

MD-2045, CHIȘINĂU, STR. STUDENȚILOR, 9/7, TEL: 022 50-99-01 | FAX: 022 50-99-05,

www.utm.md**ȘTIINȚA ȘI ANALIZA DATELOR****1. Date despre unitatea de curs/modul**

Facultatea	Calculatoare, Informatică și Microelectronică				
Catedra/departamentul	Informatică și Ingineria Sistemelor				
Ciclul de studii	Studii de master, CICLUL II				
Programul de studiu	Știința Datelor				
Anul de studiu	Semestrul	Tip de evaluare	Categoria formativă	Categoria de opționalitate	Credite ECTS
II	III	E	F	O	5

2. Timpul total estimat

Total ore în planul de învățământ	Din care				
	Ore auditoriale		Lucrul individual		
	Curs	Lucrări practice	Proiect de an	Studiul materialului teoretic	Pregătire aplicații
150	20	20	---	110	-

3. Precondiții de acces la unitatea de curs/modul

Conform planului de învățământ	Studiul se bazează pe cunoștințele studenților acumulate la următoarele discipline: teoria probabilității, statistica matematică, metode de programare, tehnologii Web, Baze de Date, precum și managementul proiectelor, modelarea, principiile cercetării științifice etc. Adicional, masteranzii trebuie să cunoască, structura calculatorului, cel puțin unul din SO Windows, Linux, MacOS, precum și unele limbaje de programare cum ar fi: C#, C++, HTML, PHP, Python.
Conform competențelor	Masteranzii trebuie să aibă următoarele cunoștințe, competențe și abilități pentru a stăpâni disciplina academică "Știința și Analiza Datelor": metode de analiză a datelor, privind aplicarea sistemelor moderne de Gestiune a Bazelor de Date, inclusiv: MySQL, MS SQL Server, modelarea multidimensională a datelor, modelarea și aplicarea bazelor de date Operaționale și a Depozitelor de date, tehnologiilor OLTP, OLAP, BI, (e.g. VS, Power BI), un anumit grad de explorare a datelor cu algoritmi Data Mining. Python este principalul limbaj de programare pentru cursul Știința și Analiza Datelor. În mod deosebit este vorba despre astfel de biblioteci Python, cum ar fi, de exemplu - NumPy, Scipy, Pandas. Bibliotecile, precum SQLAlchemy, Psycopg2, MySQLdb și SQLite3, ce oferă modalități eficiente de conectare la o bază de date, de a crea tabele, de a insera și de a prelua date și de a efectua alte operațiuni de bază de date, - toate sunt de asemenea condiții de lucru eficient în cadrul cursului și acumulare de cunoștințe.

	Adițional, masteranzii trebuie să poată lucra în grup, să aibă capacități de utilizare corectă a TIC în rezolvarea problemelor profesionale cu caracter aplicativ, folosind abilitățile de modelare matematică ale fenomenelor social-economice, ale mediului ambiant, din domeniul tehnic, precum și din domeniile: financiar-bancar, învățământ, medicină, asigurări, marketing etc. Rezultatele principale ale disciplinei “Știința și Analiza Datelor”, sunt utilizarea cunoștințelor, competențelor și abilităților în activitățile de Știința Datelor / Data Science, Data Engineering, Business Analytics, precum și în procesul de studiu al disciplinelor: “Analiza datelor”, “Știința Datelor”, “Analitica Big Data”, “Informatica în afaceri”, etc.
--	---

