Python - BigData a Hadoop

Kód kurzu: PYTHON_BIGDATA

Pojem Big Data se nyní odráží v mnoha oblastech a sektorech národního hospodářství, a proto je nutné nejen získat teoretické znalosti, ale i praktické zkušenosti se zpracováním velkého množství dat. Kurz představuje vývoj Big Data v prostředí programovacího jazyka Pythonu. Python je nyní nejdůležitějším jazykem nejen pro vývojáře aplikací, analytiky dat, ale především pro profesionály v oboru Data Science.

Pro koho je kurz určen:

- Data Scientist, datoví analytici, zejména v Big Data prostředí, jsou primárním auditoriem pro tento intenzivní kurz.
- Software vývojáři, kteří ovládají jazyk Python alespoň na střední az pokročilé úrovni a kteří mají za cíl vytvářet data-intenzivní aplokace pomocí enginu SPARK v prostredi Big Dat (Cloud).
- Datoví architekti

Požadované vstupní znalosti:

- Základní znalost jazyka Python na úrovni kurzu PYTHON_INTRO

Metody výuky:

- Odborný výklad s praktickými ukázkami, cvičení na počítačích.

Studijní materiály:

- Tištěné prezentace probírané látky.

Osnova:

Hadoop Distribuovaný Souborový Systém (HDFS) a Python

- Technologie MapReduce v Python-e
- Jazyk Pig a Python
- Spark / Yarn v Python-e
- Konfigurace virtuálního prostředí pro Spark
- Tvorba Batch a streamování aplikací s Spark
- Žongolování" s daty pomocí Spark
- Základy strojového učení z dat pomocí Spark
- Streamování "živých" dat se Spark
- Vizualizace dat
- Workflow Management v Python (Oozie)

GOPAS Praha
Na Strži 2097/63
140 00 Praha 4 - Krč
Tel.: +420 226 201 390
info@gopas.cz

GOPAS Brno
Nové sady 996/25
602 00 Brno
Tel.: +420 530 513 590
info@gopas.cz

GOPAS Bratislava
Dr. Vladimíra Clementisa 10
Bratislava, 821 02
Tel.: +421 902 903 132
info@gopas.sk

Copyright © 2026 GOPAS, a.s.,
All rights reserved