

Технології штучного інтелекту та безпека даних

Освітня програма «Економіка і аналітика», 1 курс · 4 кредити ECTS

Лекція 1. Що таке штучний інтелект: від алгоритмів до генеративних моделей

Мета лекції: сформувати в студентів базове й коректне розуміння того, що таке штучний інтелект (ШІ), чим він відрізняється від машинного навчання, глибокого навчання та генеративного ШІ, а також простежити історичний шлях технології — від перших ідей до появи великих мовних моделей (LLM). Лекція оглядова, без заглиблення в математику й код.

Формат конспекту: тези для кожного з 17 пунктів плану лекції. Кожен блок можна розгортати усно прикладами, ілюстраціями з презентації та короткою дискусією зі студентами (позначено як «Питання для аудиторії»).

1. Чому виник поточний інтерес до ШІ

- Google Trends фіксує різкий стрибок пошукового інтересу до запитів «AI» та «LLM», що почався наприкінці 2022 року — після публічного релізу ChatGPT.
- До цього моменту термін «штучний інтелект» існував у науці й індустрії понад 70 років, але залишався нішевим.
- **Ключова хибна теза, яку варто одразу розвінчати:** «ШІ» і «LLM (великі мовні моделі)» — це не синоніми. LLM — лише один із сучасних різновидів ШІ, хоч і найпомітніший.
- Курс і ця лекція покликані розвести терміни й показати, що ШІ — набагато ширше поняття, ніж чат-боти.
- Питання для аудиторії: «Що ви уявляєте, коли чуєте слово „штучний інтелект“?» — зазвичай відповідь буде «ChatGPT» або «робот» — це і є ілюстрація проблеми термінології.

2. Різниця між ШІ, машинним навчанням, глибоким навчанням і генеративним ШІ

Це поняття не рівнозначні, а вкладені одне в одне, як матрьошка (концентричні кола):

- **Штучний інтелект (AI)** — найширше поняття: будь-яка спроба відтворити машиною процес прийняття рішень, властивий людині.
- **Машинне навчання (ML)** — підмножина ШІ, де система вчиться на даних, а не виконує жорстко прописані правила.
- **Глибоке навчання (DL)** — підмножина ML, що використовує багатоварові штучні нейронні мережі.
- **Генеративний ШІ (Generative AI)** — підмножина DL, що вміє генерувати новий контент (текст, зображення, код); великі мовні моделі (LLM) — окремий випадок генеративного ШІ, спеціалізований на тексті.

Формула для запам'ятовування: **LLM $\subset$ Generative AI $\subset$ Deep Learning $\subset$ Machine Learning $\subset$ AI.**

Дидактична примітка: Варто намалювати або показати на слайді 4 концентричні кола й попросити студентів самим навести приклад технології для кожного рівня.

3. Визначення ШІ. Приклад типового ШІ — гра в хрестики-нулики

- **Визначення:** штучний інтелект — це спроба за допомогою математичних формул та алгоритмів відтворити логіку людського міркування та прийняття рішень.
- Важливо: ШІ не обов'язково означає «навчання» чи «розум» — досить, щоб система імітувала процес прийняття рішень.

Приклад: алгоритм гри у хрестики-нулики, побудований на чотирьох простих правилах:

- Правило 1 (Перемога) — якщо є два свої знаки в ряд, поставити третій і виграти.
 - Правило 2 (Блок) — якщо два знаки в ряд у суперника, заблокувати.
 - Правило 3 (Кути/центр) — якщо центр вільний, зайняти його (відкриває найбільше варіантів).
 - Правило 4 (Рандом) — якщо жодне правило не спрацювало, обрати випадкову клітинку.
- Такий алгоритм повністю підпадає під визначення «штучного інтелекту», хоча це просто послідовність команд «якщо..., то...» без жодного навчання чи статистики.

Дидактична примітка: Гарний момент для «розвінчання магії»: студенти часто дивуються, що настільки просте правило вже вважається ШІ.

4. Визначення машинного навчання

- **Визначення:** машинне навчання — це різновид ШІ, що має подобу «пам'яті» і використовує її для покращення результатів.
- По суті ML — це поєднання статистики та алгоритмів: замість того, щоб людина прописувала правила (як у хрестиках-нуліках), система сама виводить закономірності з великої кількості прикладів.
- Головна відмінність від «класичного» ШІ (п. 3): джерелом «інтелекту» стають дані, а не набір заздалегідь написаних правил.
- ML поділяється на два основні підходи: **навчання з учителем** (supervised) і **навчання без учителя** (unsupervised) — розглянемо обидва на прикладах.

