

Titulli i lëndës: Përgatitja dhe vizualizimi i të dhënave

Informatat themelore për lëndën	
Njësia akademike:	Fakulteti i Inxhinierisë Elektrike dhe Kompjuterike
Titulli i lëndës:	Përgatitja dhe vizualizimi i të dhënave
Niveli:	M.Sc
Statusi i lëndës:	Obligative
Viti i studimeve:	Semestri I-rë
Numri i orëve në javë:	2+0+2
Kredit ECTS:	6
Koha/ Vendi:	TD / Klasa
Mësimdhënësi:	PhD Mërgim H. HOTI
Të dhënat kontaktuese:	mergim.hoti@uni-pr.edu
Përshkrimi i lëndës:	<p>Grumbullimi i të dhënave: qasja në të dhëna (API-të), shprehjet e rregullta, crawling, crowdsourcing.</p> <p>Para-procesimi për përgatitjen e të dhënave për analizë: tipet e të dhënave, mbledhja e të dhënave, kualiteti i të dhënave, integrimi, agregimi, mostrimi, pastrimi, vlerat e zbrazëta dhe të dhënat e rradha, redukimi i dimensionit, zgjedhja e nën bashkësisë së vetive, krijimi i vetive, diskretizimi dhe binarizimi, transformimi. Detektimi i përjashtuesit. Të dhënat me klasa të shtrembëra. Rëndësia për vëllim të dhënash. Madhësitë e ngjashmërisë dhe jo ngjashmërisë. Mënjanimi i zbulimeve jo të sakta. Eksplorimi i të dhënave: statistika përmbledhëse, statistika përmbledhëse multivariate. Vizualizimi i të dhënave: vizualizimi sipas tipeve të dhënave, vizualizimi statik dhe ineraktiv, vizualizimi i të dhënave shumë-dimensionale, dashboard-et dhe storyboard-et, librabritë dhe veglat e ndryshme të vizualizimit.</p>
Qëllimet e lëndës:	Ky kurs mbulon përgatitjen e të dhënave për t'u analizuar përmes teknikave të machine learning, si dhe vizualizimin e tyre para, përgjatë dhe pas analizës.
Rezultatet e pritshme të nxënies:	Studenti pas përfundimit të këtij kursi do të jetë në gjendje të:

		<ol style="list-style-type: none"> 1. Kuptoj dhe përdor teknikat e përshtatshme për të grumbulluar dhe përgatitur të dhënat për analizë të caktuar. 2. Vizualizon të dhënat për eksplorim, dhe të dhënat në dalje të analizës. 3. Interpretoj vizualizimin. 4. Njoh dhe trajtoj problemet e ndryshme gjatë analizës shkaku i të dhënave. 		
Rëndësia dhe aktualiteti i lëndës:	Lënda është bazike në lëmine e shkencës së të dhënave, e cila është në rritje e sipër për nga relevanca në pothuajse secilin sektor të zhvillimit në shoqëri.			
Ngarkesa e studentit (duhet të jetë në përputhje me Rezultatet e Nxënies të studentit)				
Aktiviteti	Orë mësimore	Ditë/ Javë	Gjithsej	
Ligjëratat	3	15	45	
Teori/Punë në laborator/Ushtrime	1	15	15	
Punë praktike	1	15	15	
Përgatitje për test intermediar				
Konsultime me mësimdhënësin	1	5	5	
Puna në terren				
Testi, punimi seminarik				
Detyrë shtëpie				
Mësimi individual (në bibliotekë apo në shtëpi)	3	15	45	
Përgatitja për provimin final				
Koha e vlerësimit (testi, kuizi, provimi final)				
Projektet, prezantimet, etj.	2	15	30	