4. Obiectivele unității de curs

Obiectivul general	Cursul “Știința și Analiza Datelor”, împreună cu cursul “Tehnologii Big data” prezintă o componentă importantă în acumularea cunoștințelor, dezvoltarea deprinderilor, abilităților practice și dobândirea de competențe, în înțelegerea impactului transformator din aceste domenii asupra afacerilor, precum și care este diferența semnificativă dintre ele și metodele tradiționale de analiză, pentru aplicarea lor, în soluționarea celor mai diverse probleme din economia națională, utilizând simularea proceselor și fenomenelor din domenii de activitate proxime viitorilor specialiști în aplicarea TIC, instruiți în cadrul facultății Calculatoare Informatică și Microelectronică, la programul de studii – Știința Datelor. Astfel, centrată pe Analiza unei oarecare afaceri, procesul de aplicare a Științei în Analiza Datelor reprezintă rezultatele obținute de afacere într-o anumită perioadă de timp, care este direct proporțional cu viteza cu care sunt procesate volumele mari de date, Big Data, ce o reprezintă și care, bazată pe știință, face posibilă ca Analiza Datelor să fie una profesionistă, relevantă, credibilă, validată și predictivă, argumentată științific, capabilă să ofere recomandări bazate pe cunoștințe obținute urmare a procesării lor, generatoare de succese și valoare adăugată în afaceri. În acest context, Tehnologiile Big Data, curs complementar cursului dat de “Știința și Analiza Datelor”, permit la rândul lor a Analiza mai multe tipuri de date comparativ cu instrumentele de Business Intelligence, ceea ce face posibilă concentrarea nu numai pe stocarea structurată a datelor, dar și pe cea nestructurată. <u>Obiectivul general al cursului:</u> <i>cunoașterea</i> abordărilor științifice Știința Datelor / Data Science, a tehnologiilor de lucru cu date brute, Big Data, <i>și dezvoltarea</i> unor capacități și abilități de a le analiza și utiliza în orice domenii de activitate profesională, în particular, pentru suport în procesul decizional al unei entități implicate în menținerea climatului optim al dezvoltării socio-economice continue și durabile. <u>Obiectivele disciplinei:</u> • Studiul și analiza conceptelor de bază ale activității științifice, precum și a metodelor de cercetare științifică; • Studiul și analiza principiilor fundamentale ale proceselor de modelare, (modelare conceptuală, logică și fizică), a unor probleme din cele mai diverse domenii problematice de activitate profesională. • Studiul și analiza algoritmilor de Analiză a Datelor, inclusiv a abordărilor științifice, Știința Datelor / Data Science, și tehnologiilor de lucru cu date brute, Big Data.
---------------------------	---

	• Studiul și analiza sarcinilor de bază ale Analizei Datelor, utilizând instrumentele și tehnologiile din Știința Datelor / Data Science. • Studiul și analiza diverselor instrumente și tehnologii pentru aplicarea Științei Datelor / Data Science în procesarea și Analiza Datelor, cum ar fi ML și DML, inclusiv și a abordărilor științifice și tehnologiilor de lucru cu date brute, Big Data. • Aplicarea cunoștințelor cursului “Știința și Analiza Datelor” pentru a rezolva probleme cu caracter aplicativ, din orice domenii problematice de activitate profesională, în particular, pentru suport în procesul decizional al unei entități implicate în menținerea climatului optim al dezvoltării socio-economice continue și durabile. În particular masteranzii, urmare a acestui curs, trebuie să poată: • efectua lucrări practice de cercetare, privind implementarea și evaluarea eficienței tehnologiilor și instrumentelor Științei Datelor / Data Science și Big Data în entitățile economice (Business Big Data management), prin prisma cunoștințelor și abilităților obținute la cursul “Știința și Analiza Datelor”; • Implementa și aplica instrumente de analiză și suport decizional bazate pe tehnologiile și instrumentele Științei Datelor / Data Science și Big Data în managementul deciziilor (Decision management); • dezvolta noi modele ale infrastructurii informaționale ale unei entități, ținând cont de posibilitățile tehnologiilor Big Data și Știința Datelor / Data Science (Model management). • modela economic și matematic cele mai diverse domenii problematice de activitate profesională (economic and mathematical modeling) • aplica principiile, cunoștințele și abilitățile obținute la cursul dat pentru Afaceri Inteligente (Applying Data Analysis methods and Big Data in Business Intelligence)
Obiectivele specifice	Masteranzii, urmare a acestui curs, vor fi ancorați și în atingerea unor obiective specifice, ce derivă din faptul că procedura de lucru cu Big Data în Știința Datelor / Data Science poate fi prezentată ca un ciclu ce constă din următoarele cinci etape: • Colectarea datelor: la această etapă se determină volumul și structura datelor, care se colectează într-o formă convenabilă pentru următoarele etape. Tot aici se decide asupra obiectivelor/ipotezelor pentru analiza datelor și să selectează metodele necesare pentru etapele următoare; • Pregătirea datelor: această etapă include crearea unei arhitecturi de BD, sau DWH și a unui sistem de stocare/pipeline/conducă de date, ETL/ELT, precum și verificarea și ștergerea valorilor irelevante (validare); • Prelucrarea/procesarea datelor: la această etapă, se realizează gruparea și clasificarea (diviziunea în eșantioane și grupuri), sunt utilizate modele matematice și metode software, cum ar fi ML și DML; • Analiza datelor: un cercetător de date / Data Scientist studiază rezultatele modelării și caută răspunsuri la ipotezele inițial formulate, efectuează analize calitative și cantitative și după caz, identifică erori în domeniul problematic precăutat, fie în procesele socio-economice studiate și supuse analizei, fie în formularea de ipoteze, sau prognozează comportamentul unor fapte observate pentru viitor, în dependență de setul de caracteristici independente prestabilit.

