



**МІНІСТЕРСТВО ВНУТРІШНІХ СПРАВ УКРАЇНИ**  
**Харківський національний університет внутрішніх справ**  
**Факультет № 4**  
**Кафедра протидії кіберзлочинності**

**ЗАТВЕРДЖЕНО**

на засіданні кафедри  
протидії кіберзлочинності  
протокол № 21 від 18 вересня 2023 р.  
Завідувач кафедри  
**Олександр МАНЖАЙ**



**Лучик Світлана Дмитрівна**

**ПРИКЛАДНИЙ АНАЛІЗ ВЕЛИКИХ ДАНИХ У РОЗСЛІДУВАННІ  
ЗЛОЧИНІВ (ВК. 24)**

**ОПИС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ**

<b>Кафедра</b>	Кафедра протидії кіберзлочинності ( <a href="https://univd.edu.ua/uk/dir/1740">https://univd.edu.ua/uk/dir/1740</a> )
<b>Розробник</b>	Лучик Світлана Дмитрівна, професор кафедри протидії кіберзлочинності, доктор економічних наук, професор
<b>Контактний телефон</b>	+38 0977883470 (моб.)

<b>E-mail</b>	luchiksvitlana@gmail.com
<b>Навчальна дисципліна</b>	Прикладний аналіз великих даних у розслідуванні злочинів
<b>Назва освітньо-професійної програми</b>	"Кібербезпека" («Безпека інформаційних та комунікаційних систем») - CyberSecurity (Security of Information and Communication Systems)
<b>Рівень вищої освіти</b>	Перший (бакалаврський) (НРК України – 6 рівень, FQ-ЕНЕА – перший цикл вищої освіти Рамки кваліфікацій Європейського простору вищої освіти)
<b>Галузь знань</b>	12 Інформаційні технології
<b>Спеціальність</b>	125 Кібербезпека
<b>Статус дисципліни</b>	Вибіркова компонента освітньо-професійної програми, вивчається в 6 семестрі 3 курсу навчання
<b>Обсяг дисципліни в кредитах ECTS</b>	3 кредити ECTS (загальний обсяг - 90 год.)
	аудиторна робота: 44 год. для денної форми навчання (8 год. для заочної форми навчання), з них:
	<ul style="list-style-type: none"> <li>● лекції: 20 год. для денної форми навчання (4 год. для заочної форми навчання)</li> </ul>
	<ul style="list-style-type: none"> <li>● лабораторні заняття: 24 год. для денної форми навчання (4 год. для заочної форми навчання)</li> </ul>
	<ul style="list-style-type: none"> <li>● самостійна робота: 46 год. для денної форми навчання (82 год. для заочної форми навчання)</li> </ul>
<b>Час і місце проведення навчальної дисципліни</b>	Аудиторія та час проведення заняття згідно розкладу.
<b>Консультації з навчальної дисципліни</b>	Аудиторні консультації: аудиторія згідно графіку консультацій. Он-лайн-консультації: письмово в системі дистанційного навчання Moodle, в груповому чаті в Telegram або електронною поштою викладача.
<b>Мета вивчення дисципліни</b>	Вивчення моделей, методів і засобів аналітичної обробки великих даних, а також придбання практичних навичок розв'язання задач професійної діяльності на ПЄОМ з їх використанням.
<b>Завдання вивчення дисципліни</b>	Розвиток у здобувачів вищої освіти практичних навичок використання методів і алгоритмів

