МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Львівський національний університет імені Івана Франка

Факультет прикладної математики та інформатики

Кафедра кібербезпеки

Затверджено
На засіданні кафедри кібербезпеки
факультету прикладної математики
та інформатики
Львівського національного університету
імені Івана Франка
(Протокол №9/24 від 29 серпня 2024 р.)

Завідувач кафедри .

Петро ВЕНГЕРСЬКИЙ

Силабус з навчальної дисципліни

“Конфіденційність та етика в ШІ для кібербезпеки”,

що викладається в межах ОПП Технології штучного інтелекту в кібербезпеці другого (магістерського) рівня вищої освіти для здобувачів з спеціальності 125 – кібербезпека та захист інформації

Львів - 2024

Назва дисципліни	Конфіденційність та етика в ШІ для кібербезпеки
Адреса викладання дисципліни	м. Львів, вул. Університетська 1
Факультет та кафедра, за якою закріплена дисципліна	Факультет прикладної математики та інформатики Кафедра кібербезпеки
Галузь знань, шифр та назва спеціальності	12 – інформаційні технології 125 – кібербезпека та захист інформації
Викладачі дисципліни	Журавчак Даниїл Юрійович, доцент кафедри кібербезпеки
Контактна інформація викладачів	danyil.zhuravchak@lnu.edu.ua; https://ami.lnu.edu.ua/employee/zhuravchak-d-yu Головний корпус ЛНУ ім. І. Франка, каб. 380. м. Львів, вул. Університетська, 1
Консультації з питань навчання по дисципліні відбуваються	Консультації в день проведення лекцій/лабораторних занять (а також за розкладом консультацій кафедри).
Сторінка курсу	https://ami.lnu.edu.ua/course/konfidentsiynist-ta-etyka-v-shi-dlia-kiberbezpeky
Інформація про дисципліну	Дисципліна “Конфіденційність та етика в ШІ для кібербезпеки” є вибірковою дисципліною з спеціальності 125 – кібербезпека та захист інформації для освітньої програми «Технології штучного інтелекту в кібербезпеці», яка викладається в 3-му семестрі в обсязі 3-ох кредитів (за Європейською Кредитно-Трансферною Системою ECTS).
Коротка анотація дисципліни	Дисципліна “Конфіденційність та етика в ШІ для кібербезпеки” охоплює питання захисту конфіденційних даних та етичні аспекти використання штучного інтелекту в кібербезпеці. Студенти ознайомляться з основами конфіденційності, принципами цілісності та доступності даних, а також регуляціями захисту інформації (GDPR, ISO27001). Курс включає теоретичні основи конфіденційності, аналіз етичних викликів у використанні ШІ та практичні завдання з оцінки ризиків і захисту даних.
Мета та цілі дисципліни	Метою викладання дисципліни є формування у студентів навичок захисту конфіденційних даних, розуміння етичних аспектів використання ШІ у сфері кібербезпеки, а також ознайомлення з сучасними стандартами та регуляціями захисту даних.
Література для вивчення дисципліни	**Базова література** 1. Regan P. M. “Ethics and Data Privacy in the Age of Smart Technologies”, Springer, 2021. 2. Allen B., Hicks T. “Building a Cyber Risk Management Program”, O’Reilly, 2024. 3. Sarker I.H. “Machine Learning: Algorithms, Real-World Applications and Research Directions”, SN Computer Science, 2021. 4. Floridi L. “The Ethics of Artificial Intelligence”, Springer, 2020. **Допоміжна література**