• **Comunicarea rezultatelor obținute:** această etapă include vizualizarea faptelor obținute în baza modelelor identificate, precum și crearea de rapoarte analitice și propuneri de comportament: ce decizii pot fi luate pe baza cunoștințelor obținute, în domeniul de activitate profesională, în particular, pentru o entitate implicată în menținerea climatului optim al dezvoltării socio-economice continue și durabile.

În același timp, masteranzii vor avea posibilitatea să-și dezvolte și alte capacități de lucru și abilități specifice cum ar mai fi:

• extinderea cunoștințelor cu noi noțiuni fundamentale pentru cursul “Știința și Analiza Datelor”, cum ar fi: Big Data (date mari), Modelarea matematică, Data Mining/Miningul de date, Algoritmi Data Mining, Machine Learning, Feature Engineering, Deep Learning

• însușirea și a obținerea de cunoștințe noi, privind utilizarea tehnologiilor și instrumentelor ce stau la baza Data Science și Big Data, necesare pentru activități în domenii tangente cum ar fi Data Engineering și Inteligența Artificială;

• aplicarea TIC în curs de dezvoltare, cum ar fi: cloud computing, realitatea augmentată și virtuală, în activități de inovare și cercetare profundă a fenomenelor social-economice pentru a fi pregătiți să aplice tendințele tehnologice emergente în transformarea modelelor de cercetare și producere, reieșind din provocările la zi;

• dezvoltarea de noi capacități de aplicare a principiilor procesării distribuite a datelor și a utilizării arhitecturilor de calcul de înaltă performanță în acest context;

• dezvoltarea de noi capacități de a adapta soluții bazate pe abordări de tip data-driven pentru probleme specifice unui anumit domeniu de aplicabilitate.

• dezvoltarea de noi capacități de a planifica și organiza activitatea de cercetare într-un domeniu prin organizarea lucrului în grup; înțelegerea responsabilităților și respectarea regulilor de etică profesională, a celor de securitate și confidențialitate a datelor.

La nivel de cunoștințe:

• obținerea de cunoștințe și competențe pentru a utiliza tehnologiile și instrumentele ce derivă din esența și specificul cursului “Știința și Analiza Datelor” în activitățile de Data Engineering, Data Science și de Business Intelligence;

• dezvoltarea de competențe interdisciplinare, prin utilizarea Ip. Python, platformelor Power BI, Visual Studio, tehnologiilor cloud și platformelor cloud cum ar fi: AWS, Azure, Google Cloud, pentru Analiza Datelor și pentru Analitica Big Data;

• dezvoltarea de competențe care oferă o înțelegere a arhitecturii afacerii, impactul introducerii noilor tehnologii IT, inclusiv instrumentelor, tehnologiilor Data Science și Big Data, asupra eficienței managementului întreprinderii, schimbarea lanțului valoric;

• dezvoltarea de competențe de management în domeniul implementării sistemelor și serviciilor de Data Science și Big Data;

• dezvoltarea și obținerea de noi cunoștințe și competențe în utilizarea unor metode și tehnici eficiente de învățare, informare și specializare; abilitatea de a selecta și analiza critic resursele bibliografice; asumarea unui plan de dezvoltare profesională continuă prin actualizarea permanentă a cunoștințelor

	și abilităților în corelare cu progresele științifice și tehnologice înregistrate în domeniu; • dezvoltarea și obținerea de noi cunoștințe și competențe, atât pe dimensiunea data engineering /ingineria datelor (<i>practica de a proiecta și construi sisteme pentru colectarea, stocarea și analizarea datelor la scară</i>), cât și analitică (<i>modelare statistică, metode de instruire supervizate și nesupervizate, tehnici de învățare automată și profundă</i>) ce permite, viitorilor specialiști, să aplice cunoștințele acumulate, la soluționarea celor mai diverse probleme din economia națională, utilizând simularea proceselor și fenomenelor din domenii de activitate proxime viitorilor specialiști în aplicarea TIC; <u>La nivel de abilități:</u> <u>Tehnologii</u> - Vizualizarea datelor. - Modelarea matematică, informatică, inclusiv și econometrică. - Aplicarea tehnologiilor Știința Datelor / Data Science în modelare predictivă. - Procesarea limbajului natural. - Cloud Computing - Descoperirea cunoștințelor (din Big Data) - Aplicarea ML și DML <u>Management</u> - Crearea și gestionarea activelor informaționale ale unei entități - Gestionarea modernă a datelor unei entități. - Marketing Analytics, bazat pe Big Data. - Risk Analytics, bazat pe Big Data. - Gestionarea procesului decizional bazat pe Data Science și Big Data.
--	--