	Data Mining для розв'язання практичних задач за допомогою сучасних інструментальних систем; підготовка здобувачів вищої освіти до ефективного розв'язання задач, що постають в подальшій професійній діяльності.
<b>Форми та види проведення навчальних занять</b>	Форма навчання – денна, заочна. Види навчальних занять: лекції, лабораторні, самостійна робота.
<b>Самостійна робота</b>	Виконання індивідуальних завдань до лабораторних робіт. Опрацювання питань для самостійного освоєння з використанням рекомендованої літератури,
<b>Необхідне обладнання</b>	Мультимедійне обладнання (ноутбук та проектор), комп'ютерне забезпечення з виходом у мережу Інтернет.
<b>Індивідуальні завдання</b>	Індивідуальні завдання для лабораторних робіт, реферати.
<b>Контроль</b>	Поточний та підсумковий контроль Поточний: опитування на лабораторних заняттях; участь в дискусіях, обговоренні доповідей, рефератів; підготовка рефератів та доповідей, тестування, виконання самостійних робіт, захист лабораторних робіт. Критерії оцінки поточного контролю викладач повідомляє на першому занятті та перед кожними оцінюванням. Підсумковий контроль: залік.
<b>Політика навчального курсу</b>	Пропущені заняття відпрацьовуються за темою заняття усно і на ПЕОМ згідно графіка консультацій викладача; перевірка тексту робіт на наявність текстових запозичень здійснюється програмою UNICHECK
<b>Інтегральна компетентність, загальні компетентності, спеціальні (фахові) компетентності</b>	Здатність розв'язувати складні спеціалізовані задачі та практичні проблеми у галузі забезпечення інформаційної безпеки та/або кібербезпеки, що характеризується комплексністю та неповною визначеністю умов. ЗК 1. Здатність застосовувати знання у практичних ситуаціях. ЗК 2. Знання та розуміння предметної області та розуміння професії. ЗК 4. Вміння виявляти, ставити та вирішувати проблеми за професійним спрямуванням.

	<p>ЗК 5. Здатність до пошуку, оброблення та аналізу інформації.</p> <p>ФК 2. Здатність до використання інформаційно- комунікаційних технологій, сучасних методів і моделей інформаційної безпеки та/або кібербезпеки</p> <p>ФК 12. Здатність аналізувати, виявляти та оцінювати можливі загрози, уразливості та дестабілізуючі чинники інформаційному простору та інформаційним ресурсам згідно з встановленою політикою інформаційної та/або кібербезпеки.</p>
<b>ЗМІСТ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ ЗА ТЕМАМИ</b>	
<p><b>ТЕМА № 1. Актуальність і тенденції аналізу Великих Даних у правоохоронній сфері.</b></p> <p>Місце, роль та задачі дисципліни. Предмет, структура і зміст дисципліни.</p> <p>Технологічні передумови: Четверта технологічна революція та інформаційний вибух. Вплив технологічних факторів на правоохоронну сферу.</p> <p>Актуальність і тренди аналізу Великих Даних у правоохоронній сфері: зарубіжний досвід. Прогнозування злочинності на основі Big Data. Великі Дані у протидії кіберзлочинності. Протидія фінансовим шахрайствам.</p> <p>Стан аналізу Великих Даних у діяльності Національної поліції України.</p>	
<p><b>ТЕМА № 2. Задачі обробки Великих Даних.</b></p> <p>Специфіка проблем, що виникають під час обробки Великих Даних. Задачі обробки Великих Даних: асоціація, класифікація, кластеризація, аналіз викидів.</p> <p>Етапи опрацювання Великих Даних. Препроцесінг даних, аналітична обробка даних. Типи оброблюваних даних: числові, категоріальні, рангові, текстові, змішані. Оцінювання близькості точок даних у просторі. Масштабування та нормування даних.</p>	
<p><b>ТЕМА № 3. Методи і алгоритми чіткої кластеризації.</b></p> <p>Постановка задачі кластеризації. Критерії кластеризації. Внутрішньокластерна дисперсія. Міжкластерна дисперсія. Оцінки міжкластерної відстані.</p> <p>Ієрархічні алгоритми. Агломеративні алгоритми. Алгоритм K-means. Алгоритм C-means. Алгоритм J-means. Графові алгоритми.</p> <p>Технологія Анакапа.</p>	
<p><b>ТЕМА № 4. Статистичне дослідження великих даних.</b></p> <p>Основні поняття і етапи статистичного дослідження. Статистичні ряди розподілу. Ряди динаміки. Виявлення основної тенденції і тренду зміни в часі досліджуваного явища.</p>	

