

Emri i Lëndës : Datamining II							
Kodi	Tipi	Semestri	Leksione (orë/javë)	Seminare (orë/javë)	Lab (orë/javë)	Kredite	ECTS
CMP 509	C	Vjeshtë	3.00	1.00	0.00	3.50	6.00
Lektori		Eris Zeqo, PhD					
Asistenti		Damiana Teliti, Msc					
Gjuha e kursit		Shqip					
Niveli i lëndës		Master					
Përshkrimi		"Datamining II" fokusohet në teknikat e avancuara të zbulimit të njohurive nga të dhënat, përfshirë analizën e të dhënave të mëdha, klasifikimin e avancuar, grupimin, zbulimin e anomalive dhe teknikat për të dhëna të paqëndrueshme ose me rrjedhje të vazhdueshme (data streams). Përdorimi i algoritmeve të përparuara në kontekste reale do të jetë fokusi kryesor.					
Objektivat		Të eksplorojnë teknikat e avancuara të analizës së të dhënave. Të aplikohen algoritme klasifikimi dhe grupimi në skenarë të ndërlukuar. Të analizohen të dhëna të paqëndrueshme dhe të dhëna që rrjedhin në kohë reale. Të kuptohet zbulimi i anomalive dhe teknikat e personalizimit.					
Konceptet Kryesore		Klasifikim i avancuar dhe përmirësim i performancës Grumbullim hierarkik, fuzzy, DBSCAN Minierimi i të dhënave në rrjedhë (stream mining) Zbulimi i anomalive dhe outlier-ëve Reduktimi i dimensioneve (PCA, t-SNE) Përzgjedhja dhe kombinimi i modeleve (ensemble methods)					
Programi i Lëndës							
Java	Tema						
1	Hyrje dhe rishikim i teknikave nga Data Mining I						
2	Klasifikim i avancuar: Boosting dhe Bagging						
3	Përzgjedhja e veçorive dhe reduktimi i dimensioneve						
4	Grumbullimi fuzzy dhe DBSCAN						
5	Minierimi i të dhënave të rrjedhshme						
6	Zbulimi i anomalive dhe outlier-ëve						
7	Teknikat e personalizimit dhe rekomandimi						
8	Provimi Gjysëmfinal						
9	Minierimi i të dhënave tekstuale dhe të pa strukturuara						
10	Metoda ensemble: Random Forests, Stacking						
11	Vlerësimi i performancës dhe validimi ndërkryq						
12	Përdorimi i minierimit në të dhënat e mëdha						
13	Mjete praktike (Weka, Scikit-learn, RapidMiner)						
14	Etika dhe privatësia në analizën e të dhënave						
15	Prezantimi i projekteve finale / Rishikim						
16	Provimi Final						

Parakushtet	Studenti duhet të frekuentojë lëndën në masën minimale prej 75%.		
Literatura	<ul style="list-style-type: none"> • Jiawei Han, Micheline Kamber, Jian Pei – Data Mining: Concepts and Techniques, 4th Edition, Morgan Kaufmann, 2022. 		
Referenca të tjera	<ul style="list-style-type: none"> • Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar – Introduction to Data Mining, 2nd Edition, Pearson, 2019. • Charu C. Aggarwal – Data Mining: The Textbook, Springer, 2015. • Trevor Hastie, Robert Tibshirani, Jerome Friedman – The Elements of Statistical Learning, Springer, 2017. 		
Rezultatet e Lëndës dhe Kompetencat			
1	Studentët do të jenë në gjendje të zbatojnë teknika të avancuara të minierimit të të dhënave në të dhëna reale dhe të mëdha.		
2	Do të jenë në gjendje të përzgjedhin algoritmet më të përshtatshëm për klasifikim, grupim ose zbulim anomalish.		
3	Do të përdorin mjete praktike për analiza të dhënash të strukturuar dhe të pa strukturuar.		
4	Do të kuptojnë rëndësinë e etikës dhe mbrojtjes së privatësisë në analizën e të dhënave.		
Mënyra e Vlerësimit të Lëndës			
Notat e Ndërmjetme		Sasia	Përqindja
Gjysmë finale		0	0
Kuize		0	0
Projekte		1	20
Projekte semestrale		1	20
Punë laborator		0	0
Pjesëmarrja në mësim		0	0
Kontributi i notave të ndërmjetme mbi vlerësimin final			40
Kontributi i provimit final mbi vlerësimin final			60
Total			100
Ngarkesa ECTS (Në Bazë të Ngarkesës së Studentit)			
Aktivitetet	Sasia	Kohëzgjatja (orë)	Ngarkesa Totale (orë)
Kohëzgjatja e kursit (Duke përfshirë edhe javën e provimeve : 16x Orët totale të kursit)	16	4	64
Orët e studimit jashtë klase (Parapërgatitje, Praktika etj)	14	2	28
Detyra	2	24	48
Gjysmë finale	0	0	0
Provimi final	1	10	10
Të tjera	0	0	0
Ngarkesa totale e orëve			150
Ngarkesa totale e orëve / 25 (orë)			6.00
ECTS			6.00