5. Competențe specifice acumulate

<u>Competențe profesionale</u>	
	CPM1: Elaborarea și proiectarea arhitecturii sistemului CPM2: Monitorizarea tendințelor tehnologice. Inovație. Dezvoltarea durabilă. CPM3: Dezvoltarea aplicațiilor. Integrarea componentelor. Ingineria sistemelor. CPM5: Îmbunătățirea proceselor.
<u>Competențe transversale</u>	
	CTM1: Autonomie și responsabilitate CTM2: Interacțiune socială CPM5: Îmbunătățirea proceselor.

6. Conținutul unității de curs/modulului

Tematica activităților didactice	Numărul de ore
----------------------------------	----------------

Tematica prelegerilor	
TP1. Tema 1: Introducere în Știința și Analiza Datelor: Date. Cele 4 tipuri de date. Managementul datelor. Ecosistemul datelor cu 4 nivele. Cele patru scale de măsurare a datelor, nominal, ordinal, interval și rație. Caracteristicile celor 4 scale de date. Analiza/Analitica Datelor. Două tipuri de Analitică a datelor.	2
TP2. Tema 2: Nivele de analiză a datelor - Analiza descriptivă, Analiza diagnostică, Analiza predictive și Analiza prescriptive. Ce este Știința Datelor/ Data Science – (ȘD/DS). Domenii de aplicabilitate. Tehnologii. Instrumente. Metodologia cercetării. Template-ul cercetării la cursul ȘAD – Lucrare de Totalizare Curs. Despre programul cursului ȘAD. Fundamentarea cursului ȘAD (introducere în curs). Instrumente și tehnologii în ȘD/DS: Ipoteze în ȘD/DS:	2
TP3. Tema 3: Caracterul aplicativ și de cercetare al Științei Datelor: rolul Științei Datelor în TIC; domenii de aplicare a Științei Datelor; perspectivele dezvoltării ȘD. Ciclul de viață al ȘD.	2
TP4. Tema 4: Ciclul de viață al ȘD: formularea problemei; caracteristici / features, observații ale unui fenomen supus cercetării. Statistici. O analiză pas cu pas a ciclului de viață a ȘD: de la înțelegerea afacerii, la monitorizarea modelului.	2
TP5. Tema 5: Modelarea datelor. Tipuri de modele de date. Ce este un flux de date /data flow? Importanța unui flux de date/data flow pentru ȘD. Analiza exploratorie a datelor - EDA - Exploratory Data Analysis (8 pași);.	2
TP6. Tema 6: Analiza exploratorie a datelor - EDA - Exploratory Data Analysis (8 pași);. EDA - prezentare generală, tehnici, instrumente și aplicații. EDA – pas cu pas. Ghid de efectuare a EDA. Studiu de caz.	2
TP7. Tema 7: Tehnologia BD și a depozitelor de date – DWH. DWH + Data Lake sau ce este LakeHouse: Ce este o conductă/ data pipeline? Aree de utilizare. Instrumente si tehnologii pentru conducte de date. Procesul ETL/ELT. Statistica, metrice. Metode de Selectare a caracteristicilor.	2
TP8. Tema 8: Metode de Selectare a caracteristicilor. ML/DML. Instrumente si tehnologii. Tipuri de Analize da Date: selectarea modelului; evaluarea și validarea modelului; implementarea modelului; suportul și monitorizarea modelului.	2
TP9. Tema 9: Optimizarea procesului în ȘD/DS: recomandări și elaborarea Strategiei de comportament privind optimizarea procesului ȘD/DS. Selectarea modelului optim în baza metricilor. Hyperparametri. Tuning.	2
TP10. Tema 10: Hyperparametri. Tuning. Tuning/Reglarea Hyperparametrilor. Tehnici de reglaj/tuning a hyperparametrilor - GridSearchCV, RandomSearchCV și căutare manuală. Selectarea modelului optim în baza hyperparametrilor. Cum să identificăm hyperparametrii optimi pentru modelele noastre?	2
Total prelegeri:	20

7. Conținutul lucrărilor de laborator

Concept.

