

# 18AAD

## Aplikovaná analýza dat

LS 1+1 z volitelný

Zajímá vás oblast **zpracování a analýzy dat**?

**Chtěli byste se dozvědět** více o tom, jak své teoretické základy uplatnit v praxi na reálné problémy?

Jak probíhá **vývoj (datových) aplikací** a co byste měli vědět, než do svého prvního týmu nastoupíte?

Rádi byste si zapsali **prakticky zaměřený předmět** bez zkoušky, který vás provede tématy **Big Data, neuronových sítí, paralelních výpočtů, analýzou grafů, cloudových technologií, nasazováním a vývojem softwaru** či IoT?

## Zapište si 18AAD!

Předmět je vyučován ve spolupráci se společností **BigHub**, která se věnuje vývojem datových aplikací s využitím umělé inteligence v energetice, logistice, financích či medicíně a má v týmu mnoho absolventů **Jaderky**.

1. **Big Data & Data Science:** Využití moderních nástrojů na akvizici, zpracování a vyhodnocení dat. Úložiště dat a jejich využití (SQL, NoSQL, objektové databáze, fulltextové databáze). Formáty pro uložení dat a jejich vhodnost.
2. **Frameworky pro paralelní výpočty** (MapReduce, Spark): Zpracování tokových dat. Komerční i volně dostupné nástroje, příklady použití.
3. **Neuronové sítě:** Využití neuronových sítí v praxi. Rekapitulace vývoje a nezbytný teoretický základ. Přehled dostupných nástrojů pro jejich optimalizaci/inferenci.
4. **Využití GPU pro zrychlení a paralelizaci výpočtů:** Srovnání s nástroji pro klasické strojové učení. Ukázka na zpracování obrazu, NLP či predikce časových řad.
5. **Analýza grafů:** Analytické výpočty nad grafy na reálných příkladech. Rekapitulace teorie grafů, základní algoritmy. Page Rank, hledání influencerů, statistiky nad hlubokými organizačními strukturami.
6. **Využití frameworku Pregel a nástroje GraphX:** Využití grafových databází, grafových jazyků pro efektivní dotazování (GraphQL/Cipher). Vizualizace grafů (např. v PowerBI, Neo4j apod.).
7. **Zpracování dat a AI v Cloudu:** Porovnání Cloud a On-premise. PaaS a IaaS. Infrastructure as a Code. Analytické nástroje v Cloudu (analytické nástroje, databáze, frameworky, rozhraní). Hands-on využití Cloudu pro řešení reálných problémů.
8. **Nasazení datových a AI řešení do firemních procesů:** Nasazení datových řešení do produkce. Projektové metodologie (Scrum/Agile/Waterfall). Způsob vedení projektů/týmů (DevOps/DataOps).
9. **Nástroje pro podporu DevOps. Configuration Management:** jak a kde udržovat kód, jakým způsobem řídit verzování, oprava a sledování defektů. Release Management – automatizované nasazení infrastruktury a kódu.
10. **Internet of Things:** IoT a jeho využití v praxi. Jakým způsobem postavit datové řešení založené na IoT čidlech.

## Doporučená literatura

1. Hadoop: The Definitive Guide, Autor Tom White
2. Deep Learning, Autor Ian Goodfellow, Yoshua Bengio, Aaron Courville
3. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, Autor Aurélien Géron
4. Effective DevOps: Building a Culture of Collaboration, Affinity, and Tooling at Scale, Autor: Jennifer Davis, Ryn Daniels