

Міністерство освіти і науки України
Національний університет «Києво-Могилянська академія»
Факультет соціальних наук та соціальних технологій
Кафедра соціології

Кваліфікаційна робота

освітній ступінь — бакалавр

на тему: **«ВПРОВАДЖЕННЯ ШТУЧНОГО ІНТЕЛЕКТУ ЯК
СКЛАДОВА НЕРІВНОСТІ НА РИНКУ ПРАЦІ УКРАЇНИ»**

Виконав: студент 4-го року навчання

Спеціальності 054 Соціологія

Лесик Данило Максимович

Науковий керівник: Артикуца С.С.,

ст. викладач кафедри соціології НаУКМА

Рецензентка: Прохорова А. С.

Кваліфікаційна робота захищена

з оцінкою _____

Секретар ЕК: _____

« ___ » _____ 2024 р.

Київ – 2024

ЗМІСТ

ВСТУП	3
РОЗДІЛ I	7
ТЕОРЕТИЧНІ ЗАСАДИ ДОСЛІДЖЕННЯ ШТУЧНОГО ІНТЕЛЕКТУ, АЛГОРИТМІВ ТА ПОНЯТТЯ НЕРІВНОСТІ НА РИНКУ ПРАЦІ	7
1.1. Поняття штучного інтелекту у наукових дослідженнях на тему нерівності	7
1.2. Поняття нерівності та підходи до вивчення нерівності на ринку праці	11
1.3. Алгоритми штучного інтелекту як засіб відтворення нерівності на ринку праці. Механізми дії та імплементації	17
РОЗДІЛ 2	23
ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ ЯК СКЛАДОВА	23
НЕРІВНОСТІ НА РИНКУ ПРАЦІ УКРАЇНИ	23
2.1. Глобальний вплив штучного інтелекту на трансформацію ринку праці	23
2.2. Поточний статус ринку праці України. Ситуація з нерівністю	26
2.3. Впровадження алгоритмів штучного інтелекту у контексті відтворення нерівності	28
2.4. Шляхи мінімізації негативного впливу штучного інтелекту на ринку праці України. Негативні наслідки використання штучного інтелекту на ринку праці	33
ВИСНОВКИ	37
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	38
ДОДАТОК А	42
Транскрипт інтерв'ю з Майком Зайко	42

ВСТУП

Розробка протомоделей штучного інтелекту та запровадження дискусійної платформи з даної тематики сягають 60-х років ХХ століття. Вивчення, моделювання та робота над прогресією штучного інтелекту є неперервним процесом, що триває й донині. Однак, варто зазначити, що активний розвиток та подальша доволі значна й фундаментальна успішність методу глибинного навчання (Deep Learning) призвела до покинутості інших методів машинного навчання, а найголовніше – до вагомого збільшення інтересу та залучення інвестицій до сфери штучного інтелекту.

До прикладу, згідно з даними наукового звіту ЮНЕСКО, протягом 2015-2019 років частка досліджень на тематику машинного навчання зросла на 50% (виміряно за кількістю публікацій). Також вагомим показником зростання інтересу й всесвітнього охоплення штучного інтелекту є оцінка Zion Market Research щодо об'єму ринку штучного інтелекту. Так, згідно з даними, у 2021 році всесвітній ринок штучного інтелекту оцінювався у 59.67 мільярдів доларів, натомість за оцінками очікується зростання вартості до 422.37 мільярдів до 2028 року. Німецька компанія Statista, що спеціалізується на зборі великих масивів даних та їхній візуалізації, у свою чергу, зазначає, що ринок технологій штучного інтелекту до 2030 зросте до орієнтовно 1,8 трильйонів доларів, паралельно оцінюючи нинішнє становище ринку у майже 300 мільярдів доларів США.

Без сумніву, настільки швидка популяризація та розвиток технологій штучного інтелекту сприяє позитивним змінам й на ринку праці: автоматизації значної кількості процесів із відповідним полегшенням роботи працівників та покращенням продуктивності; генерації нових робочих місць в силу імплементації штучного інтелекту, як складової наукового прогресу в цілому тощо. Крісталіна Георгієва – директорка-розпорядниця Міжнародного валютного фонду у звіті за

січень 2024 року зазначає, що згідно з останньою аналітичною роботою, близько 40% працевлаштованих у світі так чи так пов'язані як з ризиками, так і з позитивними наслідками імплементації штучного інтелекту. Причому, більш розвинуті країни мають як в цілому більшу частку робочих місць, що вразливі до впливу штучного інтелекту (близько 60%), так і більшу (приблизно половина з 60%) ймовірність бути доповненими імплементаційним штучним інтелектом, що безумовно є позитивним аспектом та підвищить продуктивність. Одночасно із цим, менш розвинені країни, як от країни, що розвиваються, матимуть загалом нижчий очікуваний вплив імплементації штучного інтелекту (близько 40%). Георгієва зазначає, що це свідчить як про потенційне стикання з меншим обсягом потрясінь через впровадження штучного інтелекту, так і про відсутність достатньої кількості необхідної інфраструктури та кваліфікованих фахівців задля ефективного використання переваг штучного інтелекту, що підвищує ризик збільшення нерівності між країнами з часом.

Водночас, зі стрімким зростанням використання та імплементації штучного інтелекту, зростає кількість дискусій, що спрямовані на обговорення етичності, загроз та негативних наслідків, що можуть повпливати на економіку, соціальне становище та, не в останню чергу, ринок праці. Поточна ситуація, що стосується імплементації штучного інтелекту, диктує умови, що ведуть за собою потенційне та фактичне заміщення професій шляхом автоматизації, продукування нерівності на підґрунті доходу, акумуляцію ресурсів та набуття статусу шляхом володіння інструментами імплементації штучного інтелекту та продукування упередженості алгоритмами глибинного навчання.

Так, Рудра Тіварі у своєму дослідженні зазначає, що імплементація штучного інтелекту, автоматизація та роботизація призводили до дехуманізації професій, зниження заробітних плат та збільшення нерівності прибутку. Аджемоглу та Рестрепо, власне, надають докази даному тезису на американському ринку праці.