Lucrările de laborator sunt axate pe realizarea practică a **Obiectivului general al cursului**, - cunoașterea abordărilor științifice Știința Datelor/Data Science, tehnologiilor de lucru cu date brute, Big Data, și dezvoltarea unor capacități și abilități de a le analiza și utiliza în orice domenii de activitate profesională, în particular, pentru suport în procesul decizional al unei entități implicate în menținerea climatului optim al dezvoltării socio-economice continue și durabile.

Reieșind din Obiectivul general al cursului și obiectivele/sarcinile cursului, **scopul lucrărilor de laborator** este formarea de competențe academice și abilități în domeniul analizei și evaluării impactului noilor tehnologii informaționale, inclusiv tehnologiilor și instrumentelor Știința Datelor / Data Science și Big Data și a celor conexe, asupra performanței unui oarecare domeniu de activitate profesională, în particular, pentru suport în procesul decizional al unei entități implicate în menținerea climatului optim al dezvoltării socio-economice continui și durabile, precum și aplicarea practică a acestora în cercetare.

Principalele obiective ale **lucrărilor de laborator** sunt:

1. Transmitere studenților, prin instruire și materiale metodice, a abilităților muncii academice, inclusiv pregătirea și desfășurarea cercetării științifice, cu toate etapele ei pas-la-pas (from scratch);
2. Transmitere studenților, prin instruire și materiale metodice, a abilităților de discuție/dezbatere științifică și de prezentare a ipotezelor/ideilor, conceptelor, rezultatelor cercetării, proiectelor și lucrărilor de cercetare;
3. Transmitere studenților, prin instruire și materiale metodice, instruire în utilizarea instrumentelor și tehnologiilor Știința Datelor / Data Science și Big Data în cercetarea științifică;
4. Transmitere studenților, prin instruire și materiale metodice, instruire în metodele și abilitățile de utilizare a previziunii științifice pentru a determina tendințele tehnologice în domeniul TIC,

Succint, scopul practic al lucrărilor de laborator este: a face din munca științifică a studenților un element permanent și sistematic al procesului de învățământ, de a-i include în viața comunității științifice, de a ajuta la stăpânirea metodologiei, tehnologiei și instrumentelor activităților de cercetare.

Din punct de vedere tematic, lucrările de laborator sunt focusate pe următoarele 5 compartimente:

1. Bazele matematice și tehnologice ale instrumentelor și tehnologiilor Data Science - un studiu al caracteristicilor aplicării instrumentelor și tehnologiilor la sarcinile legate de formarea unei proceduri eficiente, a unei entități din cadrul unui oarecare domeniu de activitate profesională, în baza oportunităților noi, create de aplicarea agregată a Big Data și Știința Datelor/ Data Science, privind procesul de analiză și luare a deciziilor.
2. Bazele matematice și tehnologice ale instrumentelor și tehnologiilor Big Data - un studiu al caracteristicilor aplicării tehnologiilor la sarcinile legate de formarea unei infrastructuri informaționale eficiente, a unei entități din cadrul unui oarecare domeniu de activitate profesională, pentru generarea de noi oportunități de analiză a datelor istorice și curente în procesul de luare a deciziilor.
3. Identificarea unor soluții arhitecturale bazate pe Big Data și Știința Datelor/Data Science pentru întreprindere: prezentarea modului de transformare a sistemului de management al datelor, formarea activelor informaționale ale unei entități din cadrul unui oarecare domeniu de activitate profesională, pentru un sistem de colectare a datelor privind procesele de afaceri, prelucrarea informațiilor externe, design nou de interacțiune, interacțiunea sistemelor.
4. Prezentarea economiei instrumentelor și tehnologiilor Știința Datelor / Data Science și Big Data: evaluarea eficienței economice a soluțiilor pentru managementul unei entități din cadrul unui oarecare domeniu de activitate

profesională, bazate pe tehnologia Big Data, posibilitatea utilizării activelor informaționale ale întreprinderii, posibilitatea utilizării soluțiilor bazate pe informații nestructurate din diverse surse în managementul întreprinderii.

5. Perspectivile de dezvoltare a funcționalității și domeniilor de aplicare a instrumentelor și tehnologiilor Știința Datelor / Data Science și Big Data în cercetarea științifică și impactul ei asupra unui oarecare domeniu de activitate profesională, în particular, pentru suport în procesul decizional al unei entități implicate în menținerea climatului optimal dezvoltării socio-economice continue și durabile.

