

Emri i Lëndës : Datamining I							
Kodi	Tipi	Semestri	Leksione (orë/javë)	Seminare (orë/javë)	Lab (orë/javë)	Kredite	ECTS
CMP 401	B	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		This course explores the concepts and techniques of knowledge discovery and data mining. As a multidisciplinary field, data mining draws on work from areas including statistics, machine learning, pattern recognition, database technology, information retrieval, network science, knowledge-based systems, artificial intelligence, high-performance computing, and data visualization. This course focuses on issues relating to the feasibility, use-fulness, effectiveness, and scalability of techniques for the discovery of patterns hidden in large data sets. As a result, this course is not intended as an introduction to statistics, machine learning, database systems, or other such areas, although it does provide some background knowledge to facilitate the reader's comprehension of their respective roles in data mining.					
Objektivat							
Konceptet Kryesore		Data Mining, Llojet e të Dhënave, Metodot Kernel, Të dhënat Shumë Dimensionale, Gjetja e Modeleve të Shpeshta, Grupimi, Reduktimi i Dimensioneve, Rregullat e Shoqërimit, K-means dhe Pemët e vendimit					
Programi i Lëndës							
Java	Tema						
1	Hyrje në Data Mining Në këtë leksion do të realizohet një njohje e përgjithshme e lëndës si dhe do të trajtohen tema të tilla si çfarë është Data Mining, origjina dhe arsyeja e zhvillimit të Data Mining si dhe detyrat kryesore që mund të kryhen me anë të Data Mining. (Lit. bazë, fq. 21-42)						
2	Llojet e të Dhënave Në këtë leksion do të trajtohen tema të tilla si vetitë, llojet dhe vlerat e attributeve, kategorizimi dhe transformimi i të dhënave, bashkësitë e të dhënave dhe llojet e tyre si dhe analizimi i cilësës së të dhënave për shkak të problematikave të matjeve dhe grumbullimit të të dhënave. (Lit. bazë, fq. 43-69)						
3	Parapërpunimi i të Dhënave - 1 Në këtë leksion do të trajtohen masat e ndryshme të ngjashmërisë dhe distancave. Llojet kryesore të distancave që do të trajtohen janë Distanca Euklidiane, Distanca Minkoëski, Distanca Mahalanobis, llojet kryesore të ndashmërive që do të trajtohen janë Ngjashmëria midis vektorëve binare, Ngjashmëria Kosinusoidale si dhe Korrelacioni Pearson. (Lit. bazë, fq. 91-110)						
4	Parapërpunimi i të Dhënave - 2 Në këtë leksion do të trajtohen teknika të para përpunimit të të dhënave dhi Agregimi, marrja e mostrave, reduktimi i dimensioneve, përzgjedhja e nëngrupit të veçorive, krijimi i veçorive të reja, diskretizimi, binarizimi, transformimi i variablave si dhe njësitet matëse të bazuara në informacion. (Lit. bazë, fq. 70-90)						
5	Eksplorimi i të Dhënave - 1 Në këtë leksion do të trajtohen elementet bazë të statistikave përmbledhëse të të dhënave, si llojet e mesatareve, llojet e shpërndarjeve, masat e ndryshme të ngjashmërisë dhe ndryshimeve midis objekteve të ndryshme të të dhënave, llojet e njësive matëse të afërsisë, informacioni i ndersjelltë si dhe teknikat në përzgjedhjen e njësisë matëse të duhur. (Lit. e Rekomanduar, fq. 44 – 55)						