	1. Stallings W. "Cryptography and Network Security: Principles and Practice", 7th Edition, Pearson, 2017. 2. Dwork C., Roth A. "The Algorithmic Foundations of Differential Privacy", Foundations and Trends® in Theoretical Computer Science, 2014. 3. Elliott M. "An Introduction to Privacy Engineering and Risk Management in Federal Systems", National Institute of Standards and Technology (NIST), Special Publication 800-37, 2020. 4. Sweeney L. "k-Anonymity: A Model for Protecting Privacy", International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2002. 5. O'Neil C. "Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy", Crown, 2016. 6. Goodman B., Flaxman S. "European Union regulations on algorithmic decision-making and a “right to explanation”", AI Magazine, 2017. 7. Regan P. M. "Ethics and Data Privacy in the Age of Smart Technologies", Springer, 2019. 8. ISO/IEC 27001:2022. "Information Security Management Systems — Requirements". 9. EU GDPR (General Data Protection Regulation), 2016/679. 10. The European Parliament and Council. "Directive (EU) 2016/1148 concerning measures for a high common level of security of network and information systems across the Union (NIS Directive)". 11. Goodfellow I., Bengio Y., Courville A. "Deep Learning", MIT Press, 2016. 12. Floridi L. "The Ethics of Artificial Intelligence", Springer, 2020. 13. Allen B., Hicks T. "Building a Cyber Risk Management Program", O'Reilly, 2024. 14. Sarker I.H. "Machine Learning: Algorithms, Real-World Applications and Research Directions", SN Computer Science, 2021. **Інтернет-ресурси** 15. The European Data Protection Board (EDPB). "Guidelines on Data Protection Impact Assessment (DPIA)". https://edpb.europa.eu 16. National Institute of Standards and Technology (NIST). "Privacy Framework: A Tool for Improving Privacy through Enterprise Risk Management". https://www.nist.gov 17. The Future of Privacy Forum (FPF). "Privacy and Artificial Intelligence". https://fpf.org
Обсяг курсу	Загальний обсяг: 90 годин. Аудиторних занять: 48 год., з них 16 год. лекцій та 32 год. лабораторних робіт. Самостійної роботи: 42 год.
Очікувані результати навчання	У результаті вивчення навчальної дисципліни студент має набути таких компетентностей: Знання: 1. Знати основи конфіденційності та етики в кібербезпеці: ○ Розуміти принципи конфіденційності, цілісності та доступності даних (CIA-триада).

	- Знати основні підходи до захисту персонально ідентифікованої інформації (PII). - Знати міжнародні стандарти кібербезпеки та захисту даних, включаючи GDPR, PCI DSS, ISO27001, SOC2, а також регуляції, що стосуються конфіденційності в Україні (Держспецзв'язок та КСЗІ). 2. Ознайомитися з етичними аспектами використання штучного інтелекту: - Знати етичні ризики використання ШІ, включаючи питання упередженості, дискримінації, впливу на приватність. - Розуміти законодавчі аспекти регулювання використання ШІ, включаючи вимоги "права на пояснення" алгоритмічних рішень. 3. Мати знання щодо захисту даних у хмарних середовищах: - Розуміти архітектуру даних у хмарах та основи безпеки в хмарних середовищах. - Ознайомитися із загальними ризиками, пов'язаними з обробкою та зберіганням даних у хмарних платформах. Уміння: 1. Аналізувати та оцінювати ризики конфіденційності: - Уміти проводити аналіз загроз конфіденційності та визначати відповідні ризики. - Використовувати підходи до оцінювання ризиків, зокрема, якісну та кількісну оцінку ризиків. 2. Застосовувати етичні принципи при роботі з ШІ: - Ідентифікувати потенційні етичні проблеми при використанні ШІ у різних аспектах кібербезпеки. - Розробляти та застосовувати рекомендації для мінімізації етичних ризиків у використанні штучного інтелекту. 3. Використовувати інструменти для забезпечення конфіденційності: - Використовувати інструменти та алгоритми ШІ для виявлення загроз та аналізу ризиків. - Проводити аналіз даних для ідентифікації аномалій та виявлення потенційних загроз конфіденційності. Навички: 1. Розробка та реалізація політик безпеки: - Вміти створювати політики безпеки даних, що відповідають сучасним стандартам та регуляціям. - Застосовувати практики захисту персональних даних на різних рівнях системи (локально, хмарно тощо). 2. Критичний аналіз і прийняття рішень: - Розвивати здатність критично аналізувати інформацію та оцінювати відповідність етичним принципам у сфері кібербезпеки. - Визначати ефективні методи зниження ризиків для забезпечення конфіденційності. Професійні компетенції: 1. Робота з нормативними документами: - Орієнтуватися в нормативно-правових актах, що стосуються конфіденційності та захисту даних, і застосовувати їх у професійній діяльності. 2. Комунікація та співпраця:

	○ Вміти працювати в команді для аналізу та оцінювання ризиків, проведення аудиту конфіденційності та захисту даних. ○ Пояснювати складні технічні аспекти конфіденційності та етики, зважаючи на різноманітну аудиторію (технічну та нетехнічну).
Ключові слова	Конфіденційність, етика, штучний інтелект (ШІ), персонально ідентифікована інформація (PII), CIA-триада (Confidentiality, Integrity, Availability), кібербезпека, машинне навчання, етичні ризики, міжнародні стандарти безпеки (GDPR, PCI DSS, ISO27001), захист даних, хмарні технології, анонімізація даних, диференційна конфіденційність, управління ризиками, нормативно-правові акти, регуляції конфіденційності, алгоритмічна упередженість, права на пояснення (Right to Explanation), кіберризики, комплаєнс.
Формат курсу	Очний Проведення лекцій, лабораторних робіт і консультацій.
Теми	Теми подані у Схемі курсу нижче
Підсумковий контроль, форма	Залік в кінці 3 семестру
Пререквізити	Для вивчення курсу студенти потребують базові знання з дисципліни: • Кібербезпека. • Штучний інтелект.
Навчальні методи та тех.-ніки, які будуть використовуватися під час викладання курсу	Лекції — інформаційно-рецептивний метод для викладу теоретичних знань. Практичні заняття — практичне застосування набутих знань, включаючи використання інструментів ШІ для аналізу загроз. Самостійна робота — виконання індивідуальних завдань та підготовка рефератів. Воркшопи — обговорення етичних кейсів, робота у групах над практичними завданнями.
Необхідне обладнання	Комп'ютерний клас із вільно-доступним програмним забезпеченням, локальна комп'ютерна мережа, доступ до Internet мережі.
Критерії оцінювання (окремо для кожного виду навчальної діяльності)	Оцінювання проводиться упродовж семестру за 100-бальною шкалою. Бали нараховуються за такими видами робіт з наступним співвідношенням: • робота під час лабораторних занять: 43% семестрової оцінки; максимальна кількість балів 43; • самостійна робота: 32% семестрової оцінки; максимальна кількість балів 32; • контрольна робота: 25% семестрової оцінки; максимальна кількість балів 25. Підсумкова максимальна кількість балів 100. **Академічна доброчесність:** очікується, що роботи студентів будуть оригінальними дослідженнями чи міркуваннями. Списування та втручання в роботу інших студентів становлять, але не обмежують, приклади можливої академічної недоброчесності. Виявлення ознак академічної недоброчесності в написанні завдань є підставою для її незарахуванння викладачем, незалежно від масштабів плагіату чи обману. Жодні форми порушення академічної доброчесності не толеруються. **Відвідання занять** є важливою складовою навчання. Очікується, що всі

	студенти відвідають усі лекції та лабораторні заняття курсу. Студенти повинні інформувати викладача про неможливість відвідати заняття. У будь-якому випадку студенти зобов'язані дотримуватися термінів визначених для виконання всіх видів робіт, передбачених курсом. **Література.** Уся література, яку студенти не зможуть знайти самостійно, буде надана викладачем виключно в освітніх цілях без права її передачі третім особам. Студенти заохочуються до використання також й іншої літератури та джерел, яких немає серед рекомендованих. **Політика виставлення балів.** Враховуються бали, набрані при поточному контролі та бали за виконання лабораторних робіт. При цьому обов'язково враховуються присутність на заняттях та активність студента під час лабораторного заняття; недопустимість пропусків та запізнень на заняття; користування мобільним телефоном, планшетом чи іншими мобільними пристроями під час заняття в цілях не пов'язаних з навчанням; списування та плагіат; несвоєчасне виконання поставленого завдання і т. ін. **Оцінювання роботи на лабораторних заняттях:** студенти на 8 лабораторних заняттях виконують різноманітні вправи та завдання. У підсумку максимальна кількість балів студента за роботу на лабораторних заняттях - 43. Бали оцінювання роботи на практичних заняттях нараховуються за наступним співвідношенням: 3-4 – студент в повному обсязі володіє навчальним матеріалом, має повне розуміння досліджуваної проблеми, надає правильні відповіді на запитання по темі, має свої ідейні міркування щодо реалізації даної проблеми; 1-2 – студент не достатньо розуміє приведені ним результати, вагається та надає неточні/не конкретні відповіді на запитання по темі; 0 - студент безвідповідально відносився до виконання роботи, студент виявляє нульовий рівень компетентності та зовсім не засвоїв розглянутий матеріал. Також за запропоновані новітні методи, активність і креативність під час лабораторних занять студент може набрати додаткових 9 балів. **Оцінювання самостійної роботи:** студенти самостійно виконують завдання для 8 домашніх робіт. У підсумку максимальна кількість балів студента за самостійну роботу - 32. Бали оцінювання самостійної роботи нараховуються за наступним співвідношенням: 3-4 – робота цілком і повністю відображає індивідуальне завдання студента, містить правильні висновки, ілюстрований (за потреби) відповідними графіками, студент має повне розуміння розглянутої теми, надає правильні відповіді на запитання по темі; 2 – робота в достатній мірі відображає індивідуальне завдання студента, містить допустимі висновки, ілюстрований (за потреби) відповідними графіками, які частково відображають суть виконаного завдання, присутні неточності та незначні помилки у відповідях на запитання по темі; 1 – звіт містить загальні формулювання завдання, висновки нечіткі, необхідні ілюстрації відсутні, студент не досить добре розуміє розглянутий матеріал, надає неточні/не конкретні відповіді на запитання по темі; 0 - робота відсутня/не відповідає темі, студент зовсім не засвоїв розглянутий матеріал.