Tematica activităților didactice	Numărul de ore
Tematica lucrărilor de laborator	
LP1. Tema 1: Crearea unui set de date nu este o performanță ușoară, dar destul de importantă în ȘD. Despre 6 strategii de creare a unui set de date. Ce este Parsing si Scraping. Instrumente si Tehnologii.	4
LP2. Tema 2: Modelarea în ȘD. Modelul si fitting-ul lui în ȘD, sau cum arată un model care se potrivește bine cu obiectul fizic real / procesul cercetat (adică cu sistemul real dorit). Instrumente si Tehnologii. SMOTE (tehnica de supraeșantionare a minorității sintetice) / SMOTE (synthetic minority over-sampling technique).	4
LP3. Tema 3: Data Science Framework. Primul proiect ȘD. 7 pași ai procesului ce determină ȘD. EDA - Instrumente si Tehnologii. Instrumentul Yellowbrick in EDA prin vizualizări.	4
LP4. Tema 4: Metode de selectare a caracteristicilor. Metode supravegheate și nesupravegheate. Sfaturi și trucuri pentru selectarea caracteristicilor. Suprafitting si underfitting.	4
LP5. Tema 5: Validarea modelului și evaluarea calității lui. Acuratețea predictivă. Parametru si hiperparametru. Reglarea hiperparametrului și tehnicile sale. Tehnici de optimizare a hiperparametrilor - GridSearch CV, RandomSearch CV și Manual. Pași pentru efectuarea reglării hiperparametrului.	4
Total lucrări de laborator:	20

8. Lucrare de Totalizare Curs

Alegerea temei lucrării de totalizare curs se realizează în conformitate cu tema lucrării de laborator Nr.1, oferită de lector, sau se realizează conform temei propuse de student de comun acord cu profesorul.

Subiectele aproximative ale Lucrărilor de Totalizare Curs, ținând cont de utilizarea conceptului, instrumentelor/metodelor și tehnologiilor Știința Datelor / Data Science și Big Data, sunt formate în următoarele domenii:

- Analiza activelor informaționale ale unei entități economice (pe industrii concrete).
- Modelarea informațională a arhitecturii și proceselor de afaceri ale unei entități economice.
- Metode și tehnologii pentru prelucrarea unor volume mari de date eterogene.
- Aplicarea tehnologiilor Știința Datelor / Data Science și Big Data în managementul entităților economice.
- Aspecte economice ale utilizării Știința Datelor / Data Science și Big Data.

- SIAD-uri în managementul întreprinderilor, bazate pe utilizarea Știința Datelor / Data Science și Big Data.
- Alte domenii
 - Învățământ
 - Domeniul Bancar
 - Medicină
 - Psihologie
 - etc

Temele Lucrării de Totalizare Curs sunt considerate o etapă importantă în procesul de pregătire a tezei finale de master și, de regulă, pregătește studentul pentru dezvoltarea acesteia.

9. Referințe bibliografice

Principale	1. Миркин Б. Г. Введение в анализ данных. Учебники и практикум[Электронный ресурс] : М.:Издательство Юрайт , 2019 - 174 - Режим доступа: https://biblioonline.ru/book/vvedenie-v-analiz-dannyh-432851 2. Peters Morgan, Data analysis from scratch with Python. Step by step guide © Copyright 2016 by AI Sciences LLC, All rights reserved. First Printing, 2016, Edited by Davies Company Ebook Converted and Cover by Pixels Studio Published by AI Sciences LLC, ISBN-13: 978-1721942817, ISBN-10: 1721942815 3. John Wiley & Sons, Data Science & Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data, 10475 Crosspoint Boulevard, Indianapolis, IN 46256, www.wiley.com Copyright © 2015 by John Wiley & Sons, Inc., Indianapolis, Indiana, Published simultaneously in Canada ISBN: 978-1-118-87613-8, ISBN: 978-1-118-87622-0 (ebk) ISBN: 978-1-118-87605-3 (ebk) 4. Building Big Data and Analytics Solutions in the Cloud International Technical Support Organization December 2014 5. Wes McKinney, Python for Data Analysis, Copyright © 2013 Wes McKinney. All rights reserved. Printed in the United States of America. Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472. 6. Michael Z. Zgurovsky Yuriy P. Zaychenko, Big Data: Conceptual Analysis and Applications, SSN 2197-6503 ISSN 2197-6511 (electronic) Studies in Big Data ISBN 978-3-030-14297-1 ISBN 978-3-030-14298-8 (eBook) https://doi.org/10.1007/978-3-030-14298-8 Library of Congress Control Number: 2019933181 © Springer Nature Switzerland AG 2020 7. Zeng Deze, Huan Huang, Rui Hou, Seungmin Rho, Naveen Chilamkurti, Big Data Technologies and Applications, 10th EAI International Conference, BDTA 2020 and 13th EAI International Conference on Wireless Internet, WiCON 2020 Virtual Event, December 11, 2020. ISSN 1867-8211 ISSN 1867-822X (electronic) Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering ISBN 978-3-030-72801-4 ISBN 978-3-030-72802-1 (eBook) https://doi.org/10.1007/978-3-030-72802-1 8. Sherif Sakr, Albert Y. Zomaya, Encyclopedia of Big Data Technologies, ISBN 978-3-319-77524-1 ISBN 978-3-319-77525-8 (eBook) ISBN 978-3-319-77526-5
------------	---

