



FIȘA UNITĂȚII DE CURS/MODULULUI

1. Date despre program

1.1. Instituția de învățământ superior	Universitatea de Stat din Moldova
1.2. Facultatea	Matematică și Informatică
1.3. Departamentul	Informatică
1.4. Domeniul de formare profesională	0613 Dezvoltarea produselor program și a aplicațiilor
1.5. Ciclul de studii	Studii superioare de master, ciclul II
1.6. Programul de studii	Inteligența Artificială și Știința Datelor

2. Date despre unitatea de curs/modul

2.1. Denumirea disciplinei	Procesarea BIG DATA						
2.2. Titularul activităților de curs	?						
2.3. Titularul activităților de laborator	?						
2.4. Anul de studii	2	2.5. Semestrul	1	2.6. Tipul de evaluare	E	2.7. Categoria formativă	SA

3. Timpul total estimat (ore pe semestru al activităților didactice)

3.1. Număr de ore pe săptămână	4	din care: 3.2 curs	2	3.3. laborator	2
3.4. Total ore în planul de învățământ	40	din care: 3.5 curs	20	3.6. laborator	20
<b>Distribuția fondului de timp pentru studiul individual:</b>					<b>ore</b>
Studiu după manual, suport de curs, bibliografie și notițe					40
Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate					35
Pregătire laboratoare, teme, referate, portofolii					35
3.7. Total ore studiu individual	110				
3.8. Total ore pe semestru	150				
3.9. Număr de credite	5				

4. Precondiții de acces la unitatea de curs/modul

4.1. de curriculum	Baze de date, programare(Python), statistică pentru știința datelor, analiza și vizualizarea datelor, învățarea automată, depozite de date și data mining.
4.2. de competențe	Cunoștințe de programare, statistică și învățare automată.

5. Condiții de desfășurare

5.1. a cursului	Sală de curs cu tablă mare și proiector, laptop(cu aplicații instalate – Big Data).
5.2. a laboratorului	Sală de laborator, calculatoare care au instalate aplicații Big Data.

6. Obiectivele disciplinei – rezultate așteptate ale învățării la formarea cărora contribuie parcurgerea și promovarea disciplinei

Cunoștințe	(C1) Dobândirea unor cunoștințe profunde despre diverse tehnologii și instrumente utilizate în stocarea, procesarea, analiza și vizualizarea datelor mari, inclusiv ecosistemele Spark, bazele de date NoSQL, și instrumentele de vizualizare cum ar fi Tableau și PowerBI. (C2) Dobândirea unor cunoștințe fundamentale cu privire la proiectarea diferitor arhitecturi Big Data
------------	--



	<p>scalabile și sigure, înțelegând diferite modele arhitecturale, cum ar fi Lambda și Kappa.</p> <p>(C3) Familiarizarea cu tehnici analitice avansate și despre aplicarea algoritmilor de învățare automată în contexte de Big Data.</p>
Abilități	<p>(A1) Abilitatea în utilizarea instrumentelor și tehnologiilor de procesare a datelor mari, cum ar fi MapReduce, Apache Kafka, și Apache Spark.</p> <p>(A2) Abilitatea de a crea vizualizări eficiente și dashboard-uri interactive pentru a comunica eficient rezultatele analitice.</p> <p>(A3) Abilitatea în a aplica tehnici analitice pentru a extrage informații valoroase din seturi mari de date și să interpreteze rezultatele..</p> <p>(A4) Abilitatea de a proiecta și implementa soluții scalabile de Big Data, care să poată gestiona creșterea volumului și diversității datelor.</p>
Responsabilitate și autonomie	<p>(R1) Capacitatea de a implementa și gestiona soluții de Big Data în mod autonom, asigurând scalabilitatea și securitatea acestora.</p> <p>(R2) Capacitatea de a contribui la inovarea și cercetarea în domeniul Big Data, propunând soluții noi și adaptând tehnologiile existente la cerințele specifice ale industriei..</p>

## 7. Conținutul unității de curs/modulului

7.1. Curs	Numărul de ore
<i>C1. Introducere în Big Data.</i> Ce este Big Data și de ce este important.Cele 5 V-uri ale Big Data: Volum, Viteză, Varietate, Veridicitate, Valoare.Prezentare generală a ecosistemului Big Data.	2
<i>C2. Tehnologii de Stocare Big Data.</i> Sisteme de fișiere distribuite (HDFS).Baze de date NoSQL: MongoDB, Cassandra.Data Lakes vs. Data Warehouses.	2
<i>C3. Prelucrarea Big Data.</i> Introducere în MapReduce.Procesarea datelor în flux: Apache Kafka, Apache Storm. Apache Spark și PySpark.	2
<i>C4. Big Data Analytics.</i> Tehnici analitice pentru Big Data.Instrumente analitice: Apache Hive, Apache Pig.Introducere în machine learning cu Big Data.	2
<i>C5. Vizualizarea Big Data.</i> Principii de bază în vizualizarea Big Data.Instrumente de vizualizare: Tableau, PowerBI, Qlik.Crearea dashboard-urilor eficiente pentru Big Data.	2
<i>C6.Arhitectura Big Data.</i> Proiectarea unei arhitecturi Big Data scalabile.Considerații privind securitatea și confidențialitatea datelor.Înțelegerea arhitecturilor Lambda și Kappa.	2
<i>C7. Big Data în Cloud.</i> Servicii cloud pentru Big Data: AWS, Azure, Google Cloud.Avantajele și provocările prelucrării Big Data în cloud.Casus de utilizare Big Data în cloud.	2