Total			155
Metodat e mësimdhënies:	45 orë ligjërata, 30 ushtrime laboratorike dhe praktike, 5 orë konsultime. Afërsisht 75 orë punë të pavarur përfshirë punimin e projektit.		
Metodat e vlerësimit:	Detyra e projektit 40% (2 x 15%, 1 x 10). Provimi final 50%.		
Vijueshmëria:	Ligjërata 5% Ushtrime 5%		
Literatura Primare:	<ol style="list-style-type: none"> 1. Introduction to Data Mining (2nd Edition). Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar. Pearson, 2018. 2. Principles of Database Management: The Practical Guide to Storing, Managing and Analyzing Big and Small Data (1st Edition). Wilfried Lemahieu, Bart Baesens, Seppe vanden Broucke. Cambridge University Press. 2018. 3. Mining of Massive Datasets (2nd Edition). Jure Leskovec, Anand Rajaraman, Jeff Ullman. Cambridge University Press. 2014. 4. Storytelling With Data: A Data Visualization Guide for Business Professionals (1st Edition). Cole Nussbaumer Knaflic. Wiley, 2015. 		
Literatura shtesë:			
Hartimi i planit mësimor			
Java	Titulli i ligjërates	Ushtrimet	
Java 1	Grumbullimi i të dhënave: qasja në të dhëna (API-të), shprehjet e rregullta, crawling, crowdsourcing.	Grumbullimi i të dhënave: qasja në të dhëna (API-të), shprehjet e rregullta, crawling, crowdsourcing.	
Java 2	Grumbullimi i të dhënave: qasja në të dhëna (API-të), shprehjet e rregullta, crawling, crowdsourcing.	Grumbullimi i të dhënave: qasja në të dhëna (API-të), shprehjet e rregullta, crawling, crowdsourcing.	
Java 3	Para-procesimi për përgatitjen e të dhënave për analizë: tipet e të dhënave, mbledhja e të dhënave, kualiteti i të dhënave, integrimi, agregimi, mostrimi, pastrimi, vlerat e zbrazëta dhe të dhënat e rradha, redukimi i dimensionit, zgjedhja e nën bashkësisë së vetive, krijimi i vetive,	Para-procesimi për përgatitjen e të dhënave për analizë: tipet e të dhënave, mbledhja e të dhënave, kualiteti i të dhënave, integrimi, agregimi, mostrimi, pastrimi, vlerat e zbrazëta dhe të dhënat e rradha, redukimi i dimensionit, zgjedhja e nën bashkësisë së vetive, krijimi i vetive,	

	diskretizimi dhe binarizimi, transformimi.	diskretizimi dhe binarizimi, transformimi.
Java 4	Para-procesimi për përgatitjen e të dhënave për analizë: tipet e të dhënave, mbledhja e të dhënave, kualiteti i të dhënave, integrimi, agregimi, mostrimi, pastrimi, vlerat e zbrazëta dhe të dhënat e rradha, redukimi i dimensionit, zgjedhja e nën bashkësisë së vetive, krijimi i vetive, diskretizimi dhe binarizimi, transformimi.	Para-procesimi për përgatitjen e të dhënave për analizë: tipet e të dhënave, mbledhja e të dhënave, kualiteti i të dhënave, integrimi, agregimi, mostrimi, pastrimi, vlerat e zbrazëta dhe të dhënat e rradha, redukimi i dimensionit, zgjedhja e nën bashkësisë së vetive, krijimi i vetive, diskretizimi dhe binarizimi, transformimi.
Java 5	Para-procesimi për përgatitjen e të dhënave për analizë: tipet e të dhënave, mbledhja e të dhënave, kualiteti i të dhënave, integrimi, agregimi, mostrimi, pastrimi, vlerat e zbrazëta dhe të dhënat e rradha, redukimi i dimensionit, zgjedhja e nën bashkësisë së vetive, krijimi i vetive, diskretizimi dhe binarizimi, transformimi.	Para-procesimi për përgatitjen e të dhënave për analizë: tipet e të dhënave, mbledhja e të dhënave, kualiteti i të dhënave, integrimi, agregimi, mostrimi, pastrimi, vlerat e zbrazëta dhe të dhënat e rradha, redukimi i dimensionit, zgjedhja e nën bashkësisë së vetive, krijimi i vetive, diskretizimi dhe binarizimi, transformimi.
Java 6	Detektimi i përjashtuesit. Dorëzimi dhe prezantimi i projektit - pjesa 1-rë. (03.11.2024)	Detektimi i përjashtuesit. Dorëzimi dhe prezantimi i projektit - pjesa 1-rë.
Java 7	Të dhënat me klasa të shtrembëra.	Të dhënat me klasa të shtrembëra.
Java 8	Rëndësia për vëllim të dhënash.	Rëndësia për vëllim të dhënash.
Java 9	Madhësitë e ngjashmërisë dhe jo ngjashmërisë.	Madhësitë e ngjashmërisë dhe jo ngjashmërisë.
Java 10	Mënjanimi i zbulimeve jo të sakta.	Mënjanimi i zbulimeve jo të sakta.
Java 11	Eksplorimi i të dhënave: statistika përmbledhëse, statistika përmbledhëse multivariate. Dorëzimi dhe prezentimi i projektit - pjesa 2. (08.12.2024)	Eksplorimi i të dhënave: statistika përmbledhëse, statistika përmbledhëse multivariate. Dorëzimi dhe prezentimi i projektit - pjesa 2.
Java 12	Eksplorimi i të dhënave: statistika përmbledhëse,	Eksplorimi i të dhënave: statistika përmbledhëse,