Майк Зайко ж, застерігає зосередитися на питанні концентрації «влади штучного інтелекту» в руках кількох великих корпорацій, що складаються, переважно з білих чоловіків-розробників програмного забезпечення Кремнієвої долини. Застереження вживається у контексті продукування потенційних упереджень (bias) при написанні алгоритмів. Джуді Вайкман додає, що новітні технології у даному контексті створюють «не менше роботи, натомість більше гірших робочих місць».

Що стосується України, дослідження вказують на певну відстороненість та навіть необізнаність українців щодо штучного інтелекту. Згідно з дослідженням Центру Разумкова за червень 2023 року, 33,9% опитаних заявили, що не знають що таке штучний інтелект, 42% відповіли, що приблизно уявляють і лише 24% виявили впевненість у своїх знаннях стосовно штучного інтелекту. Щодо текстового штучного інтелекту, що використовує глибинне навчання (як-от ChatGPT), переважна більшість опитаних заявила, що не користуються технологією зовсім, 36,5% – що користуються (з яких лише 8,6% використовують текстовий штучний інтелект у своїй роботі), 8,7% заявили, що не знають що це таке. В принципі, це приблизно підтверджує тезис дослідження Міжнародного валютного фонду щодо ризиків залучення штучного інтелекту до робочого процесу в країнах, що розвиваються.

Загалом, серед українського суспільства наявна тенденція до наголошення позитивних аспектів імплементації штучного інтелекту, у парадигмі наукових робіт/статей та публікацій переважає позитивна конотація та акцентуація корисності зокрема й глибинного навчання. Щодо виокремлень проблем продукування нерівності або наявних упереджень – дискусійна база значно менша й повноцінно не розкрита.

Дана робота актуальна і корисна перш за все з точки зору новизни тематики штучного інтелекту і в особливості поняття нерівності крізь призму імплементації

штучного інтелекту на ринку праці України. Також, варто зазначити про малу кількість досліджень й публікацій, що охоплюють дану конотацію тематики імплементації штучного інтелекту. Робота може бути використана, як превентивний базис до запобігання створення нових проблем, що можуть бути пов'язані з імплементацією штучного інтелекту на ринку праці України.

Об'єктом дослідження є штучний інтелект та його використання на ринку праці України, **предметом** є впровадження штучного інтелекту як складова нерівності на ринку праці України.

Метою є визначити **наявні та потенційні** прояви нерівності, у **контексті** впровадження штучного інтелекту на ринку праці України

Завдання дослідження:

- Концептуалізувати поняття штучного інтелекту у контексті відтворення нерівностей
- Концептуалізувати поняття алгоритмів штучного інтелекту
- Окреслити підходи до вивчення нерівності на ринку праці
- Опрацювати наукові дослідження на тему штучного інтелекту та відтворення нерівності
- Визначити шляхи подолання нерівності у контексті імплементації штучного інтелекту на ринку праці України
- Визначити подальший вплив штучного інтелекту на ринок праці України

РОЗДІЛ І.

ТЕОРЕТИЧНІ ЗАСАДИ ДОСЛІДЖЕННЯ ШТУЧНОГО ІНТЕЛЕКТУ, АЛГОРИТМІВ ТА ПОНЯТТЯ НЕРІВНОСТІ НА РИНКУ ПРАЦІ

1.1. Поняття штучного інтелекту у наукових дослідженнях на тему нерівності

Перш за все, на мою думку варто зазначити й обґрунтувати використання терміну «штучний інтелект» у даній роботі. Дані дії продиктовані тим, що в переважній більшості у даному дослідженні йтиметься саме про нейронні мережі, що базуються на субметоді машинного навчання (machine learning) – глибинному навчанні (deep learning). Зрозуміло, що у патернах відтворення нерівності та залученості до ринку праці беруть участь не лише моделі, що базуються на глибинному навчанні, проте основний наголос у роботі буде саме на них. Проте, переважна кількість науковців, що специфікують напрямок своїх досліджень на нерівність та штучний інтелект, у роботах обирають саме останній термін «Artificial intelligence/AI», подекуди узагальнюючи, а подекуди спрощуючи задля кращого розуміння.

Отже, почати із, власне, визначення терміну **штучного інтелекту**, котрий застосовується у дослідженнях на тему нерівності, виглядає найбільш доцільним. Зробити це варто головним чином задля кращого усвідомлення суті та головних складових проблеми нерівності на ринку праці, ключовою ціллю є розуміння суб'єкту апеляції науковців у дослідженнях щодо штучного інтелекту та нерівності.

Головна проблематика полягає у тому, що остаточного і доконаного визначення поняття «штучний інтелект» не існує, існують різноманітні інтерпретації різноманітних науковців, котрі мають високий рівень схожості проте

різняться окремими деталями (West & Allen, 2018, ch.1). Цим фактом зокрема й можна пояснити варіативність та, де-факто, невідповідність використання терміну «штучний інтелект» у різних роботах з даної тематики. Проте, найбільш уніфікованим й влучним на мою думку визначенням штучного інтелекту може бути «машини, які відповідають на стимуляцію відповідно до традиційних реакцій людей, з урахуванням людської здатності до роздумів, суджень і намірів» (Shubhendu & Vijay, 2013, p. 28). Тобто штучний інтелект – це по суті здатність комп'ютерних систем виконувати певні задачі, які зазвичай вимагають сприйняття й мислення людського рівня. Ба більше, подібно до механізму функціонування людського мозку штучний інтелект також здатний адаптивно підходити до вирішення проблеми, виявляти гнучкість у процесі вирішення. Вест (West) та Аллен (Allen) виокремлюють дану ознаку штучного інтелекту, підкреслюючи важливість побудови моделі виконання та оцінки ситуації в реальному часі (2018). Науковці наводять приклад штучного інтелекту, імplementованого до напівавтономних та автономних транспортних засобів, де адаптаційні ознаки штучного інтелекту дозволяють сповіщати водіїв про затори попереду, ями та вибоїни, переносити адаптований досвід на інші транспортні засоби з імplementованим штучним інтелектом або ж навіть повністю переймати контроль над керуванням транспортного засобу (у випадку автономних транспортних засобів) завдяки сенсорам алгоритмам та іншим інструментам.