	Оцінювання контрольної роботи: 10 тестових теоретичних питань (по 1 балу за кожне) та 3 практичних завдання (по 5 балів кожне). Бали оцінювання практичного завдання залікової (контрольної) роботи нараховуються за наступним співвідношенням: 5 – студент в повному обсязі володіє навчальним матеріалом, має повне розуміння досліджуваної проблеми, надає правильні відповіді на запитання по темі, має свої ідейні міркування щодо реалізації даної проблеми; 3-4 – студент достатньо розуміє розглянутий матеріал, демонструє достатній рівень обґрунтування результатів (або з несуттєвими недоліками); 2 – студент не достатньо розуміє приведені ним результати, вагається та надає неточні/не конкретні відповіді на запитання по темі; 1 – студент погано розуміє приведені результати, у більшості надає помилкові відповіді на питання по роботі; 0 - студент безвідповідально відносився до виконання завдання, студент виявляє нульовий рівень компетентності та зовсім не засвоїв розглянутий матеріал. **Критерії оцінювання результатів неформальної освіти:** Нарахування балів відбувається за публікацію студентом тез доповідей на конференціях, наукових статей, за участь студента у діяльності наукових гуртків, семінарів, круглих столів, конкурсів, участь у заходах неформальної освіти, за отримання сертифікатів про проходження навчання на різних освітніх платформах (Coursera, Prometheus тощо). Кількість балів визначається відсотком покриття результатів відповідної активності до вимог результатів навчання з навчальної дисципліни.
Питання до контрольної роботи.	**Поняття конфіденційності у кібербезпеці: основні принципи та підходи.** • Що таке конфіденційність і чому вона важлива в кібербезпеці? • Як пов'язані між собою конфіденційність, цілісність та доступність (CIA-триада)? • Приклади порушень конфіденційності та їх наслідки. **Етичні аспекти використання штучного інтелекту в кібербезпеці.** • Які етичні ризики пов'язані з використанням штучного інтелекту у кібербезпеці? • Що таке "алгоритмічна упередженість" і як вона може вплинути на безпеку? • Як забезпечити етичне використання ШІ при обробці персональних даних? **Персонально ідентифікована інформація (PII): визначення та захист.** • Що таке персонально ідентифікована інформація (PII)? • Які основні методи захисту PII використовуються в кібербезпеці? • Наведіть приклади технологій для захисту конфіденційних даних. **Регуляції захисту даних: міжнародні та локальні стандарти.** • Що таке GDPR і які вимоги він встановлює щодо захисту персональних даних? • Яка роль Держспецзв’язку та КСЗІ в Україні? • Які основні стандарти (наприклад, PCI DSS, ISO27001) використовуються для забезпечення кібербезпеки? **Захист даних у хмарних середовищах.** • Які особливості захисту даних у хмарних середовищах? • Які основні ризики, пов'язані з обробкою та зберіганням даних у хмарах?

	• Як забезпечити конфіденційність даних при використанні хмарних платформ? **Методи оцінювання ризиків у кібербезпеці.** • Що таке якісна та кількісна оцінка ризиків? • Які основні етапи управління ризиками інформаційної безпеки? • Як машинне навчання може допомогти в оцінюванні ризиків? **Застосування диференційної конфіденційності у кібербезпеці.** • Що таке диференційна конфіденційність і в яких випадках її варто використовувати? • Як диференційна конфіденційність забезпечує захист даних? • Приклади використання диференційної конфіденційності у різних галузях. **Права користувачів на захист конфіденційності.** • Що таке "право на пояснення" у контексті алгоритмічних рішень? • Як забезпечити права користувачів на конфіденційність в умовах використання ШІ? • Які інструменти допомагають реалізувати права на доступ, виправлення та видалення даних? **Комплаєнс та його роль у забезпеченні конфіденційності даних.** • Що таке комплаєнс і яка його роль у забезпеченні кібербезпеки? • Які заходи комплаєнсу можуть бути впроваджені для дотримання вимог захисту даних? • Приклади комплаєнс-інструментів у сфері кібербезпеки. **Машинне навчання для забезпечення конфіденційності.** • Які алгоритми машинного навчання можуть використовуватися для виявлення загроз? • Як аналіз аномалій допомагає забезпечити конфіденційність даних? • Які ризики можуть виникати при використанні машинного навчання у кібербезпеці?
Опитування	Анкету-оцінку з метою оцінювання якості курсу буде надано по завершенню курсу.