(print and electronic bundle), Library of Congress Control Number: 2018960889
© Springer Nature Switzerland AG 2019 <https://doi.org/10.1007/978-3-319-77525-8>

9. **S. Ejaz Ahmed**, Big and Complex Data Analysis Methodologies and Applications, SSN 1431-1968 Contributions to Statistics ISBN 978-3-319-41572-7 ISBN 978-3-319-41573-4 (eBook) DOI 10.1007/978-3-319-41573-4 Library of Congress Control Number: 2017930198 © Springer International Publishing AG 2017
10. **Big Data Analytics with Spark**, Copyright © 2015 by Mohammed Guller, ISBN-13 (pbk): 978-1-4842-0965-3, ISBN-13 (electronic): 978-1-4842-0964-6, www.apress.com/source-code/, OR www.it-ebooks.info
11. **Thomas Erl, Wajid Khattak and Paul Buhler, Big Data Fundamentals, Concepts, Drivers & Techniques**, Copyright © 2016 Arcitura Education Inc., ISBN-13: 978-0-13-429107-9
12. **Raul Estrada Isaac Ruiz, Big Data SMACK: A Guide to Apache Spark, Mesos, Akka, Cassandra, and Kafka**, Mexico City Mexico City, ISBN-13 (pbk): 978-1-4842-2174-7 ISBN-13 (electronic): 978-1-4842-2175-4, DOI 10.1007/978-1-4842-2175-4, Library of Congress Control Number: 2016954634, Copyright © 2016 by Raul Estrada and Isaac Ruiz
13. **Big Data Principles and Paradigms, Edited by Rajkumar Buyya The University of Melbourne and Manjrasoft Pty Ltd, Australia Rodrigo N. Calheiros** The University of Melbourne, Australia Amir Vahid Dastjerdi The University of Melbourne, Australia, Morgan Kaufmann is an imprint of Elsevier 50 Hampshire Street, 5th Floor, Cambridge, MA 02139, USA Copyright © 2016 Elsevier Inc. All rights reserved. 496 pag
14. **10 Best Big Data Books in 2023 [Learn Big Data ASAP]**
<https://azure.microsoft.com/en-us/products>
<https://realtoughcandy.com/best-big-data-books/>
15. **20 Best Big Data Books for Beginners**
<https://bookauthority.org/books/beginner-big-data-books>
16. **Must Read Books for Beginners on Big Data, Hadoop and Apache Spark**
<https://www.analyticsvidhya.com/blog/2015/10/books-big-data-hadoop-apache-spark/>
<https://dokumen.pub/big-data-concepts-technology-and-architecture-9781119701828.html>
17. **Microsoft Azure for students**
<https://engineering.buffalo.edu/computer-science-engineering/information-for-students/information-technology/software-distribution/microsoft-azure-for-students.html>
<https://learn.microsoft.com/en-us/azure/education-hub/azure-dev-tools-teaching/azure-students-program>
<https://azure.microsoft.com/en-us/resources/students?activetab=pivot:githubtab>
<https://learn.microsoft.com/en-us/azure/education-hub/azure-dev-tools-teaching/program-faq>
<https://www.techopedia.com/definition/26434/azure-service-platform>
18. **Big Data Analytics - Statistical Methods.**
https://www.tutorialspoint.com/big_data_analytics/statistical_methods.htm
19. **Майер-Шенбергер, В. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим / Виктор Майер-Шенбергер,**