У більш прикладному сенсі можна у якості прикладу адаптивності штучного інтелекту навести системи рекомендацій таких компаній, як Meta (Facebook, Instagram), Google та Amazon. У даному випадку система спостерігає перші інтеракції користувача з платформою й на базисі зібраних даних інтеракцій та показників формує вибірку контенту, що може зацікавити користувача. Чим більше інтеракцій, тим більш деталізованою стає модель.

Поруч із адаптивністю дуже важливо згадати й **інтенційність** (Intentionality) – здатність та намір миттєво аналізувати та приймати рішення базуючись на даних реального часу, та діяти на підґрунті даного миттєвого аналізу (West & Allen, 2018, ch.1).

Іншим терміном, навколо тлумачення котрого варто зосередити увагу є **глибинне навчання** (deep learning), оскільки саме на цьому понятті найчастіше зосереджуються дослідники у контексті опису процесів імплементації штучного інтелекту та нерівності. Особливість та цікавість даного терміну полягає у тому, що це також елемент штучного інтелекту, його підвид, як вже було зазначено раніше. Проте, існує чітка ієрархічна модель, яку найкраще можна зобразити шляхом побудови діаграми Ейлера (або ж кіл Ейлера), де головною множиною виступає штучний інтелект, підмножиною головної множини є машинне навчання (machine learning), а підмножиною підмножини – глибинне навчання (Darwish et al., 2019, p. 1768). Тобто глибинне навчання постає підвидом машинного навчання, що у свою чергу постає підвидом штучного інтелекту. Можна також сприймати глибинне навчання як покращену модель машинного навчання. Якщо ж спробувати уніфікувати інформацію щодо глибинного навчання у конкретне визначення, справи насправді постають краще ніж з визначенням штучного інтелекту. **Глибинне навчання** – це підвид машинного навчання, що використовує нейронні мережі для аналізу великих масивів даних (big data). Норвіг (Norvig) у своїй книзі додає, що частка «глибинне», по суті означає що “схеми зазвичай організовані у багатoshарові структури, що означає, що шляхи обчислень від входів до виходів мають багато кроків” (2021, p. 801). Варто додати аби прояснити тлумачення основного терміну, що під нейронними мережами мається на увазі модель, що наслідує патерни поведінки та механізми роботи нейронної системи людини – нейронні мережі «кормлять» величезними масивами тренувальних даних аби вони на базисі цих масивів відслідкували патерни та відтворили у кінцевому результаті (Hardesty, 2017).

Далі я маю намір прояснити поняття **машинного навчання**, оскільки даний термін виступає галуззю, до котрої входить глибинне навчання (що є базисом дослідження) і розуміння ознак зазначеної категорії штучного інтелекту може надати краще розуміння його підвиду. Хоч глибинне і машинне навчання доволі схожі види штучного інтелекту, проте все ж є чітка ієрархія та відповідна інклюзія поміж ними.

Машинне навчання – це метод, за допомогою якого комп'ютери здобувають нові знання через обробку великих обсягів даних. Чим більше даних, тим точнішими стають результати. Великі дані (big data) – це величезні масиви інформації, які компанії збирають від користувачів. Ці дані приховують багато корисної інформації про поведінку користувачів. Алгоритми машинного навчання створені для витягання цієї інформації. Вони аналізують великі обсяги даних, навчаються на них і надають рекомендації на основі цього аналізу. (Kanade, 2022)

Так, дійсно, ознайомившись із концепціями машинного та глибинного навчання у індивіда може виникнути сумнів про існування відмінності в принципі. Проте, ключова різниця все ж присутня, вона полягає у тому, що хоч і глибинне навчання це також і машинне навчання (але у зворотній бік це не завжди може бути так) – як зазначає Анірудх (Anirudh), глибинне навчання – це більш спеціалізована форма машинного навчання, яка імітує архітектуру людського мозку для навчання на даних. Структури, такі як штучні та згорткові нейронні мережі, створені для відтворення патернів нейронів і зв'язків між ними у цифровому форматі. (2022)

Глибинне навчання та нейронні мережі є категорією машинного навчання, яка використовує цей специфічний метод. Як і у випадку з відношенням між машинним навчанням і штучним інтелектом, всі методи глибинного навчання є частиною машинного навчання, але не всі моделі машинного навчання використовують техніки глибинного навчання (там само).

1.2. Поняття нерівності та підходи до вивчення нерівності на ринку праці

Нерівність та зокрема нерівність на ринку праці є широко поширеною тематикою, до обговорення та вивчення котрої зверталось безліч науковців протягом усієї історії соціології. Якщо виділяти найбільш значущі та відомі постаті у контексті вивчення нерівності, список виявиться не надто коротким. До прикладу, Карл Маркс, будучи однією із найбільш важливих фігур для соціології та філософії 19 століття, революціонував наше розуміння суспільства, економіки та політики. Одна з його основоположних праць «Капітал», критикує капіталізм та притаманну йому нерівність. Маркс стверджував, що суспільство фундаментально розділене на два класи: буржуазію, яка володіє засобами виробництва, і пролетаріат, який продає свою працю. Він стверджував, що капіталістична система експлуатує робітничий клас, що призводить до глибокої економічної та соціальної нерівності. Теорія історичного матеріалізму Маркса припускає, що економічні чинники ведуть до суспільних змін і що остаточне повалення капіталізму призведе до безкласового комуністичного суспільства (Marx, 1976). Окремі аспекти та елементи марксового *modus operandi* були розглянуті у наступних частинах даного підрозділу, тож поки лише варто зазначити, що Маркс привертає особливу увагу зокрема й тим, що пише якраз про нерівність на ринку праці у тому числі.

У свою чергу, сучасник Маркса Макс Вебер презентував більш складний аналіз соціальної стратифікації. У своїх книгах «Економіка та суспільство» та «Протестантська етика та дух капіталізму» Вебер розширив концепцію нерівності, охопивши та додавши статус і владу на додаток до економічного класу. Веберівська «тріада» комплектуючих соціальної стратифікації додає більше деталей та піднімає питання інтерсекційності нерівності.