<i>C8. Machine Learning și Big Data.</i> Aplicații ale învățării automate în Big Data.Algoritmi scalabili de învățare automată.Introducere în învățarea profundă pentru Big Data.	2
<i>C9. Securitatea și Confidențialitatea în Big Data.</i> Provocări ale securității datelor la scară largă.Reglementări și conformitate: GDPR, HIPAA.Tehnici de anonimizare și criptare.	2
<i>C10. Managementul Proiectelor Big Data.</i> Planificarea și gestionarea proiectelor Big Data.Metodologii agile și lean în proiectele Big Data.Monitorizarea și evaluarea performanței proiectelor.	1
<i>C11. Big Data în Industrii Specifice.</i> Aplicații Big Data în sănătate, finanțe, retail, și altele.Studii de caz și cele mai bune practici.Provocări și soluții specifice industriei.	1
<i>Total prelegeri:</i>	20
<b>Bibliografie:</b>	
<ol style="list-style-type: none"> <li>1. Thomas Erl, Wajid Khattak and Paul Buhler, Big Data Fundamentals, Concepts, Drivers &amp; Techniques, Copyright © 2016 Arcitura Education Inc.</li> <li>2. Big Data Analytics with Spark, Copyright © 2015 by Mohammed Guller.</li> <li>3. Marz, N., &amp; Warren, J. (2015). Big Data. Principles and Best Practices of scalable real-time systems. Manning Publications.</li> <li>4. Cielen, D., Meysman, A.D.B., &amp; Ali, M. (2016). Introducing Data Science. Big Data, machine learning, and more, using Python tools. Manning Publications.</li> <li>5. Damji, J.S., Wenig, B., Das, T., &amp; Lee, D. (2020). Learning Spark. O'Reilly Media, Inc.</li> <li>6. McCallum, Q. E. (2012). Bad Data Handbook: Cleaning Up The Data So You Can Get Back To Work. O'Reilly.</li> </ol>	
<b>7.2. Laborator</b>	
	<b>Numărul de ore</b>
L1. Manipularea datelor mari cu Python și Spark.	4
L2. Prelucrarea Big Data cu Apache Spark.	3
L3. Analiza și vizualizarea Big Data în Spark.	3
L4. Implementarea algoritmilor de machine learning în Spark.	4
L5. Optimizarea performanței în prelucrarea Big Data.	2
L6. Integrarea cu servicii cloud pentru stocarea Big Data.	4
<i>Total laborator:</i>	20
<b>Bibliografie:</b>	
<ol style="list-style-type: none"> <li>1. Thomas Erl, Wajid Khattak and Paul Buhler, Big Data Fundamentals, Concepts, Drivers &amp; Techniques, Copyright © 2016 Arcitura Education Inc.</li> <li>2. Big Data Analytics with Spark, Copyright © 2015 by Mohammed Guller.</li> <li>3. Marz, N., &amp; Warren, J. (2015). Big Data. Principles and Best Practices of scalable real-time systems. Manning Publications.</li> <li>4. Cielen, D., Meysman, A.D.B., &amp; Ali, M. (2016). Introducing Data Science. Big Data, machine learning, and more, using Python tools. Manning Publications.</li> <li>5. Damji, J.S., Wenig, B., Das, T., &amp; Lee, D. (2020). Learning Spark. O'Reilly Media, Inc.</li> <li>6. McCallum, Q. E. (2012). Bad Data Handbook: Cleaning Up The Data So You Can Get Back To Work. O'Reilly.</li> </ol>	

**8. Evaluare**

Nota semestrială
------------------



# Universitatea de Stat din Moldova



Atestarea 1	Atestarea 2	Evaluare curentă	Lucrul individual
25%	25%	25%	25%

Nota finală	
Nota semestrială	Nota la examen
60%	40%

<p style="text-align: center;"><b>Standard minim de performanță</b></p> <p>Standard minim (cunoștințe și aptitudini necesare pentru nota 5)</p> <ul style="list-style-type: none"><li>• prezența la cursuri și seminarii conform cerințelor generale ale facultății</li><li>• cunoașterea conceptelor de bază din Big Data</li><li>• abilitatea de a manipula datelor mari</li><li>• capacitatea de a implementa analiza și vizualizarea datelor utilizând Spark necesare rezolvării unei probleme reale</li></ul> <p>Evaluarea în cadrul disciplinei „Procesarea BIG DATA” se realizează formativ - activitatea la orele de laborator, verificarea orală sau scrisă curentă, două lucrări de atestare (la mijloc și la sfârșit de semestru), și sumativ (final) – 4 lucrări independente, realizate pe parcursul semestrului și examenul final la disciplină. Nota generală la disciplină însumează nota de la examen și cea semestrială, în proporție de 40 și, respectiv, 60 la sută, și apreciază gradul de corespundere cu finalitățile scontate: cunoștințele și competențele acumulate, abilitatea de a aplica cunoștințele, gradul de integrare a cunoștințelor de către studenți etc.</p>
--

Data completării  
25.08.2024

Titular de disciplină  
Dr., conf. univ. ?

Data avizării în departament

Șef departament