	statistika përmbledhëse multivariate.	statistika përmbledhëse multivariate.
Java 13	Vizualizimi i të dhënave: vizualizimi sipas tipeve të të dhënave, vizualizimi statik dhe ineraktiv, vizualizimi i të dhënave shumë-dimensionale, dashboard-et dhe storyboard-et, libraritë dhe veglat e ndryshme të vizualizimit.	Vizualizimi i të dhënave: vizualizimi sipas tipeve të të dhënave, vizualizimi statik dhe ineraktiv, vizualizimi i të dhënave shumë-dimensionale, dashboard-et dhe storyboard-et, libraritë dhe veglat e ndryshme të vizualizimit.
Java 14	Vizualizimi i të dhënave: vizualizimi sipas tipeve të të dhënave, vizualizimi statik dhe ineraktiv, vizualizimi i të dhënave shumë-dimensionale, dashboard-et dhe storyboard-et, librabritë dhe veglat e ndryshme të vizualizimit.	Vizualizimi i të dhënave: vizualizimi sipas tipeve të të dhënave, vizualizimi statik dhe ineraktiv, vizualizimi i të dhënave shumë-dimensionale, dashboard-et dhe storyboard-et, librabritë dhe veglat e ndryshme të vizualizimit.
Java 15	Vizualizimi i të dhënave: vizualizimi sipas tipeve të të dhënave, vizualizimi statik dhe ineraktiv, vizualizimi i të dhënave shumë-dimensionale, dashboard-et dhe storyboard-et, libraritë dhe veglat e ndryshme të vizualizimit. Dorëzimi dhe prezentimi i projektit - pjesa 3 (29.12.2024)	Vizualizimi i të dhënave: vizualizimi sipas tipeve të të dhënave, vizualizimi statik dhe ineraktiv, vizualizimi i të dhënave shumë-dimensionale, dashboard-et dhe storyboard-et, libraritë dhe veglat e ndryshme të vizualizimit. Dorëzimi dhe prezentimi i projektit - pjesa 3

Politikat akademike dhe Kodi i Sjelljes
Ora mësimore fillon dhe përfundon me kohë. Mjetet që përdorën gjatë orëve të mësimimit duhet të pastrohen dhe të ruhen në fund të orës mësimore. Telefonat mobil/të mençur dhe pajisjet tjera elektronike (p.sh. iPod-ët) duhet të fikën (apo të kurdisen në vibrim) dhe të mos ekspozohen gjatë orëve të mësimimit. Laptopët dhe kompjuterët tabletë lejohen të përdorën vetëm në heshtje; aktivitetet tjera siç janë kontrollimi i e-mailit personal apo shfletimi i ueb-faqeve në internet janë të ndaluara.

Shënim | Nëse 3 detyra të klasës të një studenti vlerësohen nën 50%, atëherë ai/ajo do ta humb të drejtën që t'i nënshtrohet provimit final. Vlerësimi bëhet nga 0-100 %.