Насамкінець, цікавим прикладом може слугувати доробок Жана Бодріяра. Звісно, на перший погляд постать французького соціолога не надто узгоджується із тематикою нерівності: все ж, в загальному сенсі, Бодріяр писав дещо про інакше. Проте, його дослідження у контексті сучасного споживацького суспільства дають цінні знахідки щодо понять соціальної стратифікації та нерівності. У своїй книзі «Суспільство споживання: міфи та структури» Бодріяр критикує ілюзію рівності, яку фіктивно культивує й підтримує культура споживання. Він стверджує, що розповсюдження товарів і споживання не призводить до суспільної нівеляції, як дехто може вважати. Натомість Бодріяр (Baudrillard) припускає, що споживацтво функціонує для зміцнення соціальних відмінностей і підтримки існуючих структур влади (1998). До того ж, відомі та популяризовані широкому загалу «Симулякри та симуляції» знову-таки хоч і не мають предметом дослідження нерівність, а проте можуть виступати у якості допоміжного інструменту осмислення світогляду науковця стосовно останньої. Ба більше, так само як у Маркса, поміж рядків текстів Бодріяра можна виокремити окремі елементи нерівності саме на ринку праці. Наприклад, якщо пригадати один з ключових термінів книги «Симулякри та симуляції» - гіперреальність, можна вдало спроектувати трактування ни нинішню модель ринку праці. Бодріяр (Baudrillard) зазначає, що у постмодерністському сучасному світі реальність замінена на гіперреальність – симульована дійсність, сповнена симулякрів, що рухаються у бік підміни істинної реальності, стирає межі між істинним і знаками. Цю концепцію можна пов'язати з нерівністю, оскільки гіперреальне часто маскує справжні умови соціальної нерівності. Засоби масової інформації та споживча культура створюють світ, де проєктується ілюзія рівності та можливостей, приховуючи приховану нерівність. Або ж зв'язок із нерівністю на ринку праці можна віднайти у тому сенсі, що репрезентація роботи, кар'єри та успіху в засобах масової інформації та культурі часто приховує реальні умови праці. Наприклад, засоби масової інформації можуть зображати у своїх розміщених вакансіях (що певною мірою можна вважати рекламою, це важливо у контексті

постулювання Бодріаром панування гіперреальності) певні професії як дуже бажані, не показуючи нестабільні умови праці чи нерівний доступ до цих робочих місць, таким чином маскуючи реальну нерівність на ринку праці.

Насправді, прикладів відомих соціологів, що фокусувалися на дослідженні нерівності є ще доволі значна кількість, я лише згадав найбільш популярні та найбільш релевантні до теми приклади. Надалі я хочу сконцентруватися та звузитись виключно до постатей Карла Маркса та Жана Бодріара, оскільки перш за все їхній академічний доробок стосується нерівності на ринку праці напряду чи дотично та тому що концепції даних науковців найкраще пояснюють продукування нерівності на ринку праці шляхом імплементації штучного інтелекту.

Отже, як відомо, у своїй праці «Капітал» Маркс часто звертається до тематики впровадження машин (техніки) у виробниче середовище. Він вбачає у даному явищі багато потенційної нерівності, зокрема він зазначає, що «знаряддя праці кладе робітника на лопатки. Прямий антагонізм між ними стає найбільш очевидним у моменті, коли нові машини вступають у конкуренцію з ручними ремеслами чи мануфактурами, що є передані з минулих часів» (Marx, 1976, p. 559). Маркс стверджує, що коли знаряддя праці стає машиною, воно негайно перетворюється на суперника робітника. Використання машини для збільшення капіталу призводить до скорочення робочих місць і руйнування умов існування робітників. Уся система капіталістичного виробництва базується на тому, що робітники продають свою робочу силу як товар. Поділ праці перетворює цю робочу силу на вузькоспеціалізовані навички, зосереджені на використанні певного інструменту. Коли ж виконання цих завдань переходить до машини, робітники втрачають свою споживну цінність і, відповідно, свою вартість на ринку праці. (Marx, 1976). Хоч погляди німецького соціолога є першочергово рефлексією періодів першої та другої промислової революції, що призвела до імплементації різного роду верстатів, двигунів та загальної автоматизації процесів, очевидно є аналогія із

імплементациєю штучного інтелекту нині, котрий є невід'ємною або ж навіть ключовою складовою четвертої промислової еволюції (Мба, 2023). Зараз імплементация штучного інтелекту викликає побоювання у заміні численної кількості робочих місць «нейромашинами», як зазначає Георгієва (Georgieva) – ці побоювання не є повністю безпідставними, і водночас можуть бути злегка перебільшеними (2024).

Маркс (Marx) також застерігав, що капіталістична світ-система призводить до концентрації влади й грошей у руках малих привілейованих (буржуазних) груп (1976). Нескладно провести аналогію й тут: нині спостерігається чітка тенденція до накопичення інструментів володіння системами штучного інтелекту в руках великих елітних корпорацій, часто в силу доволі великої вартості тренування моделей штучного інтелекту, що є під силу лише тим, хто здатний на згенерувати таку кількість надлишкового капіталу. І хоч розробники зазвичай вважають, що штучний інтелект видає істинну картину світу, що він є нейтральним і корисним, запламованим лише упередженнями з упередженого світу. Проте, як зазначає Каллурі (Kalluri), «в реальності байдужа система лише служить тим, хто має владу» (2020, р. 169). І дане намагання мімікрувати «жорстоку правду життя» насправді грає злий жарт стосовно штучного інтелекту: маргіналізовані групи оминаються, не включаються в модель, не виступають предметом дослідження у полі штучного інтелекту тощо.

Як вже було зазначено, Жан Бодріяр створив нове фундаментальне поняття. Він стверджував, що в сучасному суспільстві реальність замінюється симуляціями, які формують гіперреальність — стан, де симуляції здаються більш реальними, ніж сама реальність. Симуляції – є копією реального і можуть трансформуватися у симулякри Симулякри ж, у свою чергу, виступають певними посиленнями, імітацією реального без першоджерела імітування. Штучний інтелект у свою чергу виступає класичним прикладом бодріярівських симулякрів. За допомогою штучного

