

Óbudai Egyetem				
Alba Regia Műszaki Kar				
Tantárgy neve és kódja: Big Data algoritmusok és programozás (AMEBD0IBNE)				
Kreditérték: 5				
Nappali tagozat 2023/24. tanév 2. félév				
Szakok, melyeken a tárgyat oktatják: Mérnökinformatikus alapszak (BSc)				
Tantárgyfelelős oktató:	Dr. Felde Imre		Oktató:	Piglerné dr. Lakner Rozália
Előtanulmányi feltételek: (kóddal)	NIXKD1GBNE	Korszerű adatbázisok		
Heti óraszámok:	Előadás: 2	Tantermi gyak.: 0	Laborgyakorlat: 2	Konzultáció: -
Számonkérés módja:	vizsga			
A tananyag				
<i>Oktatási cél:</i> A tárgy keretében a hallgatók megismerkednek a Big Data fogalomkörhöz kapcsolódó technológiákkal, paradigmákkal, komponensekkel, alkalmazási területekkel, ezen a területen alkalmazott hardver és szoftver eszközökkel és iparági sajátosságokkal.				
<i>Témakörök:</i> Az Apache Hadoop keretrendszer, fájlrendszer, erőforrás kezelés, MapReduce paradigma. Adatelemzési alapfogalmak, előrejelzés alapok, data science. A nyílt forráskódú csomagok és lekérdező eszközök áttekintése. Adatbányászati alapok.				
<i>Gyakorlat:</i> a Python programozás alapjai és a fontosabb data science könyvtárak.				
<i>Tematika:</i>				
Témakör				Óraszám
Előadások:				
1. Bevezetés. Big Data rendszerek általános jellemzői				2
2. Hadoop, HDFS, MapReduce - alapelvek, tulajdonságok, eszközök				2
3. Big Data adatmodellek, strukturált, félig-strukturált és nem-strukturált adatok. Az adatmodellek kezelésének eszközei				2
4. Felhő alapú platformok a gyakorlatban – Adatvizualizáció Tableau-ban				2
5. Felhő alapú platformok a gyakorlatban - Databricks és Spark alapok				2
6. Felhő alapú platformok a gyakorlatban - Databricks és Spark esettanulmányok				2
7. Rektori szünet				
8. Felhő alapú platformok a gyakorlatban – Big Data kezelése Snowflake-ben				2
9. Big Data rendszerek elemzése, adatbányászati alapok				2
10. Big Data rendszerek elemzésének eszközei				2
11. Gráf adatbázisok és elemzésének eszközei				2
12. Big Data esettanulmány I.				2
13. Big Data esettanulmány II.				2
14. Önálló feladat beadása				2
Gyakorlatok:				
1. A Python és környezetének ismertetése, telepítése. Python alapok.				2
2. Adatelemzés - NumPy				2
3. Adatelemzés - Pandas				2
4. MapReduce - Mrjob				2
5. Adatvizualizáció - Matplotlib				2
6. Adatvizualizáció – Seaborn, Plotly és Cufflinks				2
7. Rektori szünet				
8. Esettanulmány				2
9. Gépi tanulás – scikit-learn				2

10. Osztályozás és csoportosítás	2
11. Regressziós modellek	2
12. Ajánló rendszerek	2
13. Természetes nyelvfeldolgozás	2
14. Önálló feladat beadása	2
Félévközi követelmények	
Aláírás feltétele:	Részvétel az előadásokon és a laboratóriumi gyakorlatokon. A laboratóriumi gyakorlatok anyagából a hallgatók kis beadandó feladatokat készítenek. Az aláírás feltétele minden kis beadandó feladat sikeres teljesítése (legalább 50%) és egy nagyobb önálló feladat 14. héten történő beadása és bemutatása.
A vizsga módja (írásbeli, szóbeli, teszt, stb): szóbeli	

Irodalom:	
Kötelező:	Az egyetem e-Learning rendszerén közzétett tananyag
Ajánlott:	Bögel Gy.: A Big Data ökoszisztémája, Typotex kiadó, 2015 Han, J., M. Kamber: Adatbányászat. Konceptiók és technikák. Panem Kft., 2004 Harrison, G.: Next Generation Databases: NoSQL, NewSQL, and Big Data. Apress, 2015 Manyika J., Chui M., Brown B., Bughin J., Dobbs R., Roxburgh C., Byers A.H.: Big Data, the Next Frontier for Innovation, Competition and Productivity. McKinsey Global Institute, 2011