<p>Кореляція і регресія. Основи машинного навчання.</p>	
<p><b>ТЕМА № 5. Інструментальні засоби обробки Великих Даних в поліцейській діяльності.</b></p> <p>Стандартні загально використовувані засоби. Спеціалізовані засоби, використовувані в поліцейській діяльності: Maltego, I2, Palantir, Ricas, Watson, PredPOL. Проект Європолу ePOOLICE.</p>	
<p><b>Програмні результати навчання</b></p>	<p>ПРН 2. Організовувати власну професійну діяльність, обирати оптимальні методи та способи розв’язування складних спеціалізованих задач та практичних проблем у професійній діяльності, оцінювати їхню ефективність.</p>
	<p>ПРН 3. Використовувати результати самостійного пошуку, аналізу та синтезу інформації з різних джерел щодо ефективного розв’язання спеціалізованих задач професійної діяльності.</p>
	<p>ПРН 4. Аналізувати, аргументувати, приймати рішення при розв’язанні складних спеціалізованих задач та практичних проблем у професійній діяльності, які характеризуються комплексністю та неповною визначеністю умов, відповідати за прийняті рішення.</p>
	<p>ПРН 5. Адаптуватися в умовах частоті зміни технологій професійної діяльності, прогнозувати кінцевий результат.</p>
	<p>ПРН 6. Критично осмислювати основні теорії, принципи, методи і поняття у навчанні та професійній діяльності.</p>
	<p>ПРН 15. Використовувати сучасне програмно-апаратне забезпечення інформаційно-комунікаційних технологій.</p>
	<p>ПРН 57. Проводити кримінальний аналіз при вирішенні завдань поліцейської діяльності.</p>
	<p><b>Методи контролю:</b></p> <ul style="list-style-type: none"> <li>- усний та письмовий;</li> <li>- тестовий.</li> </ul> <p><b>Форми контролю:</b></p> <ul style="list-style-type: none"> <li>- поточний (контрольні роботи, тестування) – 50 балів;</li> <li>- проміжний (ректорські контрольні роботи);</li> <li>- підсумковий (залік) – 50 балів.</li> </ul>
<p><b>КРИТЕРІЇ ОЦІНЮВАННЯ РЕЗУЛЬТАТІВ НАВЧАННЯ</b></p>	

Підсумкові бали з навчальної дисципліни визначаються як сума балів, отриманих здобувачем протягом семестру та балів, набраних на підсумковому контролі (екзамені).

***Підсумкові бали навчальної дисципліни = Загальна кількість балів (перед підсумковим контролем) + Кількість балів за підсумковим контролем***

### **ШКАЛА ОЦІНЮВАННЯ: НАЦІОНАЛЬНА ТА ECTS**

Оцінка в балах	Оцінка за національною шкалою	Оцінка за шкалою ECTS	
		Оцінка	Пояснення
90 – 100	Відмінно (“зараховано”)	A	<b>„Відмінно”</b> – теоретичний зміст курсу освоєний цілком, необхідні практичні навички роботи з освоєним матеріалом сформовані, всі навчальні завдання, які передбачені програмою навчання виконані в повному обсязі, відмінна робота без помилок або з однією незначною помилкою.
80 – 89	Добре (“зараховано”)	B	<b>„Дуже добре”</b> – теоретичний зміст курсу освоєний цілком, необхідні практичні навички роботи з освоєним матеріалом в основному сформовані, всі навчальні завдання, які передбачені програмою навчання виконані, якість виконання більшості з них оцінено числом балів, близьким до максимального, робота з двома – трьома незначними помилками.
75 – 79		C	<b>„Добре”</b> – теоретичний зміст курсу освоєний цілком, практичні навички роботи з освоєним матеріалом в основному сформовані, всі навчальні завдання, які передбачені програмою навчання виконані, якість виконання жодного з них не оцінено мінімальним числом балів, деякі види завдань виконані з помилками, робота з декількома незначними помилками, або з однією – двома значними помилками.
65 – 74	Задовільно (“зараховано”)	D	<b>„Задовільно”</b> – теоретичний зміст курсу освоєний не повністю, але прогалини не мають істотного характеру, необхідні практичні навички роботи з освоєним матеріалом в основному сформовані, більшість передбачених програмою навчання навчальних завдань виконано, деякі з виконаних завдань, містять помилки, робота з трьома значними помилками.
60 – 64		E	<b>„Достатньо”</b> – теоретичний зміст курсу освоєний частково, деякі практичні навички роботи не сформовані, частина передбачених програмою навчання навчальних завдань не виконані, або якість виконання деяких з них оцінено числом балів, близьким до мінімального, робота, що задовольняє мінімуму критеріїв оцінки.
21–59	Незадовільно („не зараховано”)	FX	<b>„Умовно незадовільно”</b> – теоретичний зміст курсу освоєний частково, необхідні практичні навички роботи не сформовані, більшість передбачених