Схема курсу

Тиж.	Тема, план, короткі тези	Форма діяльності (заняття)	Література	Завдання, год.	Термін виконання
1	**Тема 1. Основи конфіденційності та етики у кібербезпеці** - Визначення конфіденційності, цілісності та доступності (CIA-триада). - Етичні принципи у кібербезпеці.	Лекція, самостійна робота	[1, 3, 5, 7]	2 год. 6 год.	1 тиждень
	Тема 1. Основи конфіденційності та етики у кібербезпеці - Практичні завдання щодо аналізу конфіденційності та етичних аспектів у кібербезпеці.	лаб	[1, 3, 5, 7]	4 год.	

2	**Тема 2 Персонально ідентифікована інформація (PII) та її захист** - Поняття PII, основні методи захисту персональних даних.	лекція, самостійна робота	[1, 2, 5, 6]	2 год. 6 год.	1 тиждень
	Тема 2. Персонально ідентифікована інформація (PII) та її захист - Практичні завдання щодо ідентифікації PII та використання інструментів захисту.	лаб.	[1, 2, 5, 6]	4 год.	
3	**Тема 3. Регуляції захисту даних** - Міжнародні та національні стандарти захисту даних (GDPR, ISO27001, Держспецзв'язок).	лекція, самостійна робота	[3, 4, 6, 7]	2 год. 6 год.	1 тиждень
	Тема 3. Регуляції захисту даних - Лабораторні завдання з аналізу дотримання стандартів та регуляцій конфіденційності.	лаб	[3, 4, 6, 7]	4 год.	
4	**Тема 4 Етичні аспекти використання штучного інтелекту (ШІ) у кібербезпеці** - Потенційні ризики, пов'язані з ШІ, та етичні стандарти.	лекція, самостійна робота	[5, 7, 9, 10]	2 год. 6 год.	1 тиждень
	Тема 4. Етичні аспекти використання штучного інтелекту (ШІ) у кібербезпеці - Лабораторні завдання з оцінки етичних ризиків при використанні ШІ.	лаб.	[5, 7, 9, 10]	4 год.	
5	**Тема 5. Захист даних у хмарних середовищах** - Архітектура даних та специфіка захисту у хмарах.	лекція, самостійна робота	[4, 6, 8, 11]	2 год. 6 год.	1 тиждень
	Тема 5. Захист даних у хмарних середовищах - Лабораторні завдання з аналізу ризиків і методів захисту даних у хмарах.	лаб.	[4, 6, 8, 11]	4 год.	
6	**Тема 6. Методи оцінювання ризиків у кібербезпеці** - Якісна та кількісна оцінка ризиків, процес управління ризиками.	лекція, самостійна робота	[2, 4, 8, 12]	2 год. 6 год.	1 тиждень
	Тема 6. Методи оцінювання ризиків у кібербезпеці - Лабораторні завдання з проведення оцінки ризиків за допомогою різних методів.	лаб.	[2, 4, 8, 12]	4 год.	
7	**Тема 7 Машинне навчання для забезпечення конфіденційності** - Алгоритми виявлення загроз та	лекція, самостійна робота	[5, 6, 9, 12]	2 год. 3 год.	1 тиждень

	аналіз аномалій для забезпечення конфіденційності.				
	Тема 7. Машинне навчання для забезпечення конфіденційності - Лабораторні завдання з використання алгоритмів машинного навчання для ідентифікації загроз.	лаб.	[5, 6, 9, 12]	4 год.	
8	**Тема 8. Комплаєнс та його роль у забезпеченні конфіденційності даних** - Основні заходи комплаєнсу, приклади інструментів для дотримання конфіденційності.	Лекція самостійна робота	[4, 7, 9, 11]	2 год. 3 год.	1 тиждень
	Тема 8. Комплаєнс та його роль у забезпеченні конфіденційності даних - Лабораторні завдання з оцінки політик комплаєнсу та управління конфіденційністю.	Лаб.	[4, 7, 9, 11]	4 год.	1 тиждень
	Всього			**90**	