

**Силабус навчальної дисципліни  
«Високопродуктивні системи обробки великих даних»**

<b>Спеціальність</b>	<i>122 Комп'ютерні науки</i>
<b>Освітня програма</b>	<i>122 Комп'ютерні науки</i>
<b>Освітній рівень</b>	<i>магістр</i>
<b>Статус дисципліни</b>	<i>вибіркова</i>
<b>Мова викладання</b>	<i>українська</i>
<b>Курс / семестр</b>	<i>5 курс, 1 семестр</i>
<b>Кількість кредитів ЄКТС</b>	<i>5</i>
<b>Розподіл за видами занять та годинами навчання</b>	<i>Лекції – 20 год. Практичні (семінарські) – . год. Лабораторні – 20 год. Самостійна робота – 110 год.</i>
<b>Форма підсумкового контролю</b>	<i>екзамен</i>
<b>Кафедра</b>	<i>Кафедра інформаційних систем, 702-18-31, <a href="http://www.is.hneu.edu.ua/">http://www.is.hneu.edu.ua/</a></i>
<b>Викладач (-і)</b>	<i>Мінухін Сергій Володимирович, доктор технічних наук, професор кафедри</i>
<b>Контактна інформація викладача (-ів)</b>	<i>serhii.minukhin@hneu.net</i>
<b>Дні занять</b>	<i>вівторок</i>
<b>Консультації</b>	<i>Дистанційні: вівторок 10-15-11-50.</i>
<p>Метою викладання навчальної дисципліни "Високопродуктивні системи обробки великих даних" є формування системи теоретичних знань і придбання практичних умінь і навичок з питань використання технологій розподілених та паралельних обчислень на базі технологій високопродуктивних систем для обробки великих даних.</p>	
<p style="text-align: center;"><b>Передумови для навчання</b> <i>Комп'ютерні мережі, Операційні системи, Програмування, Розподілені та паралельні обчислення</i></p>	
<p style="text-align: center;"><b>Зміст навчальної дисципліни</b></p> <p><b>Змістовий модуль 1. Стандарти, архітектури та принципи побудови розподілених високопродуктивних систем для обробки великих даних</b></p> <p><b>Тема 1.</b> Основні поняття великих та особливості систем обробки великих даних.</p> <p><b>Тема 2.</b> Базові архітектури високопродуктивних систем.</p> <p><b>Тема 3.</b> Розподілені файлові системи.</p> <p><b>Тема 4.</b> Apache Hadoop.</p> <p><b>Тема 5.</b> Програмна модель MapReduce.</p> <p><b>Тема 6.</b> Принципи організації та функціонування фреймворку Apache Spark. Основні компоненти та їх призначення.</p> <p><b>Тема 7.</b> Принципи організації архітектури Apache Spark.</p> <p><b>Тема 8.</b> . Режими розгортання Apache Spark.</p> <p><b>Тема 9.</b> Організація роботи Apache Spark.</p>	
<p style="text-align: center;"><b>Матеріально-технічне (програмне) забезпечення дисципліни</b> Усі компоненти програмного забезпечення є забезпеченням з відкритим кодом.</p>	
<b>Сторінка курсу на платформі Moodle (персональна навчальна система)</b>	<i><a href="https://pns.hneu.edu.ua/course/view.php?id=5476">https://pns.hneu.edu.ua/course/view.php?id=5476</a></i>



## Рекомендовані джерела

### Основні

1. Blackwell M., Sen M. *Large Datasets and You*. [Електронний ресурс]. – Режим доступу: <http://www.matthewblackwell.org/files/papers/bigdata.pdf>.
2. Ross N. *FasteR! HigheR! StrongeR! – A Guide to Speeding Up R Code for Busy People*. [Електронний ресурс]. – Режим доступу: <http://www.noamross.net/blog/2013/4/25/faster-talk.html>.
3. Ryan R. Rosario. *Taking R to the Limit, Part II: Working with Large Datasets*. [Електронний ресурс]. – Режим доступу: [http://www.bytemining.com/wpcontent/uploads/2010/08/r\\_hpc\\_II.pdf](http://www.bytemining.com/wpcontent/uploads/2010/08/r_hpc_II.pdf).
4. *Big Data Specialization*. [Електронний ресурс]. – Режим доступу: <https://www.coursera.org/specializations/big-data>.
5. *Mining Massive Datasets. Onlinecourse*. [Електронний ресурс]. – Режим доступу: <https://online.stanford.edu/course/mining-massive-datasets-self-paced>.
6. Smith M.D., Telang R. *Streaming, Sharing, Stealing: Big Data and the Future of Entertainment* (MIT Press). 2016. [Електронний ресурс]. – Режим доступу: <https://www.amazon.com/Streaming-Sharing-Stealing-Future-Entertainment/dp/0262034794/>.
7. Karimi H.A. *Big Data: Techniques and Technologies in Geoinformatics*. 2014. [Електронний ресурс]. – Режим доступу: <https://www.amazon.com/Big-Data-Techniques-Technologies-Geoinformatics-ebook/dp/B00HZNQKMM/>.
8. Marz N., Warren J. *Big Data: Principles and best practices of scalable realtime data systems*. 2015. [Електронний ресурс]. – Режим доступу: <https://www.amazon.com/Big-Data-Principles-practices-scalable/dp/1617290343/>.