		<p>програм навчання, навчальних завдань не виконано, або якість їхнього виконання оцінено числом балів, близьким до мінімального; при додатковій самостійній роботі над матеріалом курсу можливе підвищення якості виконання навчальних завдань (з можливістю повторного складання), робота, що потребує доробки</p>
1–20	F	<p><b>„Безумовно незадовільно”</b> – теоретичний зміст курсу <b>не освоєно</b>, необхідні практичні навички роботи <b>не сформовані, всі виконані</b> навчальні завдання містять <b>грубі помилки, додаткова самостійна</b> робота над матеріалом курсу <b>не приведе</b> до значимого <b>підвищення якості</b> виконання навчальних завдань, робота, що потребує повної переробки</p>
<p><b>Орієнтовний перелік питань до заліку</b></p>		<ol style="list-style-type: none"> <li>1. Технологічні передумови: Четверта технологічна революція та інформаційний вибух.</li> <li>2. Вплив технологічних передумов на правоохоронну сферу.</li> <li>3. Актуальність і тренди аналізу Великих Даних у правоохоронній сфері: зарубіжний досвід.</li> <li>4. Прогнозування злочинності на основі Big Data.</li> <li>5. Великі Дані у протидії кіберзлочинності.</li> <li>6. Протидія фінансовим шахрайствам.</li> <li>7. Стан аналізу Великих Даних у діяльності Національної поліції України.</li> <li>8. Проблеми в задачах обробки великих масивів даних.</li> <li>9. Етапи обробки великих даних.</li> <li>10. Поняття графу. Вершини і ребра.</li> <li>11. Матриця з'єднань і матриця відстаней.</li> <li>12. Остовне дерево. Мінімальне остовне дерево.</li> <li>13. Задачі і алгоритми на графах. Алгоритм Пріма.</li> <li>14. Алгоритм «найближчого сусіда».</li> <li>15. Технологія Анакапа.</li> <li>16. Класифікація і кластеризація.</li> <li>17. Препроцесінг даних.</li> <li>18. Засоби відбору ознак.</li> <li>19. Засоби редукції розмірності.</li> <li>20. Проблема оцінки близькості об'єктів.</li> <li>21. Числові метрики.</li> <li>22. Метрики Мінковського.</li> <li>23. Метрика Махаланобіса.</li> <li>24. Категоріальна метрика.</li> <li>25. Рангова метрика.</li> <li>26. Масштабування значень об'єктів.</li> <li>27. Нормування значень об'єктів.</li> <li>28. Аналіз викидів.</li> </ol>

	<p>29. Критерії оптимізації в задачах кластеризації.</p> <p>30. Внутрішньо кластерна дисперсія.</p> <p>31. Міжкластерна дисперсія.</p> <p>32. Оцінки міжкластерної відстані.</p> <p>33. Ієрархічні алгоритми кластеризації.</p> <p>34. Агломеративні алгоритми кластеризації.</p> <p>35. Алгоритм K-means.</p> <p>36. Алгоритм C-means.</p> <p>37. Графові алгоритми.</p> <p>38. Генетичні алгоритми.</p> <p>39. Етапи та методи статистичного дослідження</p> <p>40. Поняття статистичного спостереження, етапи його проведення</p> <p>41. Зведення та групування матеріалів статистичного спостереження</p> <p>42. Поняття статистичного ряду розподілу, види рядів розподілу</p> <p>43. Статистичні показники.</p> <p>44. Відносні показники.</p> <p>45. Середні показники.</p> <p>46. Показники варіації ознаки.</p> <p>47. Поняття ряду динаміки, види рядів динаміки.</p> <p>48. Абсолютні та відносні показники ряду динаміки.</p> <p>49. Середні показники ряду динаміки.</p> <p>50. Виявлення основної тенденції зміни в часі досліджуваного явища</p> <p>51. Визначення сезонних коливань рівнів ряду динаміки.</p> <p>52. Прогнозування значень статистичного показника.</p> <p>53. Кореляційний аналіз.</p> <p>54. Регресійний аналіз.</p> <p>55. Спеціалізовані засоби, використовувані в поліцейській діяльності: система Maltego.</p> <p>56. Спеціалізовані засоби, використовувані в поліцейській діяльності: система I2.</p> <p>57. Спеціалізовані засоби, використовувані в поліцейській діяльності: система Palantir.</p> <p>58. Спеціалізовані засоби, використовувані в поліцейській діяльності: система Ricas.</p> <p>59. Спеціалізовані засоби, використовувані в поліцейській діяльності: суперкомп'ютер Watson.</p> <p>60. Спеціалізовані засоби, використовувані в поліцейській діяльності: система PredPOL.</p> <p>61. Проект Європолу ePOOLICE.</p>
--	---