6	Eksplorimi i të Dhënave - 2 Në këtë leksion do të trajtohen elementet bazë të statistikave përmbledhëse të të dhënave, si llojet e mesatareve, llojet e shpërndarjeve, masat e ndryshme të ngjashmërisë dhe ndryshimeve midis objekteve të ndryshme të të dhënave, llojet e njësive matëse të afërsisë, informacioni i ndersjelltë si dhe teknikat në përzgjedhjen e njësive matëse të duhur. (Lit. e Rekomanduar, fq. 56 – 64 si dhe Lit. bazë, fq. 110-132)
7	Klasifikimi: Konceptet dhe Teknikat Bazë - 1 Në këtë leksion do të trajtohen konceptet bazë të klasifikimit, klasifikimi multiklasor dhe binar, qasjet e përgjithshme për ndërtimin e modelit të klasifikimit, metodat bazë për paraqitjen e kushteve të testimit, llogaritja e masave të ndryshme të papastërtisë për lloje të ndryshme të dhënash si dhe algoritmet bazë të klasifikimit. (Lit. bazë, fq. 133-167)
8	Provimi gjysmë final
9	Klasifikimi: Konceptet dhe Teknikat Bazë - 2 Në këtë leksion do të trajtohen temat e mbipërshtatshmërisë së modelit klasifikues së përzgjedhur, vlerësimi dhe përzgjedhja i modeleve të ndryshme të klasifikimit, hiper parametrat si dhe kufizimet e algoritmeve bazë të klasifikimit. (Lit. bazë, fq. 167-212)
10	Rregullat e Shoqërimit: Konceptet dhe Algoritmat Bazë Në këtë leksion do të trajtohen konceptet dhe algoritmat bazë të rregullave të shoqërimit si principi Apriori, gjenerimi i grupeve të shpeshta të artikujve, përzgjedhja e kandidatëve, teknikat dhe metodat për gjenerimin e rregullave të shoqërimit si dhe kompleksiteti kompjuterik i algoritmeve bazë të rregullave të shoqërimit. (Lit. bazë, fq. 213-239)
11	Rregullat e Shoqërimit: Problematikat në përzgjedhjen dhe vlerësimin e modelit Në këtë leksion do të trajtohen tematika me të detajuara të rregullave të shoqërimit si përfaqësimi kompakt i grupeve të shpeshta të artikujve, metodat alternative për gjenerimin e grupeve të shpeshta të artikujve, algoritmi i rritjes FP, vlerësimi i modeleve të shoqërimit si dhe efekti i shpërndarjeve të njëanshme. (Lit. bazë, fq. 240-306)
12	Analiza e Grupimeve: Konceptet dhe Algoritmat Bazë Në këtë leksion do të trajtohen konceptet dhe algoritmat bazë të analizës së grupimeve si çfarë është Analiza e Grupimeve, llojet e ndryshme të mënyrave të grupimit, llojet e ndryshme të grupimeve si dhe një analizë e detajuar e algoritmit K-means. (Lit. bazë, fq. 307-335)
13	Analiza e Grupimeve: Problematikat në përzgjedhjen dhe vlerësimin e modelit Në këtë leksion do të trajtohen konceptet dhe algoritmat bazë të analizës së grupimeve si grumbullimi hierarkik aglomerativ, trajtimi i detajuar i algoritmit DBSCAN, si dhe teknikat dhe metodat e ndryshme për vlerësimin e grupimeve. (Lit. bazë, fq. 336-394)
14	Klasifikimi: Teknikat Alternative Në këtë leksion do të trajtohen llojet e klasifikuesve, klasifikuesit e bazuar në rregulla, klasifikuesit e fqinjit më të afërt, klasifikuesit Naive Bayes, Regresioni logjistik, Rrjeti Neural Artificial si dhe Support Vector Machine. (Lit. bazë, fq. 395-463, 478 - 498)
15	Prezantimi i Projekteve dhe Përsëritje
16	Provimi Final

Parakushtet	Studenti duhet të frekuentojë lëndën në masën minimale prej 75%.		
Literatura	• Introduction to Data Mining 2nd Edition 2018		
Referenca të tjera	• The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011		
Rezultatet e Lëndës dhe Kompetencat			
1	Studentët do të jenë të aftë të kuptojnë llojet e ndryshme të të dhënave.		
2	Studentët do të kenë njohuri mbi veçoritë dhe njësitë matëse statistikore të të dhënave.		
3	Studentët do të përvetësojnë konceptet më të rëndësishme në lidhje me modelet dhe algoritmet bazë për gjetjen e informacioneve me vlerë.		
4	Studentët do të jenë të aftë të kuptojnë rëndësinë e Data Mining në gjetjen e informacioneve me vlerë.		
5	Studentët do të jenë të gatshëm për implementuar në praktikë njohuritë bazë të dhëna.		
6	Studentët do të jenë të pajisur me njohuri të mjaftueshme teorike dhe praktike për të vijuar me lëndët e tjera pasardhëse.		
Mënyra e Vlerësimit të Lëndës			
Notat e Ndërmjetme	Sasia	Përqindja	
Gjysmë finale	0	0	
Kuize	0	0	
Projekte	1	20	
Projekte semestrare	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