

Kiberfizikai Rendszerek Intézet			Mintatanterv szerinti 6. félév 2025-26-1			
Tantárgy neve:	Kódja:	Kredit:	Óraszám			
				ea	tg	lab
Big data algoritmusok és programozás	NKXBD1HBNE	5	nappali heti	2	0	2
Tárgyfelelős: Dr. Felde Imre			Beosztás: egyetemi tanár			
Oktató(k): Tikász Ádám						
Előtanulmányi feltételek:	NSXMG1HBNE	Mély gépi tanulás				
Számonkérés módja:	Vizsga					
A tananyag						
Oktatási cél:	A tantárgy célja a hallgatók bevezetése a Big Data technológiák témakörébe bemutatva ezen rendszerek elméleti és gyakorlati működését. Bemutatásra kerül a nagy méretű adatállomány gyűjtésének, tárolásának, feldolgozásának, vizualizációjának célja, folyamatai, lehetséges kivitelezései, háttere. A témákat a hallgatók gyakorlatban is feldolgozzák előre telepített és a labor során saját maguk által telepített, továbbá felhő technológián alapuló rendszerek segítségével. Bemutatásra kerül több különböző Big Data keretrendszer és adatfeldolgozó, adatvizualizáló és elemző technológia (Apache Hadoop, Spark, Kafka, Databricks) így a hallgatók több megközelítésből megismerik és elsajátítják az érintett témaköröket mind saját telepítésű, mind felhő alapú rendszerekben.					
Tematika:	A kurzus elvégzése során a hallgatók megismerik a Big Data rendszerek elméleti háttérét, a rendelkezésre álló, feldolgozatlan és nagy mennyiségű adatok kezelésének problémakörét, ezen adatok begyűjtésének, tárolásának, feldolgozásának és megjelenítésének folyamatait és lehetőségeit, analitikai technológiákat, biztonsági és etikai kérdéseit. Az egyes témakörök bemutatják a Big Data rendszerek ezen főbb folyamatait és gyakorlati megoldásokat kínálnak azok megvalósításához. A hallgatók gyakorlat során ma is használt adatfeldolgozási, tárolási, megjelenítési és analitikai technológiát fognak használni a feladatok elvégzése során, amik bemutatják a Big Data adatkezelés főbb folyamatait. Így megismerhetnek Open Source és üzleti modellű rendszereket a gyakorlatban, továbbá saját telepítésű és felhő alapú rendszerek használatát is elsajátíthatják.					

Féléves ütemezés	
Oktatási hét (konzultáció)	Témakör
1.	Előadás: Bevezetés a Big Data technológiába Labor: Bevezetés és munkakörnyezet megismerése
2.	Előadás: Hadoop keretrendszer Labor: Python és Jupiter alapok adatfeldolgozáshoz és analitikához
3.	Előadás: BigData adattárolás és fájlformátumok Labor: Fileformátumok, adat import és export
4.	Előadás: Adatfeldolgozás általánosan Labor: Spark API feladatok
5.	Előadás: Batch processing Spark keretrendszerrel Labor: Batch processing
6.	Előadás: Apache Kafka Labor: Stream processing Kafka forrásból
7.	Előadás: Stream processing Spark és Storm keretrendszerekkel

	Labor: Cassandra
8.	Előadás: Adatelérés és vizualizáció Labor: Databricks platform
9.	Előadás: Analitika és gépi tanulás Labor: Databricks Delta Lake
10.	Előadás: Folyamatok automatizálása Labor: Databricks vizualizáció és analitika
11.	Előadás: Big Data esettanulmányok, etikai kérdések Labor: Databricks Workshop
12.	Előadás: Hadoop cluster management Labor: Biztonsági kérdések, féléves feladat konzultáció
13.	Zárthelyi dolgozat
14.	Pót zárthelyi dolgozat
Félévközi követelmények	
Évközi jegy / aláírás megszerzésének feltételei:	Az aláírás megszerzéséhez a zárthelyi dolgozatokon és a beadandó feladatokon külön-külön legalább 50%-os eredmény elérése szükséges. Az órákon való részvétel legalább 70%-ban.
Zárthelyi dolgozatok	
Oktatási hét	Témakör
6.	Kis zárthelyi dolgozat az első öt elméleti és gyakorlati témakör tananyagából
13.	Zárthelyi dolgozat – a teljes félév elméleti és gyakorlati tananyagából
14.	Pót zárthelyi dolgozat – a teljes félév elméleti és gyakorlati tananyagából
Vizsgaidőszak	Aláírás pótló dolgozat – a teljes félév elméleti és gyakorlati tananyagából
Az évközi jegy kialakításának módszere (csak évközi jegyes tárgyak esetében töltendő ki)	
-	
Pótlás módja	
A ZH / évközi jegy / aláírás pótlásának módja:	Amennyiben a zárthelyi dolgozat nem éri el az 50%-os eredményt vagy nem lett megírva, a 14. héten lehetőség van egy pót zárthelyi dolgozat megírására. Amennyiben sem a zárthelyi, sem a pót zárthelyi dolgozat nem éri el az 50 %-ot, úgy a vizsgaidőszakban lehetőség van egy aláírás pótló dolgozat megírására. A féléves feladat pótlása (dokumentáció és/vagy bemutatás) lehetséges: • Pót zárthelyi dolgozat (14. hét) idejében 25 % pontlevonással. • Aláírás pótló dolgozat idejében (vizsgaidőszak) 50 % pontlevonással. Az aláírás pótló zárthelyi vagy az aláírás pótláskor leadott féléves feladat a fent leírtak szerint és a HKR-ben előírtaknak megfelelően, a vizsgaidőszak első 10 munkanapjának valamelyikén az aláírás pótlási díj ellenében pótolható.
Vizsga módja (csak vizsgás tantárgy esetében töltendő ki)	
Írásbeli vizsga	
Vizsgajegy kialakítása (csak vizsgás tantárgy esetében töltendő ki)	
A tantárgy teljesítéséhez a zárthelyi dolgozaton, beadandó feladatokon és a vizsgán külön-külön legalább 50%-os eredmény elérése szükséges. Ha ezek mind teljesülnek, akkor az elért pontok összegéből alakul ki a végső érdemjegy.	

Feladat		Maximum pont
Gyakorlaton írt zárthelyi dolgozatok eredménye		28
Gyakorlaton beadott féléves feladatok eredménye		22
Vizsga eredménye		50
Összesen		100
Az egyes érdemjegyek ponthatárai:		
	Elérhető eredmény	Érdemjegy
	0-59	Elégtelen (1)
	60-69	Elégséges (2)
	70-79	Közepes (3)
	80-89	Jó (4)
	90-100	Jeles (5)
Irodalom		
Kötelező:	Alex Holmes: Hadoop In Practice, 2nd Edition, September 2014, ISBN 978-1-617-29222-4	
Ajánlott:	– Dirk deRoos, Paul C. Zikopoulos, Roman B. Melnyk PhD, Bruce Brown, Rafael Coss: Hadoop for Dummies, 2014 John Wiley & Sons, Inc., Hoboken, New Jersey, ISBN 978-1-118-65220-6 – Elmasri, R., Navathe, S. B.: Fundamentals of Database Systems 7th Edition, ISBN: 978-0133970777	
Egyéb segédletek:		