інтелекту, ми можемо імітувати реальних людей, висловлювання, текстові документи. Навіть наша особиста реальність зокрема в силу алгоритмів машинного та глибинного навчання медіапрограм та соціальних мереж якими ми користуємося не є істинною, а саме наповнена безліччю копій реальності, якої в реальності не існує. Наприклад, алгоритми рекомендацій з часом формують міцну бульбашку наших уподобань, за межі якої індивіду дуже складно вийти, оскільки вони мають усе що їм потрібно. Або ж, якщо взяти саме соціальні мережі, нормою є ділитися позитивними, успішними моментами з життя, що створює симулякр того, чого насправді не існує – створюється ілюзія радісного й вдалого, контрастного життя на противагу особистому досвіду користувача, який, оскільки є особистим, виступає абсолютно повним і полярним за спектром емоцій. Щодо нерівності на ринку праці, імплементація штучного інтелекту, зокрема його алгоритмів можуть продукувати нерівність. Враховуючи згадану паралель штучного інтелекту та симулякрів й симуляції, що формують гіперреальність, доречним буде згадати Бодріарову модель трансформації симуляції у симулякр. Він стверджує, що перехід відбувається у чотири етапи: «Ось які були б послідовні фази образу: він є відображенням глибокої реальності; він маскує і спотворює глибоку реальність; він маскує відсутність глибокої реальності; він не має жодного відношення до будь-якої реальності: він є симулякром.» (Baudrillard, 1994, p. 6).

Подальше пояснення участі штучного інтелекту як симулякру вдало здійснює Ніда (Nida, 2019). Вона наводить приклад моделей ймовірності рецидивізму злочинців, які широко використовуються в Сполучених Штатах для визначення термінів ув'язнення. Ніда розглядає розвиток алгоритму крізь призму чотирьох стадій перетворення симуляції у симулякр. Мета алгоритму - створити детальний профіль правопорушника, що дозволяє оцінити ймовірність його рецидиву. Намір є повністю об'єктивним і прагне надати точне відображення реальності, що є першою стадією (відображення глибокої реальності). На етапі розробки алгоритму, вимагається трансфер даних задля оцінки ризику (рецидиву). Коли відсутні

безпосередньо шукані дані, що трапляється досить часто, розробники використовують замітники. Це може включати використання непрямих даних, таких як рідне місто людини та рівень злочинності в ньому, для визначення ймовірності повторного правопорушення. В результаті люди з районів з високим рівнем злочинності автоматично вважаються високоризиковими. Інші замітники включають частоту зупинок та обшуків поліцією, ігноруючи той факт, що афроамериканці непропорційно більше піддаються таким законам як у Великій Британії, так і в США. Змішуючи кореляцію з каузальністю, алгоритм створює спотворену версію реальності (другий етап). Прикладом третього етапу є той факт, що алгоритм не просто прогнозує, а конструює такі умови, що неминуче приведуть до справдження прогнозу. Причиною цьому є петля позитивного зворотнього зв'язку, що змушує вважати успішною модель. Наприклад, скажімо що особа, котра очікує вироку проживає у районі з високим рівнем злочинності, що скоріш за все означатиме наявність у сім'ї родичів, що вже відбували покарання у місцях позбавлення волі. Це вважатиметься алгоритмом тригером до здійснення рівня ризику та ймовірності рецидиву особи, що очікує на вирок – відповідно термін буде оголошений якнайдовшим. Відбуваючи покарання, особа з кожним роком втрачатиме шанси на працевлаштування після виходу. Якщо ця особа у відчаї піде на повторне правопорушення – алгоритм перемаже, це вважатиметься успіхом, модель визнають надійною. Наостанок, для людини зі сторони, справжня (реальна) істина повністю прихована. Як можна побачити, незалежно від дійсної об'єктивної ймовірності рецидиву засудженої особи, симуляційна модель алгоритму у будь-якому випадку визначить рівень ризику високим. (він не має жодного відношення до будь-якої реальності: він є симулякром)

Отже, гіперреальність у даному контексті стала реальністю. Ми зараз ступаємо по світу симулякрів, що у свою чергу породжують нові симулякри, а ті у свою чергу – треті, усе завдяки штучному інтелекту. Межа між реальним та уявним стає ледь видимою, нечіткою й слабкою. Четверта індустріальна революція створила

математичні моделі й нейромережі, котрі активно роблять цю межу непомітною зовсім.

1.3. Алгоритми штучного інтелекту як засіб відтворення нерівності на ринку праці. Механізми дії та імплементації

У попередньому підрозділі було неодноразово згадано алгоритми штучного інтелекту, як такі, що є одним із способів відтворення нерівності. Дійсно, алгоритми моделей машинного чи глибинного навчання часто відтворюють нерівність в силу механізму тренування й побудови моделі, що насправді виступає серйозною проблемою (Zajko, 2022). Враховуючи наростаюче хвилювання й увагу до даної складової штучного інтелекту є доцільним пояснити та описати що узагалі являє собою алгоритм штучного інтелекту, який його механізм дії та звідси яким чином алгоритми штучного інтелекту можуть бути складовою нерівності.

Зрозуміло, що поняття алгоритму – не суто термін з сфери штучного інтелекту, натомість це поняття максимально широке у математичному контексті. Водночас з тим, як звужене трактування алгоритмів у контексті штучного інтелекту не означає, що усі алгоритми поза звуженням відповідають усім критеріям даного поняття, більш узагальнене поняття містить у собі все, чому відповідає звужене. Ситуація схожа із описаною ієрархією понять штучного інтелекту у підрозділі 1.1. Отже, якщо описати визначення узагальнено, алгоритми – це «набір інструкцій, яких слід дотримуватися під час обчислень або інших операцій» (Tableau Software, 2021, sec. What is an AI algorithm?). І як можна зрозуміти з попередніх концептуалізацій штучного інтелекту, машинного та глибинного навчання - це визначення дійсно підходить. Проте, якщо говорити конкретно щодо визначення алгоритмів штучного інтелекту, по суті, алгоритми отримують навчальні дані

(позначені чи не позначені, надані розробниками або зібрані автономно) і використовують їх для розвитку/тренування своїх можливостей. Вони виконують завдання, використовуючи ці дані як базис, від якого потрібно рухатися. Деякі алгоритми можуть самовдосконалюватися шляхом включення нових даних і вдосконалення свого підходу, тоді як інші потребують керівництва програміста, щоб покращити свою продуктивність (supervised/unsupervised) (Tableau Software, 2021). Задля повнішого розуміння варто поглибитись та визначити особливості supervised та unsupervised методів тренування алгоритмів штучних мереж, спираючись переважно на визначення, що надає Делуа (Delua) у своїй статті, оскільки джерело є максимально релевантним до тематики (2021, March 12).

