

BIG DATA

1. Date despre unitatea de curs/modul

| | | | | | |
|--------------------------------|---|------------------------|-------------------------------------|-----------------------------------|---------------------|
| Facultatea | Calculatoare, Informatică și Microelectronică | | | | |
| Catedra/departamentul | Informatică și Ingineria Sistemelor | | | | |
| Ciclul de studii | Studii superioare de licență, Ciclul I | | | | |
| Programul de studiu | 0613.5 Informatica Aplicată | | | | |
| Anul de studiu | Semestrul | Tip de evaluare | Categoria formativă | Categoria de opționalitate | Credite ECTS |
| III (învățământ cu frecvență); | V | E | S – unitate de curs de specialitate | A - unitate de curs la alegere | 4 |

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 |
| 120 | 30 | 30 | | 30 | 30 |

3. Precondiții de acces la unitatea de curs/modul

| | |
|--------------------------------|--|
| Conform planului de învățământ | Studiul se bazează pe următoarele discipline: Matematică discretă, Algebră, Teorie probabilități și statistici matematice, Metode de programare, Metode de analiză a datelor. Adicional, studenții trebuie să cunoască, structura calculatorului, SO Windows 10, limbajele de programare C#, C++, HTML, PHP, Python, utilizarea unui redactor de texte. Cunoștințe avansate despre sisteme moderne de Gestionare a Bazelor de Date, cum ar fi: MySQL, MS SQL Server 2019, CMS-uri Wordpress, framework-uri Laravel, Code Igniter, Symphony ș.a., abilități de a crea aplicații DeskTop sau Web, Front End și Back End, utilizând Baze de Date Operaționale, precum și modelarea multidimensională a datelor, Depozitele de date, tehnologia OLAP, Instrumente BI, (VS, Power BI), tehnologia Data Mining, lucrul cu tehnologia cloud, platformele Azure, AWS (Amazon Web Service) și Google Cloud. |
| Conform competențelor | Studenții trebuie să aibă următoarele cunoștințe și competențe pentru a stăpâni disciplina academică "Big Data": capacitatea de a utiliza corect TIC în rezolvarea problemelor profesionale, cu caracter aplicativ, folosind abilitățile de modelare matematică ale fenomenelor social-economice, ale mediului ambiat, din domeniul tehnic, precum și din domeniile: financiar-bancar, învățământ, medicină, asigurări, marketing etc, abilități formate la disciplinele menționate mai sus. Rezultatele principale ale disciplinei "Big Data" pot fi utilizate în viitor la studierea disciplinelor: "Analiza datelor", "Metode de protecție a informațiilor", "Analitica Big Data", "Informatica în afaceri", "Procesarea Big Data", "Analiza Big data", "Tehnologii Big Data", |

4. Condiții de desfășurare a procesului educațional pentru

| | |
|----------------------------|---|
| Curs | Pentru prezentarea materialului teoretic sala de curs trebuie să dispună de un proiector și calculator. Nu vor fi tolerate întârzierile studenților, precum și convorbirile telefonice în timpul cursului. |
| Laborator/seminar/practice | Studenții vor perfectă rapoarte conform condițiilor prevăzute de indicațiile metodice. Termenul de predare a lucrării de laborator/practice, individual sau de grup, se prezintă la următoarea ședință de lucru, în format electronic. Pentru predarea cu întârziere a lucrului efectuat, aceasta se depuncea cu 1pct./săptămână de întârziere. |

