

Strojové učenie nad veľkými dátami

Kód kurzu: MLC_BDATA

Cieľom tohto kurzu je predstaviť rôzne nástroje a koncepty strojového učenia nad veľkými dátami. Po dokončení tohto kurzu by účastníci mali byť schopní určiť, aký nástroj použiť pre daný problém, zistiť, či neexistuje jednoduchšie riešenie, a poznať časté chyby a vedieť sa im vyhnúť. Špeciálnu pozornosť venujeme Sparku ako univerzálnemu nástroju, ktorý možno použiť na spracovanie veľkých dát, ako aj na strojové učenie nad veľkými dátami.

Pobočka	Dní	Katalógová cena	ITB
Praha	1	4 990 Kč	0

Všetky ceny sú uvedené bez DPH.

Termíny kurzu

Dátum	Dní	Cena kurzu	Typ výučby	Jazyk výučby	Lokalita
-------	-----	------------	------------	--------------	----------

Všetky ceny sú uvedené bez DPH.

Požadované vstupné znalosti

- Základy práce v Pythone a v nástroji Google Colab
- Znalosti strojového učenia na úrovni kurzu Úvod do strojového učenia.

Študijné materiály

Študijný materiál spoločnosti Machine Learning College.

Osnova kurzu

- Prehľad konceptov a nástrojov v spracovaní veľkých dát
- Od malých k veľkým dátam a odhad ich hodnoty
- Riadkové a stĺpcové databázy
- HDFS (Hadoop Distributed File System)
- Formáty dát – Parquet, ORC, Avro
- Komprimácia – gzip, snappy, zstd
- SQL databázy – BigQuery, Redshift, Clickhouse, Snowflake, Vertica
- Praktický príklad na porovnanie malých a veľkých dát
- Úvod do Sparku
- MapReduce
- Spark Computing Engine a RDDs (Resilient Distributed Datasets)
- DataFrames
- Spark ekosystém
- Najčastejšie chyby
- Kde spustiť Spark
- Alternatívy – Apache Beam (Dataflow), Dask, lambdas
- Praktický príklad so Sparkom
- ML stratégie pre veľké dáta
- Inkrementálne učenie
- Dávkové učenie pre neurónové siete
- Distribuované tréningovanie
- Federated learning
- Alternatívne stratégie
- Náhodné vzorkovanie
- Podmodely

GOPAS Praha

Kodaňská 1441/46
101 00 Praha 10
Tel.: +420 234 064 900-3
info@gopas.cz

GOPAS Brno

Nové sady 996/25
602 00 Brno
Tel.: +420 542 422 111
info@gopas.cz

GOPAS Bratislava

Dr. Vladimíra Clementisa 10
Bratislava, 821 02
Tel.: +421 248 282 701-2
info@gopas.sk



Copyright © 2020 GOPAS, a.s.,
All rights reserved

Strojové učenie nad veľkými dátami

- Väčšie výpočtové kapacity
- Frameworky
- Scikit-learn a partial_fit
- MLlib
- Dask-ML
- Praktické príklady s frameworkmi
- Najčastejšie chyby

GOPAS Praha

Kodaňská 1441/46
101 00 Praha 10
Tel.: +420 234 064 900-3
info@gopas.cz

GOPAS Brno

Nové sady 996/25
602 00 Brno
Tel.: +420 542 422 111
info@gopas.cz

GOPAS Bratislava

Dr. Vladimíra Clementisa 10
Bratislava, 821 02
Tel.: +421 248 282 701-2
info@gopas.sk



Copyright © 2020 GOPAS, a.s.,
All rights reserved