**Рекомендована література (основна), інформаційні ресурси в мережі Інтернеті**

**Основна література.**

1. Струков В.М., Узлов Д.В., Гнусов Ю.В., Колісник Т.П., Можаяєв М.О. Інформаційні технології у правоохоронній діяльності. Частина 1: Високотехнологічні тренди у правоохоронній сфері зарубіжних країн. Навчальний посібник. Харків: ХНУВС, 2020. 256 с.
2. Конспект лекцій.
3. Aggarwal C.C. Data Mining. – Cham: Springer Ltd. Publ. Switzerland, 2015. – 734p.
4. Aggarwal C.C., Reddy C.K. Data Clustering. Algorithms and Applications.- New York: CRC Press, Taylor & Francis Group, 2014. – 648p.
5. Капітонова Ю.В., Кривий С.Л., Летичевський О.А., Луцький Г.М., Печорін М.К. „Основи дискретної математики”, -Київ.: Наук.думка, 2002. – 579 с.
6. Р.М.Трохимчук. Теорія графів. Навчальний посібник для студентів факультету кібернетики - К.: РВЦ “Київський університет”, 1998. - 43 с.

**Допоміжна література.**

7. Зацеркляний М.М. Інформаційні технології у правозастосовній діяльності: Навч. посібник / М.М. Зацеркляний, В.М. Струков. : Х.: ТОВ „Східно-регіональний центр гуманітарно-освітніх ініціатив”; 2010. 332 с.
8. Узлов Д. Ю. Прикладний кримінальний аналіз на базі інформаційно-аналітичної системи «RICAS»: Методичні рекомендації щодо проведення кримінального аналізу/ Узлов Д. Ю., Дегтярьова І.В., Струков В.М. та ін.// Харків: «Юрайт», 2018. – 92 с.
9. Зацеркляний М.М. Основи комп’ютерних технологій для економістів: Навч. посібник / М.М. Зацеркляний, О.Ф. Мельников, В.М. Струков. – К.: ВД „Професіонал”, 2006 р. – 672 с.
10. Han L., Kamber M. Data Mining: Concepts and Techniques. – Amsterdam: Morgan Kaufman Publ., 2006. – 754p.
11. Westphal C. Data Mining for Intelligence, Fraud and Criminal Detection. Advanced Analytic & Information Sharing Technologies / C. Westphal. – Boca Raton : CRC Press, 2009. – 426p.
12. Mena J. Investigative Data Mining for Security and Criminal Detection. – Amsterdam: Elsevier Science, 2003. – 452p.

**Інформаційні ресурси в Інтернеті.**

13. Третя промислова революція: сайт. URL:<https://uk.wikipedia.org/wiki/>
14. Четверта промислова революція: сайт. URL: <https://uk.wikipedia.org/wiki/>
15. Biryukov A. Neural network clustering methods to evaluate the totality of taxpayers according to their degree of creditworthiness // Artificial societies. 2017. V. 12. Issue 1-2 : сайт. URL: [http://artsoc.jes.su/s207751800000103-2-1\(circulationdate:22.07.2018\).DOI:10.18254/S0000103-2-1](http://artsoc.jes.su/s207751800000103-2-1(circulationdate:22.07.2018).DOI:10.18254/S0000103-2-1)