5. Competențe specifice acumulate

| | |
|-------------------------|--|
| Competențe profesionale | CPL 1. Proiectarea aplicațiilor (A6)** K1 Tehnici de modelare a cerințelor și tehnici de analiză a nevoilor. K2 Metodele de dezvoltare a software-ului și argumentarea acestora (de exemplu, prototipuri, metode agile, retroinginerie etc.). K3 Metricile care se referă la dezvoltarea aplicațiilor. K4 Principiile de proiectare a interfeței pentru utilizator. K5 Limbajele pentru formalizarea specificațiilor funcționale. K6 Aplicațiile existente și arhitectura lor aferentă. K7 Sisteme de gestionare a bazelor de date (DBMS), depozite de date, informații de business etc. |
|-------------------------|--|

| | |
|-------------------------|--|
| | <p>K8 Tehnologiile mobile</p> <p>CPL 2. Proiectarea și dezvoltarea aplicațiilor (B.1)</p> <p>K1 Programe/module software adecvate.</p> <p>K2 Componente hardware, instrumente și arhitecturi hardware.</p> <p>K3 Proiectarea funcțională și tehnică.</p> <p>K4 Tehnologiile de ultimă oră.</p> <p>K5 Limbaje de programare.</p> <p>K6 Baze de date (DBMS).</p> <p>K7 Sisteme de operare și platforme software.</p> <p>K8 Mediul de dezvoltare integrat (IDE - integrated development environment).</p> <p>K9 Dezvoltarea rapidă a aplicațiilor.</p> <p>K10 Problemele legate de drepturile de proprietate intelectuală (IPR).</p> <p>K11 Tehnologia de modelare tehnică și limbaje.</p> <p>K12 Limbajele de definire a interfeței (IDL).</p> <p>K13 Probleme de securitate.</p> |
| Competențe transversale | <p>CTL1 Autonomie și responsabilitate Demonstrează executarea responsabilă a sarcinilor profesionale în condiții de autonomie.</p> <p>CTL2 Interacțiune socială Execută rolurile și activitățile specifice muncii în echipă și distribuie sarcinile între membri pe niveluri subordonate.</p> <p>CTL3 Dezvoltare personală și profesional Conștientizează nevoia de formare continuă cu utilizarea eficientă a resurselor și tehnicilor de învățare pentru dezvoltarea personală și profesională</p> |

6. Obiectivele unității de curs/modulului

| | |
|-----------------------|--|
| Obiectivul general | <p>familiarizarea, inițierea și instruirea studenților pentru a înțelege și utiliza în practică care este esența și specificul cursului “Big Data”, cum se poziționează în acest proces utilizarea practică a cunoștințelor și abilităților obținute la cursurile precedente ce l-au precedat, Baze de Date și Învățarea Automată Procesarea Informației, cum se aplică pentru managementul Big Data, Bazele de Date Operaționale - BDO, Depozitele de Date - DWH și sistemele moderne de gestiune a lor, precum și să dezvolte deprinderi, abilități practice și să asigure dobândirea de competențe, atât pe dimensiunea analitică (modelare statistică, metode predictive, tehnici de învățare automată) cât și pe cea tehnologică (limbaje de programare, arhitecturi și platforme). ce le-ar permite studenților, viitorii specialiști, să aplice cunoștințele acumulate, de-a lungul cursurilor menționate, la 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 - 0613.5 Informatica Aplicată</p> |
| Obiectivele specifice | <p><i>La nivel de cunoștințe:</i></p> <ul style="list-style-type: none"> • Capacitatea de a opera cu concepte fundamentale din domeniul modelării matematice și analizei statistice precum și abilitatea de a le folosi în contexte practice • Capacitatea de a identifica, implementa și utiliza algoritmi de extragere a modelelor din date folosind metode statistice și tehnici de învățare automată. • Capacitatea de a înțelege și a aplica principiile procesării distribuite a datelor și a utilizării arhitecturilor de calcul de înaltă performanță. • Abilitatea de a utiliza platforme și tehnologii specifice prelucrării unor volume mari de date și de a implementa aplicații scalabile. • Capacitatea de a adapta soluții bazate pe abordări de tip data-driven pentru probleme specifice unui anumit domeniu de aplicabilitate. • Capacitatea de a planifica și organiza activitatea de cercetare într-un domeniu prin organizarea lucrului în grup; înțelegerea responsabilităților și respectarea regulilor de etică profesională și a celor de securitate și confidențialitate a datelor. • 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 <p><i>La nivel de abilități:</i></p> <ul style="list-style-type: none"> • să elaboreze strategii de date Big Data, reieșind din contextul entității ce le utilizează, scopul și obiectivele formulate pentru dezvoltarea ei și generarea de valoare adugată; |

