

Tantárgy neve: Haladó big data	Kreditértéke: 5
A tantárgy besorolása: kötelezően választható	
A tantárgy elméleti vagy gyakorlati jellegének mértéke, „képzési karaktere”: 50% (kredit%)	
<p>A tanóra típusa: ea. / gyak. / konz. és óraszám: 2 / 2 / 1 az adott félévben, (ha nem (csak) magyarul oktatják a tárgyat, akkor a nyelve:)</p> <p>Az adott ismeret átadásában alkalmazandó további (sajátos) módok, jellemzők: A gyakorlat során Kaggle vagy hasonló big data/data science versenyek feladatainak önálló megoldása és beszámoló készítése.</p>	
<p>A számonkérés módja (koll. / gyj. / egyéb): koll. / gyj.</p> <p>Az ismeretellenőrzésben alkalmazandó további (sajátos) módok: A gyakorlat során egy nagyobb feladat önálló megoldása, dokumentálása és bemutatása szükséges a gyakorlati jegy megszerzéséhez.</p>	
A tantárgy tantervi helye (hányadik félév): 3. félév	
Előtanulmányi feltételek (ha vannak): nincs	
Tantárgy-leírás: az elsajátítandó ismeretanyag tömör, ugyanakkor informáló leírása	
<p>Az informatika hatalmas fejlődésen ment keresztül az elmúlt tíz évben. Ma már életünk és világunk szinte minden entitásáról és változásáról valamilyen szintű digitális lenyomattal is rendelkezünk (pl. SmartCity, banki rendszerek, viselhető szenzorok, IoT, mobil hálózatok, közösségi hálózatok, stb.). Az informatikai rendszerek által generált egyre nagyobb adat tömegek a tárolási problémák mellett is számos kihívást jelentenek. Ezen adatok sokszor önmagukban semmitmondók, értéket csak megfelelő tisztítás, elemzés, értelmezés és együttes használat során nyernek. A kis adattömegek esetén jól működő algoritmusok és módszerek sok esetben nagy, akár több terrabájtos adat mellett már nem nyújtanak megfelelő teljesítményt vagy egyáltalán nem alkalmazhatóak. Ezen folyamat eredményeképp az informatikai szektor is megváltozott. A piacon egyre nagyobb az igény olyan szakemberek iránt, akik az adatok megfelelő feldolgozását és értelmezését el tudják végezni, megfelelő gyakorlati és elméleti ismeretekkel rendelkeznek nagy adattömegek tárolására, hatékony feldolgozására használható rendszerek üzemeltetéséről és használatáról.</p> <p>A tárgy elsősorban a big data rendszerek hatékony üzemeltetésének és felhasználásának problémáival foglalkozik és az alapok helyett elsősorban a „legjobb gyakorlatok” bemutatására fókuszál gyakorlati példákon keresztül.</p> <p>Részletes tematika:</p> <ol style="list-style-type: none"> 1. Hadoop klaszter menedzsment: telepítés, konfiguráció 2. Környezetek: Hadoop, Spark 3. NoSQL a Hadoop környezetben: HBase 4. Elosztott üzenet kezelés: Kafka, RabbitMQ 5. Adatbegyűjtés és aggregáció: Flume 6. Tárolás Hadoop környezetben: HIVE, PIG 7. MapReduce algoritmusok 8. Hadoop ökoszisztéma összekapcsolása 9. Gyakorlati Hadoop problémák: célzott hirdetések, tranzakció analízis, churn elemzés, IoT, tudományos számítások, anomália detektálás, adatvizualizáció 	
A legfontosabb kötelező, illetve ajánlott irodalom (jegyzet, tankönyv) felsorolása bibliográfiai adatokkal (szerző, cím, kiadás adatai, (esetleg oldalak), ISBN)	
<ol style="list-style-type: none"> 1. Sandeep Karanth: Mastering Hadoop, ISBN 9781783983643 © Packt Publishing Limited. 2. Chuck Lam: Hadoop in Action 3. Alexey Yakubovich, Boris Lublinsky, and Kevin T. Smith : Professional Hadoop Solutions 	

Azoknak az előírt szakmai kompetenciáknak, kompetencia-elemeknek a felsorolása, amelyek kialakításához a tantárgy jellemzően, érdemben hozzájárul

a) tudása

- Átfogóan és naprakészen ismeri és érti az informatikai szakterületének általános elméleteit, összefüggéseit, tényanyagát és az ezekhez szükséges felépítő fogalomrendszert, különösen - választott specializációjának megfelelően - az alábbi területeken: osztott rendszerek, az adatbázis-kezelő rendszerek, big data környezetek.
- Átfogóan és naprakészen ismeri az informatikai szakterületének tervezési, fejlesztési, működtetési és irányítási folyamatainak feladat-megoldási elveit, módszereit és eljárásait, különösen - választott specializációjának megfelelően - a következő területeken: adatbázis rendszerek tervezése, fejlesztése menedzselése, osztott rendszerek tervezése, felépítése, menedzselése

b) képességei

- Képes az informatikai szakterületen felmerülő komplex szakmai problémák formalizálására, a szükséges elvi és gyakorlati háttér feltárására és a probléma megoldására.
- Képes tervezési, fejlesztési, üzemeltetési és irányítási feladatok ellátására komplex szoftver rendszerek, adatbázis kezelő rendszerek, vállalati információs rendszerek, döntéstámogató rendszerek, szakértői rendszerek működtetése esetében.

Tantárgy felelőse (név, beosztás, tud. fokozat): Gombos Gergő, Adjunktus, PhD

Tantárgy oktatásába bevont oktató(k), ha van(nak) (név, beosztás, tud. fokozat):

Laki Sándor, Adjunktus, PhD