Отже, кероване навчання використовує «позначені» (labeled) масиви даних. Вони позначені специфічними тегами задля подальшої точнішої інтерпретації алгоритмом та покращення результатів. Процес «позначання» відбувається людьми, спеціально навченими спеціалістами, тобто вимагає участі «HITL» (human-in-the-loop, людина-у-циклі). Це дозволяє моделі виокремлювати заздалегідь позначені потрібні змінні та проводити подальше тренування спираючись на них. Як зазначає Делуа, кероване навчання в загальному можна розподілити на дві категорії, що визначають поставлену задачу: класифікація та регресія (Delua, 2021).

Класифікаційні завдання використовують алгоритми для сортування тестових даних у визначені групи, як-от відрізнити котів від собак. На практиці контрольоване навчання може сортувати спам із скриньки вхідних листів. Загальні алгоритми для цих завдань включають лінійні класифікатори, метод опорних векторів, дерева рішень (decision trees) і випадкові ліси (random forests).

Регресія ж використовує алгоритми для дослідження зв'язків між залежними та незалежними змінними. Ці моделі передбачають числові результати на основі різних вхідних даних, наприклад, оцінки майбутніх продажів компанії.

Найпоширеніші методи регресії включають лінійну, логістичну та поліноміальну регресію. (Delua, 2021)

Щодо некерованого (unsupervised) навчання, усе протилежне керованому: даний тип використовує непозначені (unlabeled) масиви даних, аби провести подальший аналіз та кластеризацію. Алгоритми некерованого навчання головним чином спрямовані на виявлення прихованих патернів у даних, без потреби «керівника».

Основні цілі використання некерованих мереж – кластеризація, асоціація та зменшення розмірності. Кластеризація – один з методів глибинного аналізу даних (data mining), що бере непозначені масиви та кластеризує (групує) їх відповідно до схожостей або ж розбіжностей. Наприклад, відомий у аналізі даних метод k-середніх призначають схожі точки даних (data points) до спільних груп, при цьому k-value може розповісти про рівень деталізації та кількість кластерів. Яскравими прикладами цього методу є сегментація ринку та стиснення зображень. (Delua, 2021). Асоціація ж першочергово спрямована на виявлення зв'язків між змінними масиву, керуючись різними правилами. Даний метод широко використовується у аналізі онлайн-кошиків споживачів та механізмів рекомендацій по типу «користувачі, що купили цей товар також купують...». Якщо ж масив даних є занадто великим та має забагато ознак/функцій – використовують зменшення розмірності. Суть полягає у тому, аби зменшити кількість змінних, водночас залишивши розмір даних. Зазвичай зменшення розмірності застосовується на етапі попередньої обробки даних (препроцесінг), наприклад інструментами що зменшують кількість «шумів» зображення задля покращення якості картинки. (Delua, 2021)

По суті, різниця згаданих двох типів алгоритмів штучного інтелекту лежить у самих назвах: головна відмінність у типі використаних даних. Використання позначених даних означає людську інтервенцію у процес навчання та тренування

моделі, на противагу використання непозначених змушує алгоритми власноруч вивчати та відслідковувати тенденції непозначених даних. Кероване навчання зазвичай корисне точністю у порівнянні з некерованим, проте вимагає значних ресурсів на тренування моделі. Наприклад, модель може визначати тривалість шляху від точки А у точку Б, спираючись на зміни висоти над рівнем моря, типом пересування користувача, погодних умов тощо. Проте нічого з вищеперерахованого не є самозрозумілим для алгоритму, вдалого прогнозу без навчання тому, що, наприклад, пересування знизу догори зазвичай займає більше часу ніж згори донизу, не буде.

Врахувавши перераховані технічні особливості алгоритмів штучного інтелекту, можна перейти до роз'яснень щодо такого аспекту алгоритмів, як нерівність. Головним чином нерівність може виникати шляхом упередженості штучного інтелекту й зокрема алгоритмів. Постає питання: яким чином штучний інтелект чи алгоритми можуть стати упередженими? Відповідь криється у декількох речах, котрі з однаковою ймовірністю можуть опинитися причиною. Дана проблема може виникати через існуючу нерівність у реальному житті (або навіть у історичній перспективі, усе залежить від запиту), певні відповідні аспекти якого виступають тренувальними даними моделі. Наприклад, команда з даних та штучного інтелекту IBM (IBM Data and AI Team) зазначає, що «навчальні дані для алгоритму розпізнавання обличчя, який надмірно представляє білих людей, можуть створювати помилки під час спроби розпізнавання обличчя людей іншої раси». Також, упередження можуть ховатися у характері позначених даних: «Наприклад, інструменти рекрутингу штучного інтелекту, які використовують непослідовне позначення або виключають/надмірно представляють певні характеристики, можуть усунути достатньо кваліфікованих кандидатів на роботу з розгляду» (2023, October 6, sec. Training Data Bias).

Тобто, як стає зрозуміло, тренувальні дані та алгоритми доволі тісно пов'язані у питаннях продукування нерівності, що походить із логіки функціонування штучного інтелекту. Тож, якщо шукати причину виникнення упередженості саме в алгоритмах, використання неправильних навчальних даних може призвести до формування таких алгоритмів, котрі видають несправедливі результати, часті помилки і найголовніше – посилити упередження котре перейшло із масиву тренувальних даних. На алгоритм також може впливати й безпосередньо той, хто пише код. Наприклад, розробник може нерівномірно зважити певні аспекти в процесі прийняття рішень алгоритмом (decision-making) спираючись суто на власні суб'єктивні свідомі або ж несвідомі упередження. У результаті, алгоритм, наприклад, може помилково дискримінувати людей певної раси чи статі на основі таких характеристик, як словниковий запас чи гроші. (IBM Data and AI Team, October 6).

