projektu Kierunek Rozwój
- usługi dedykowane również uczestnikom innych programów dofinansowań
- zdobyte kompetencje dotyczą cyfrowej transformacji
- podstawa zwolnienia z VAT: Dz.U.2013.1722, art. 3, ust. 1, pkt. 14 - usł. kszt. zaw. lub przekw. zaw., fin. w co najmniej 70% ze środków publ. - podstawa zwolnienia jest każdorazowo weryfikowana w stosunku do danego Uczestnika

Warunki techniczne

- zajęcia prowadzone są w czasie rzeczywistym na platformie Zoom, wraz z dostępem do kanałów grupowych na platformie Slack
- **Minimalne wymagania sprzętowe:** komputer / laptop / lub inne urządzenie ze stałym dostępem do internetu, wyposażone w kamerę internetową
- **Minimalne wymagania dotyczące parametrów łącza sieciowego:** szybkość pobierania / przesyłania: minimalna 2 Mb/s / 128 kb/s, zalecana: 4 Mb/s / 512 kb/s

- **Niezbędne oprogramowanie umożliwiające dostęp do zajęć oraz materiałów:** przeglądarka internetowa, Zoom w wersji bezpłatnej dla użytkownika
- Uczestnicy otrzymują linki do spotkań przed każdymi zajęciami. Link umożliwiający uczestnictwo w kursie jest aktywny w godzinach wskazanych na karcie usługi

Kontakt



Barbara Zelek

E-mail biuro@codebrainers.pl

Telefon (+48) 795 952 795