5. Навчання з учителем — на прикладі розпізнавання цифр

Послідовність навчання з учителем повторює звичний навчальний процес людини:

- **Тренування** — моделі показують приклад і вже відоме правильне рішення (тисячі зображень рукописних цифр із підписом, яка це цифра).
 - **Валідація** — моделі дають аналогічний, але новий приклад і перевіряють, чи тренування дало результат.
 - **Тестування** — фінальна перевірка на прикладі, якого модель ще не бачила.
- Аналогія: підручник із розв'язаними прикладами → домашні завдання з ключем відповідей → контрольна робота без підказок.
 - Приклад: датасет рукописних цифр (0–9) — тисячі зображень, кожне розміром 28×28 пікселів, з відомою правильною міткою.

Дидактична примітка: Можна прямо спитати аудиторію: «Чи нагадує це вам щось?» — і підвести до паралелі зі шкільним/студентським навчанням.

6. Типи помилок моделі (I і II роду)

- **Помилка I роду (false positive, «хибне спрацювання»)** — модель каже «так», хоча насправді «ні».
- **Помилка II роду (false negative, «хибний пропуск»)** — модель каже «ні», хоча насправді «так».

Ілюстрація: медичний тест на вагітність — чоловіку помилково кажуть «ви вагітні» (I рід), а вагітній жінці кажуть «ви не вагітні» (II рід).

Економічний контекст (важливо для цієї аудиторії): будь-яка модель класифікації — кредитний скоринг, антифрод-система, фільтр спаму — завжди балансує між цими двома типами помилок, і «ціна» кожної з них різна:

- У кредитному скорингу: помилково відмовити надійному клієнту (I рід) коштує втраченого прибутку; помилково видати кредит ненадійному (II рід) коштує прямих збитків.
- У виявленні шахрайства: пропустити реальну шахрайську операцію (II рід) зазвичай значно дорожче, ніж один раз перестрахуватися й заблокувати легітимну (I рід).
- Немає моделі без помилок — є свідомий вибір, у який бік «зміщувати» поріг чутливості залежно від бізнес-вартості помилки.

7. Навчання без учителя — на прикладі кластеризації клієнтів магазину

- **Визначення:** у навчанні без учителя комп'ютеру не дають «правильної відповіді» — лише характеристики об'єктів і завдання «згрупувати схожі точки в групи».

Приклад: кластеризація клієнтів магазину за двома ознаками — річний дохід і рівень витрат (spending score):

- Модель сама, без підказок, виділяє групи клієнтів (кластери) — наприклад, «низький дохід + великі витрати» (схильні до імпульсивних покупок, кандидати на розстрочку) або «високий дохід + великі витрати» (преміум-клієнти для утримання лояльності).
- Практичне застосування для економіста: сегментація клієнтської бази, персоналізовані пропозиції, оцінка кредитного ризику, оптимізація маркетингового бюджету.

Дидактична примітка: Можна показати графік кластеризації (5 кольорових груп) і попросити студентів пояснити своїми словами бізнес-сенс кожного кластера.

8. Чому ШІ справляється з багатовимірними задачами краще за людину

- Людина здатна візуально й інтуїтивно аналізувати залежності максимум у 2–3 вимірах (графік, площа, 3D-об'єкт).

- Реальні економічні та бізнес-задачі містять десятки й сотні змінних одночасно (дохід, вік, регіон, історія покупок, сезонність тощо) — людина фізично не може «побачити» чи уявити таку залежність.
- Математичному алгоритму не потрібно «уявляти» простір геометрично — він оперує рівняннями й матрицями в N-вимірному просторі так само «легко», як і в двовимірному.
- Саме тому ШІ знаходить закономірності (кореляції, кластери, прогнози), недоступні людському сприйняттю, — не тому що він «розумніший», а тому що він не обмежений біологією зорового аналізатора.

Дидактична примітка: Гарна ілюстрація — спробувати показати графік залежності від чотирьох змінних одночасно і визнати, що людина такий графік прочитати не може.