	Кеннет Кукьер ; пер. с англ. Инны Гайдюк. — М. : Манн, Иванов и Фербер, 2014. — 240 20. Гусева Е. Н. Теория вероятностей и математическая статистика : Учебники и учебные пособия для ВУЗов [Электронный ресурс] - Москва : ФЛИНТА , 2016 - 220 – Режим доступа: http://biblioclub.ru/index.php?page=book_red&id=83543 21. Гутова С. Г., Алтемерова О. А. Теория вероятностей и математическая статистика: Учебники и учебные пособия для ВУЗов [Электронный ресурс] - Кемерово : Кемеровский государственный университет , 2016 - 216 - Режим доступа: http://biblioclub.ru/index.php?page=book_red&id=481538 22. Кочетков Евгений Семенович. Теория вероятностей и математическая статистика: Учебник [Электронный ресурс] : Форум , 2018 - 240 - Режим доступа: http://znanium.com/go.php?id=944923
Supli-mentare	1. Белько Иван Васильевич. Теория вероятностей, математическая статистика, математическое программирование : Учебное пособие [Электронный ресурс] , 2016 - 299 -Режим доступа: http://znanium.com/go.php?id=542521 2. Волкова Полина Андреевна. Статистическая обработка данных в учебно-исследовательских работах : Учебное пособие [Электронный ресурс] : Форум , 2019 - 96 - Режим доступа: http://znanium.com/go.php?id=1030246 3. Непомнящая Наталья Васильевна. Статистика: общая теория статистики, экономическая статистика. Практикум : Учебное пособие [Электронный ресурс] , 2015 - 376 -Режим доступа: http://znanium.com/go.php?id=549841 4. Основы теории вероятностей и математической статистики [Электронный ресурс] , 2018 - 90 - Режим доступа: https://lib.rucont.ru/efd/684477

10. Evaluare

Periodică		Curentă	Studiu individual	Proiect/teză	Examen
EP 1	EP 2				
15%	15%	15%	15%	-	40%
Standard minim de performanță - Prezența și activitatea la prelegeri și lucrările laborator; - Obținerea notei minime de „5” la fiecare dintre testări/atestări și lucrări practice individuale sau de grup; - Demonstrarea în lucrarea de examinare finală a cunoștințelor teoretice, competențelor și abilităților practice necesare pentru utilizarea celor învățate la cursul Știința și Analiza Datelor în cadrul dezvoltării unui Proiect concret, privind Lucrarea de Totalizare Curs.					

11. Criterii de evaluare

Activitate	Componente evaluare	Metodă de evaluare, Criterii de evaluare	Pondere în nota finală a activității	Ponderea în evaluarea discipline i
Învățământ cu frecvență				
Evaluare periodică I	Conținut teoretic, teme 1-3	Test pe MOODLE	100%	15%
Evaluare periodică II	Conținut teoretic, teme 4-5	Test pe MOODLE	100%	15%
Evaluare curentă	Activitatea practică	Discuții în cadrul seminarelor Test pe MOODLE	50%	15%
Dosar completat cu Rapoarte pentru fiecare Studiu de caz în discuție			50%	15%
Studiul individual	Cercetare la temă. Lucrare de Totalizare curs.	Prezentare/discurs public	100%	
Evaluarea finală	Conținut teoretic și practic	Examen oral. Prezentare Lucrare de Totalizare curs. Notare conform baremului prestabilit și prezentat studenților la începutul cursului în compartimentul "Reguli de joc".	100%	40%

12 Utilizarea IA generativă

Permisivitatea de utilizare	Utilizarea IA generative în cadrul temelor și proiectelor este permisă, cu condiția ca studenții să respecte următoarele reguli: - IA generativă poate fi utilizată pentru generarea de idei, structuri de text sau cod, dar toate materialele generate trebuie să fie revizuite și ajustate de către student pentru a se asigura că acestea corespund cerințelor academice. - Orice utilizare a IA generative trebuie să fie declarată în secțiunea de apendice a fiecărei lucrări, folosind fraza: "În timpul pregătirii acestei lucrări, autorul a utilizat [NUME INSTRUMENT / SERVICIU] în scopul [MOTIV]. După utilizarea acestui instrument/serviciu, autorul a revizuit și editat conținutul după cum a fost necesar și își asumă întreaga responsabilitate pentru conținutul lucrării."
Restricții de utilizare	Studenții nu trebuie să considere IA generativă ca o sursă de încredere pentru informații, deoarece nu oferă referințe clare sau surse documentate. Nu este permisă citarea directă a conținutului generat de IA în lucrările academice ca și cum ar fi sursă primară.

	• Activitățile în care este interzis utilizarea IA generativă sunt specificare de profesor și sunt de regulă evaluări intermediare și finale sau care nu presupun activități de dezvoltare a competențelor profesionale.
--	--