### Додаткові

9. Bahga A., Madisetti V. *Big Data Science & Analytics: A Hands-On Approach*. 2016. [Електронний ресурс]. – Режим доступу: <https://www.amazon.com/Big-DataScience-Analytics-Hands/dp/0996025537/>.
10. Marr B. *Big Data: Using SMART Big Data, Analytics and Metrics To Make Better Decisions and Improve Performance*. 2015. [Електронний ресурс]. – Режим доступу: <https://www.amazon.com/Big-Data-Analytics-Decisions-Performance/dp/1118965833/>.
11. Marr B. *Data Strategy: How to Profit from a World of Big Data, Analytics and the Internet of Things*. 2017. [Електронний ресурс]. – Режим доступу: <https://www.amazon.com/Data>.
12. Jones H. *Data Analytics: An Essential Beginner's Guide To Data Mining, Data Collection, Big Data Analytics For Business, And Business Intelligence Concepts*. 2018. [Електронний ресурс]. – Режим доступу: <https://millionbooks.best/downloads/dataanalytics-the-ultimate-beginners-guide-to-data-analytics>.

### Інформаційні ресурси в Інтернеті

13. Високопродуктивні системи обробки великих даних (122 Комп'ютерні науки). Сайт ПНС ХНЕУ ім. С.Кузнеця. - [Електронний ресурс]. – Режим доступу: <https://pns.hneu.edu.ua/course/view.php?id=5476>.
14. *Apache Spark*. [Електронний ресурс]. – Режим доступу: <https://spark.apache.org>.
15. *apache-spark-internals* [Електронний ресурс]. – Режим доступу: <https://www.sigmoid.com/apache-spark-internals/>.
16. *Spark-ecosystem* [Електронний ресурс]. – Режим доступу: <http://www.jorditorres.org/spark-ecosystem/>.
17. *Apache Hadoop YARN: The Nextgeneration Distributed Operating System* [Електронний ресурс]. – Режим доступу: <https://events.static.linuxfound.org/sites/events/files/slides/ApacheCon%2714%20YARN%20Presentation.pdf>.



18. *YARN scheduler policies* [Електронний ресурс]. – Режим доступу: <http://www.corejavaguru.com/bigdata/hadoop-tutorial/yarn-scheduler>.
19. *YARN Schedulers – FIFO, Fair, and Capacity* [Електронний ресурс]. – Режим доступу: <http://discuss.itversity.com/t/yarn-schedulers-fifo-fair-and-capacity/18098>.
20. *Apache Mesos* [Електронний ресурс]. – Режим доступу: [https://ru.bmstu.wiki/Apache\\_Mesos](https://ru.bmstu.wiki/Apache_Mesos).
21. *YARN* [Електронний ресурс]. – Режим доступу: [https://ru.bmstu.wiki/YARN\\_\(Yet\\_Another\\_Resource\\_Negotiator\)](https://ru.bmstu.wiki/YARN_(Yet_Another_Resource_Negotiator)).

#### Система оцінювання результатів навчання

Поточний контроль: мінімальна кількість балів складає 35 балів, максимальна кількість балів складає 60 балів; екзамен: мінімальна кількість балів складає 25 балів, максимальна кількість балів складає 40 балів;

Більш детальна інформація щодо оцінювання наведена в технологічній карті за адресою <https://pns.hneu.edu.ua/course/view.php?id=5476>.

#### Накопичування рейтингових балів з навчальної дисципліни

Види навчальної роботи	Мах кількість балів
Лекції	5
Лабораторні заняття	10
Захист лабораторних робіт	26
Письмові контрольні роботи	19
<b>Максимальна кількість балів</b>	<b>60</b>

#### Відповідність шкали оцінювання ЄКТС національній системі оцінювання та ХНЕУ ім. С. Кузнеця

Сума балів за всі види навчальної діяльності	Оцінка ЄКТС	Оцінка за національною шкалою	
		для екзамену (іспиту), диференційованого заліку, курсового проекту (роботи), практики, тренінгу	для заліку
90 – 100	A	відмінно	зараховано
82 – 89	B	добре	
74 – 81	C		
64 – 73	D		
60 – 63	E	задовільно	не зараховано
35 – 59	FX	незадовільно	
1 – 34	F		

#### Політики навчальної дисципліни

Політика щодо виконання завдань пізніше встановленого терміну: враховує особливості певної лабораторної роботи та може бути знижена у разі не виконання завдань у повному обсязі; зниженні оцінки за контрольну роботу можливе у разі перевищення строку її написання відповідно до технологічної карти з дисципліни.

Більш детальну інформацію щодо компетентностей, результатів навчання, методів навчання, форм оцінювання, самостійної роботи наведено у робочій програмі навчальної дисципліни <http://www.repository.hneu.edu.ua/handle/123456789/23392>.