9. Визначення глибокого навчання

- **Визначення:** глибоке навчання — це використання штучних нейронних мереж, де навчальними прикладами є «представлення» (representations) даних.
- Відмінність від «класичного» машинного навчання: у класичному ML людина сама визначає, які ознаки (features) важливі для аналізу (наприклад, дохід, вік, кредитна історія).
- У глибокому навчанні мережа сама виявляє потрібні ознаки безпосередньо із «сирих» даних — пікселів зображення, символів тексту, звукової хвилі — без ручного відбору ознак людиною.
- Це стало можливим завдяки багатошаровій структурі нейронної мережі: кожен наступний шар узагальнює інформацію попереднього.

10. Короткий огляд лінійного перцептрона

- Перцептрон — найпростіша «клітина» штучної нейронної мережі, побудована за грубою аналогією з біологічним нейроном (дендрити приймають сигнали, аксон передає результат далі).
- Механіка (без заглиблення в математику): перцептрон отримує кілька вхідних значень ($x_1 \dots x_n$), кожне з яких має свою «вагу» ($w_1 \dots w_n$), рахує їх зважену суму, застосовує до неї функцію активації й видає один вихідний сигнал.
- «Навчання» нейронної мережі — це процес поступового підбору ваг (w) так, щоб вихід мережі якомога точніше збігався з правильною відповіддю на тренувальних прикладах.
- Тисячі й мільйони таких перцептронів, з'єднаних у шари, утворюють «глибоку» нейронну мережу — принципово це та сама ідея, лише в набагато більшому масштабі.

11. Розпізнавання цифр нейронними мережами

Повертаємось до прикладу з п. 5, але тепер розв'язуємо задачу за допомогою нейронної мережі:

- **Вхідний шар** — 784 нейрони (зображення цифри 28×28 пікселів, «розгорнуте» в один довгий вектор чисел).

- **Приховані шари** — наприклад, 128 і 64 нейрони; вони поступово виявляють проміжні ознаки — лінії, вигини, форми.
- **Вихідний шар** — 10 нейронів (по одному на кожну цифру від 0 до 9); мережа обирає варіант із найвищою ймовірністю.
- Ключова відмінність від «класичного» ML: розробнику не потрібно самому пояснювати мережі, «що таке цифра 6» — вона сама виводить це з тисяч прикладів (пряме продовження ідеї з п. 9).

12. Визначення генеративного ШІ

- **Визначення:** генеративний ШІ — це різновид штучного інтелекту, здатний генерувати новий контент (текст, зображення, звук, код) за допомогою «породжуючих» моделей.
- Принцип роботи: опрацювавши величезні обсяги даних під час навчання, модель навчається продукувати власні дані, схожі за структурою й стилем на ті, що вона «бачила».
- Технологічно це ті самі нейронні мережі з п. 9–11, тільки значно більшого масштабу: вони отримують запит (prompt), «розкладають» його на складові й через ряд перетворень видають відповідь, максимально подібну на людську.
- Великі мовні моделі (LLM, наприклад ChatGPT, Claude, Gemini) — це генеративний ШІ, спеціалізований на роботі з текстом.

13. Історія розвитку штучного інтелекту — від перцептрона Розенблатта до великих мовних моделей

- **1943** — перша математична модель штучного нейрона (Мак-Каллок і Піттс).
- **1950** — «машина Тьюрінга» та ідея тесту Тьюрінга як способу перевірити, чи може машина «мислити».
- **1956** — Дартмутська конференція, де вперше офіційно пролунав термін «Artificial Intelligence» — умовне «народження» ШІ як наукової дисципліни.
- **1957–1958** — перцептрон Френка Розенблатта — перша штучна нейронна мережа, здатна навчатися на прикладах.
- **1966** — перший чат-бот ELIZA, що імітував психотерапевта простими текстовими шаблонами.
- **1980-ті** — «перший сплеск» популярності: експертні системи (детальніше — п. 15).
- **1997** — суперкомп'ютер IBM Deep Blue вперше перемагає чинного чемпіона світу з шахів.
- **2000–2010-ті** — «другий сплеск» популярності: рекомендаційні системи й таргетована реклама (детальніше — п. 16).
- **2011–2015** — голосові асистенти й перші масові AI-продукти (IBM Watson, Google Now, Amazon Echo).
- **2020-ті** — поява й вибухове поширення великих мовних моделей (детальніше — п. 17).

Дидактична примітка: Тут доречно показати слайд-таймлайн і наголосити: шлях від першої ідеї (1943) до масового продукту (2022) зайняв майже 80 років — це готує аудиторію до наступного блоку про стримуючі фактори.