Наразі, помилки програмістів, невірність тренувальних даних, недбалість на етапі валідизації тощо вже призводить до негативних наслідків та скандалів. У контексті нерівності на ринку праці варто згадати кейс Амазону: використовуючи алгоритм нейронної мережі розпізнавання тексту резюме, працівники компанії таким чином автоматизували пошук кваліфікованих підходящих кандидатів. Проте, дана кампанія вилилась в упереджені рішення штучного інтелекту на базисі гендерної нерівності: алгоритм був тренований на резюме кандидатів Амазон, поданих протягом останніх 10 років, більшість з яких у силу нерівномірного гендерного розподілу технічної індустрії були чоловіками. Як наслідок, модель навчила сама себе хибному факту, що чоловіки – більш бажані кандидати; вона бракувала резюме, де були фрагменти зі словом «жіночий/жіноча» у будь-яких відмінках (як-от у словосполученні «капітан жіночого клубу шахів»); також модель відхилила випускниць двох жіночих коледжів. Врешті, Амазон розпустив усю команду розробників, а сам проєкт зупинили. (Dastin, 2018)

Отже, у даному розділі було розглянуто достатньо теоретичного підґрунтя, аби можна було рухатися у бік звуження тематики. Було концептуалізовано поняття штучного інтелекту, стисло розглянуто його потенційні позитивні та негативні сторони. Було наведено важливі ознаки штучного інтелекту – такі як інтенційність та адаптивність. Також було концептуалізовано вужчі поняття машинного та глибинного навчання, роз'яснено чітку ієрархію понять. Було концептуалізовано поняття big data, що є невід'ємною складовою навчання штучного інтелекту. Було виокремлено відмінності між глибинним та машинним навчанням, а також комплексність підходу до специфічності вживання терміну «штучний інтелект» для позначення окремих його складових, що є вкрай важливим для кращого розуміння контексту наукових досліджень з даної тематики. Далі у розділі було розглянуто різноманітні підходи до вивчення нерівності на ринку праці. Було згадано праці та ключові тези таких класиків, як Маркс, Вебер та Бодріяр, із подальшим рішенням звужити теоретичну рамку до теорій Маркса та Бодріяра, оскільки вони є дуже цікавими та релевантними як у контексті нерівності на ринку праці, так і у контексті імплементації штучного інтелекту. Наостанок, було концептуалізовано поняття алгоритмів штучного інтелекту та їхньої ролі у продукуванні нерівності. Вищеперераховані дії дають змогу повноцінно спиратися на описані поняття й теорії при переході до наступного розділу.

РОЗДІЛ 2

ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ ЯК СКЛАДОВА НЕРІВНОСТІ НА РИНКУ ПРАЦІ УКРАЇНИ

2.1. Глобальний вплив штучного інтелекту на трансформацію ринку праці

У даному підрозділі детальніше та вужче розглядається штучний інтелект та його імплементація на ринку праці. Мається на увазі, що тут відсутній термінологічний зріз, натомість присутні описові твердження, сформовані шляхом огляду статей та ресурсів. До того ж, в якості опори та підживлення знахідок у літературі за допомогою програми Zoom було проведене неструктуроване глибинне інтерв'ю у викладача Факультету мистецтв і соціальних наук University of British Columbia та PhD соціології Майком Зайко. До його доробку зокрема входить стаття «Artificial intelligence, algorithms, and social inequality: Sociological contributions to contemporary debates», яка також використовувалася у даній роботі.

Отже, як можна зрозуміти, штучний інтелект наразі є вкрай широко репрезентований у багатьох сферах життя. Не виключенням є й ринок праці: штучний інтелект як інструмент використовують у процесі наймання на роботу, пошуку кандидатів на конкретну посаду, автоматизації деяких завдань вже існуючих позицій у компанії. Щодо останнього варто зупинитися детальніше. Зрозуміло, що автоматизація може сприйматися позитивно у контексті розвантаження працівників, даючи простір для більш креативних задач, водночас перебираючи на себе рутинну роботу. Або ж підвищення продуктивності шляхом виконання задач, що раніше були людськими зі значно більшою швидкістю та ефективністю (R. Russell, 2023). Проте доречно було б замислитися щодо того, чи потрібні людські працівники, якщо

штучний інтелект настільки успішно справляється із завданнями. З цієї відправної точки у суспільстві й починаються сумніви й побоювання про невтішне майбутнє, у якому штучний інтелект виступає загрозою людського ринку праці у сенсі перебирання значної кількості позицій на себе. Варто одразу зазначити, що дані побоювання є хибними. Як вже зазначалося раніше, дійсно, певна частка професій може бути замінена штучним інтелектом, проте ця частка не є великою (Georgieva, 2024). Як відомо, імплементація штучного інтелекту, з іншого боку призводить до створення нових робочих місць, адже моделі потрібно обслуговувати й тренувати, а дані – аналізувати й оцінювати. Майк Зайко притримується тієї ж думки щодо потенційних загроз та негативного глобального впливу штучного інтелекту на ринок праці, також зазначаючи про існування проблеми заміщення професій та водночас наголошуючи на очевидному перебільшенні ситуації. Також, науковець зазначав на протиположності побоюванням, що загалом штучний інтелект не забирає, натомість змінює тип роботи, зміщує фокус на інші завдання:

«...claims about robots or AI's taking away 50% of jobs or making lots of industries obsolete, you know, those sorts of claims have so far largely been exaggerated. It does seem, though, that there is an impact on, you know, what you call labor market or on different fields of employment. And some jobs are being affected. And the ones that are sort of the most easy to automate are the ones that are primarily being affected, changed, eliminated in some cases, but mostly it's changing the type of work that the people are doing.» (тут і далі у цьому підрозділі – Майк Зайко)

Якщо розглянути, де має місце імплементація штучного інтелекту на ринку праці, окрім виконання класичних задач під конкретні (та контекстуально унікальні) вимоги конкретних окремих компаній, однією зі спільних сфер імплементації штучного інтелекту на ринку праці постає HR, що містить у собі процеси прийому на роботу працівників, оцінку резюме кандидатів, текстова комунікація з кандидатами – завдання із якими компанії можуть призначати працювати штучний

інтелект. Потенційно, у сфері HR може відбуватися доволі багато випадків появи упереджень чи нерівності як результату, в силу тих чи тих рішень штучного інтелекту з різних причин:

«a lot of the concerns have been around hiring. So when it comes to sort of inequality, things that get labeled bias, I think that's where a lot of the concerns have rightfully been because the hiring process is being automated in a lot of places and that determines who gets what jobs»