| | |
|--|---|
| | <ul style="list-style-type: none"> • să dezvolte fluxuri și conducte de date necesare procesului de management al “Big Data”, pentru facilitarea activităților de management informational necesar procesului decizional al entității; • să analizeze utilizarea “Big Data” din punctul de vedere al eficienței, utilizării datelor, atunci când este vorba de un Proiect informațional concret, ce ține de procesul decizional al unei entități, suport al caruia “Big Data” se prezuma a fi; • să utilizeze Ip. Python, platformele Power BI, VS, tehnologii cloud și platforme cloud cum ar fi: AWS, Azure, Google Cloud, pentru Analiza și Analitica Big Data. |
|--|---|

7. Conținutul unității de curs/modulului

| Tematica activităților didactice | Numărul de ore | |
|--|-------------------------|--------------------------------|
| | învățământ cu frecvență | învățământ cu frecvență redusă |
| Tematica prelegerilor | | |
| <u>TP1</u> Introducere în curs. Reguli. Noțiuni. Definiții. Caracteristici. | 2 | |
| <u>TP2</u> Rolul Big Data în procesul decizional. Strategie de date Big Data. Big Data în SSD al unei entități de afaceri/publice. | 2 | |
| <u>TP3</u> Big Data. Tipuri de date în Big Data. Date structurate, semi-structurate, cuazi-structurate și nestructurate. | 2 | |
| <u>TP4</u> Big Data, Data Engineering. Big Data Engineer. | 2 | |
| <u>TP5</u> Arhitectura Big Data – viziune generală și detaliată. | 2 | |
| <u>TP6</u> Big Data. Lacuri de date. | 2 | |
| <u>TP7</u> Tehnologii Big Data. Instrumente | 2 | |
| <u>TP8</u> Platforma Azure Databricks. Delta Lake. | 2 | |
| <u>TP9</u> Platforma Azure Databricks Lakehouse. | 2 | |
| <u>TP10</u> Azure Databricks Lakehouse ETL | 2 | |
| <u>TP11</u> Aplicații Big Data. Convertirea datelor nestructurate în date de lucru. Depozite de date în Big Data, Instrumente de gestiune, NOSQL. | 2 | |
| <u>TP12</u> Depozite de date în Big Data, Instrumente de gestiune, NOSQL. Teorema CAP. MongoDB. Atlas MongoDB. | 2 | |
| <u>TP13</u> Hadoop și MongoDB in Big Data. Analitica Big Data | 2 | |
| <u>TP14</u> Analitica Big Data și DSS. Instrumente. | 2 | |
| <u>TP15</u> BI, Analitica Big Data și Știința Datelor. | 2 | |
| Total prelegeri: | 30 | |