14. Стримуючі фактори розвитку ШІ

Щоб технологія ШІ перейшла від наукової ідеї до масового продукту, мають одночасно «зійтися» три фактори:

- **Технічні можливості** — достатня обчислювальна потужність (сучасні GPU, дата-центри) для тренування великих моделей за прийнятний час.
- **Інтерес великого бізнесу** — готовність компаній інвестувати значні кошти в розробку й впровадження.
- **Дані для навчання** — наявність достатньої кількості якісних, оцифрованих даних, на яких можна навчати модель.

Ці три фактори — наскрізна логіка для розуміння всієї історії ШІ: кожен «сплеск популярності» (п. 15–17) можна пояснити тим, що саме тоді ці три умови збіглися в конкретній галузі.

15. Перший сплеск популярності — експертні системи 80-х у банківській та інших сферах

Аналізуємо через призму трьох факторів із п. 14:

- **Технічні можливості:** комп'ютери вже здатні зберігати порівняно великі обсяги даних.
 - **Дані:** банки накопичили достатньо даних про кредитні історії клієнтів у цифровому вигляді.
 - **Інтерес бізнесу:** банки зацікавлені приймати рішення (наприклад, про видачу кредиту) на основі цих масивів даних, а не лише експертної думки співробітника.
- Результат — «експертні системи» (expert systems): база знань (Knowledge Base) + механізм логічного виведення (Inference Engine) + інтерфейс користувача. Це системи підтримки прийняття рішень на основі заздалегідь формалізованих правил і накопичених даних.

16. Другий сплеск популярності — таргетована реклама (2000-ні – 2010-ті)

Аналізуємо через ту саму призму:

- **Технічні можливості:** великі сервери здатні зберігати значні обсяги даних про поведінку користувачів у браузері під час онлайн-покупок.
- **Дані:** історичні дані великих онлайн-магазинів вже накопичилися й дозволяють вивчати поведінку конкретних користувачів.
- **Інтерес бізнесу:** онлайн-магазини зацікавлені підвищувати конверсію через персоналізовану рекламу та рекомендації для цільової аудиторії.

- Результат — рекомендаційні системи («покупці, які придбали цей товар, також купували...») та таргетована реклама, які й донині лежать в основі бізнес-моделі більшості онлайн-платформ.

17. Презентація великих мовних моделей

Ще раз застосовуємо ту саму логіку трьох факторів — і бачимо, що вперше вони зійшлися одночасно й у величезному масштабі:

- **Технічні можливості:** сучасні GPU дозволяють тренувати величезні моделі в рази швидше, ніж раніше.
- **Дані:** дані більше не потрібно спеціально шукати чи готувати — існує весь інтернет, що містить практично будь-яку інформацію.
- **Інтерес бізнесу** з'явився не одразу з появою технології, а після того, як змінилося сприйняття інструменту користувачами: генеративний ШІ міг би залишитися нішевою «прибудою для копірайтерів», але люди почали сприймати такі інструменти не як генератор тексту з гарною пам'яттю, а як живого співрозмовника, здатного підтримати діалог.
- Саме це масове сприйняття LLM як «співрозмовника» (а не просто інструмента) і спричинило вибухову хвилю інвестицій, досліджень і продуктів 2022–2026 років — той самий Google Trends графік, з якого почалася лекція (п. 1).

Місток до наступних лекцій курсу: далі ми детальніше розберемо, що конкретно вміють і чого не вміють сучасні LLM, а також — що найважливіше для цього курсу — як безпечно й відповідально працювати з даними (своїми й користувачів) під час використання таких інструментів, з організаційної та юридичної точки зору.

Контрольні питання для самоперевірки студентів

- Чому ChatGPT не є синонімом поняття «штучний інтелект»?
- Наведіть свій приклад системи, яка є ШІ, але не є машинним навчанням.
- Чим навчання з учителем відрізняється від навчання без учителя? Наведіть по одному прикладу з економіки/бізнесу для кожного.
- Поясніть на власному прикладі різницю між помилками I та II роду і чому «ціна» цих помилок буває різною.
- Назвіть три фактори, які визначають, чи «вистрелить» технологія ШІ на ринку, і проілюструйте їх на прикладі експертних систем 80-х.
- Чому шлях ШІ до масової популярності зайняв майже 80 років, а не стався одразу?