| Tematica activităților didactice | Numărul de ore | |
|--|-------------------------|--------------------------------|
| | învățământ cu frecvență | învățământ cu frecvență redusă |
| Tematica lucrărilor practice | | |
| <u>LP1.</u> Big Data în SSD al unei entități de afaceri/publice. Elaborare Strategie de date Big Data. Studiu de caz. | 2 | |
| <u>LP2.</u> Statistici. Python. Aplicații în Big Data. Importanța statisticilor pentru Big Data. Studiu de caz/Exemple. | 2 | |
| <u>LP3.</u> Modelare Big Data. Proiect. Elemente de bază. Structură. Studiu de caz. | 2 | |
| <u>LP4.</u> Flux de date. Conductă de date. Studiu de caz. | 2 | |
| <u>LP5.</u> Big Data. Proiect. EDA. Studiu de caz/Exemple. | 2 | |
| <u>LP6.</u> Maparea datelor. Convertirea datelor nestructurate în date de lucru. Instrumente. Studiu de caz/Exemplu. | 2 | |
| <u>LP7.</u> Conducta de date SSIS/ETL pentru MS SQL SERVER și VS. Studiu de Caz/Exemplu. | 2 | |
| <u>LP8.</u> Parsing-ul și Scraping-ul datelor. Studii de caz/Exemple. | 2 | |
| <u>LP9.</u> Construirea conducte ETL cu Azure Databricks și Delta Lake. Studiu de Caz/Exemplu. | 2 | |
| <u>LP10.</u> Procesarea datelor cu ajutorul Platformei Azure Databricks. Studiu de Caz/Exemplu. | 2 | |
| <u>LP11.</u> Convertirea datelor nestructurate în date de lucru. Depozite de date în Big Data, Instrumente de gestiune. NOSQL. MongoDB. Atlas MongoDB | 2 | |
| <u>LP12.</u> Depozite de date în Data Data. Instrumente de gestiune. NOSQL. MongoDB. Atlas MongoDB. Studii de caz/Exemple. | 2 | |
| <u>LP13.</u> MongoDB in Big Data. Analitica Big Data. Studiu de caz/Exemplu. | 2 | |

| | | |
|--|----------|-----------|
| LP14. Analitica Big Data si DSS. Instrumente. Exemple. | 2 | |
| LP15. BI, Analitica Big Data și Știința Datelor. Studiu de caz/Exemple. | 2 | |
| Total lucrări de laborator: | | 30 |

8. Referințe bibliografice

| | |
|------------|--|
| Principale | <ol style="list-style-type: none"> 1. 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 2. 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 3. 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 4. 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. 5. 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) 6. 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 7. Building Big Data and Analytics Solutions in the Cloud International Technical Support Organization December 2014 8. 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 9. 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/ 10. 20 Best Big Data Books for Beginners https://bookauthority.org/books/beginner-big-data-books 11. 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 12. 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 13. Big Data Analytics - Statistical Methods. https://www.tutorialspoint.com/big_data_analytics/statistical_methods.htm 14. 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 15. 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 16. 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 17. Кочетков Евгений Семенович. Теория вероятностей и математическая статистика: Учебник [Электронный ресурс] : Форум , 2018 - 240 - Режим доступа: http://znanium.com/go.php?id=944923 18. Миркин Б. Г. Введение в анализ данных. Учебник и практикум [Электронный ресурс] : М.:Издательство Юрайт , 2019 - 174 - Режим доступа: https://biblioonline.ru/book/vvedenie-v-analiz-dannyh-432851 |
|------------|--|

| | |
|-------------------|--|
| Supli- mentare | <ol style="list-style-type: none"> 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 |
|-------------------|--|

9. Utilizarea IA generativă

| | |
|------------------------------------|--|
| Permisivitatea de utilizare | <p>Utilizarea IA generative în cadrul temelor și proiectelor este permisă, cu condiția ca studenții să respecte următoarele reguli:</p> <ul style="list-style-type: none"> • 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 appendice 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 | <p>Studenții nu trebuie să considere IA generativă ca o sursă de încredere pentru informații, deoarece nu oferă referințe clare sau surse documentate.</p> <ul style="list-style-type: none"> • 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. |

10. Evaluare

| Periodică | | Curentă | Studiu individual | Proiect/teză | Examen |
|---|------|---------|-------------------|--------------|--------|
| EP 1 | EP 2 | | | | |
| 15% | 15% | 15% | 15% | - | 40% |
| <p>Standard minim de performanță Prezența și activitatea la prelegeri și lucrările practice; Obținerea notei minime de „5” la fiecare dintre atestări și lucrări practice individuale sau de grup; Demonstrarea în lucrarea de examinare finală a cunoștințelor teoretice și competențelor practice necesare pentru utilizarea Big Data în cadrul dezvoltării unui Proiect concret.</p> | | | | | |