Тобто наразі, поміж позитивного глобального впливу на ринок праці, штучний інтелект має й багато аспектів негативного впливу, де сфера HR виступає суттєвою складовою. Зазвичай, коли мова йде про певні випадки упередженості штучного інтелекту та внаслідок утворення нерівності стосовно елементів, за які відповідає HR (наймання, резюме, листи кандидатам тощо), трапляється це із великими компаніями, оскільки крізь них постійно проходить величезний обіг листів/заявок/аплікацій/кандидатів, який потрібно автоматизувати. Можна згадати приклад Амазону з минулого розділу, який є яскравою репрезентацією негативного впливу штучного інтелекту. Тож проблема саме у великій кількості автоматизованих процесів, адже штучний інтелект повинен власноруч визначати на перших етапах комунікації із кандидатами, чи варті їхні навички зустрічі з реальними людьми на наступних ітераціях. Майк Зайко пояснює це наступним чином:

«...primarily by large employers who have big HR operations... the danger is that the sort of pattern they're going to be trained on is going to be some sort of historical set of inequalities that then becomes perpetuated. ...it's very easy for one of these pattern recognition systems to pick up on lots of proxies, lots of indicators that someone is a woman or someone is an ethnic minority...»

Як і було згадано раніше, настільки масштабні масиви даних, з котрими працює штучний інтелект, враховуючи специфіку упереджень у тренувальних

даних, можуть змушувати модель не виправдано дискримінувати певні групи етносу, гендерні групи та інші за аналогією.

2.2. Поточний статус ринку праці України. Ситуація з нерівністю

У даному підрозділі варто зосередитися на ситуації із нерівністю на ринку праці України, аби зрозуміти потенційний вплив імплементації штучного інтелекту на її відтворення, на формування несправедливих упереджень тощо.

Почати потрібно із гендерної нерівності на ринку праці України, оскільки тема гендерної нерівності наразі є актуальною і активно обговорюються шляхи її подолання. Згідно з Всесвітнім звітом щодо гендерного розриву (Global Gender Gap Report), Україна посіла 66-е місце у загальному рейтингу зі 146 країн (2023). Розрив складав 29%, що лише на 0,7% краще ніж минулого звіту. Тобто можна стверджувати, що гендерна нерівність все ще виступає суттєвою проблемою.

Якщо звертатися до уточнення ситуації з гендерною нерівністю на ринку праці України, варто зазначити різницю у заробітній платні чоловіків та жінок. Далі я керуватимуся даними зі статті Національного інституту стратегічних досліджень, де висвітлені дані щодо гендерної нерівності на ринку праці України (2023, June 2). Отже, згідно з даними, у 2021 р. жінки отримували середньомісячну заробітну платню на 18,4 % нижчу, ніж відповідна заробітна платня у чоловіків (13802 грн – у жінок і 16912 грн – у чоловіків). Враховуючи динаміку ідентичних показників, можна стверджувати, що тенденція вищої зарплатні чоловіків зберігається, особливо якщо додати результати Global Gender Gap Report. Ще однією важливою проблемою гендерної нерівності на ринку праці є безробіття жінок. Станом на 01 квітня 2023 р., мали статус офіційно зареєстрованого безробітного 137,3 тис осіб. З них чоловіки становили 28 %, жінки – 72 % (НІСД, 2023). Насправді в умовах

повномасштабного вторгнення поясненням такої різкої відмінності між показниками безробіття може бути офіційність даних. Оскільки аби отримати статус й, відповідно, допомогу по безробіттю потрібно стати на військовий облік. Тож чоловіки зазвичай самостійно шукають роботу, почасти й неофіційну. До того ж, як зазначає НІСД, чоловіки, які проходять військову службу (на умовах строкової служби, мобілізації або контракту), вважаються задіяними в економічній діяльності країни. (2023).

Інший вид нерівності на ринку праці, про який варто зазначити – етнічний, оскільки як можна згадати із минулих підрозділів роботи, етнічні групи також можуть піддаватися упередженням з боку алгоритмів штучного інтелекту. У намаганнях зрозуміти становище із етнічною нерівністю на ринку праці України, варто зосередитись на конкретній маргіналізованій етнічній групі – на ромів. Звісно, дана національна меншина України доволі часто стикається з дискримінацією, стигматизацією та упередженнями з боку оточення. Проте яке становище у даної групи наразі, під час повномасштабного вторгнення? Згідно з оцінками, до повномасштабного вторгнення в Україні проживало близько 200-400 тисяч ромів, близько 100 тисяч стали біженцями у інших країнах після 24 лютого 2022 року. (Тужанський, 2022) Попри те, зіткнувшись із дискримінацією та неповагою до себе з боку населення європейських країн, більшість ромів була змушена повернутися на територію України (Ellena & Makszimov, 2022). На своїх рідних землях дана етнічна група також нерідко стикається із випадками дискримінації. Згідно з дослідженням КМІС, роми є однією з найменш бажаних етнічних груп на території України у контексті близькості соціальної дистанції (2022). Щодо ринку праці, багато ромів втратили роботу під час повномасштабного вторгнення, оскільки більшість з них постраждали від зупинки роботи підприємств (Уповноважений Верховної Ради України з прав людини, 2022). Тобто певна частка національної меншини вже не є складовою працевлаштованого населення, а якщо взяти до уваги той факт, що приблизно 10-20% з орієнтовних 400 тисяч ромів, що

проживають в Україні не мають паспортів (European Network on Statelessness, 2022) – ситуація складається така, що відносно загальної кількості працевлаштованого населення України, відсоток працевлаштованих ромів може бути близький до нуля. Відповідно, падає до критично низьких показників й імовірність, що у випадково взятій компанії працюватимуть представники ромської етнічної групи. Даний логічний ряд будувався з метою наголошення на принципах роботи алгоритмів штучного інтелекту при імплементації останнього у сферу HR. Для зручності можна знову провести аналогію з Амазон: чим менше у тренувальних даних буде представників тої чи тої категорії – тим більший ризик формування алгоритмами упередження та визначення домінуючої в даних категорії єдиною правильною. Є зрозумілим і факт наявності етнічної нерівності на ринку праці України, тут грає роль інтерсекційність сфер нерівності етнічних меншин – через дискримінацію у інших сферах життя, етнічні меншини не можуть бути повноцінно представлені на ринку праці України.

2.3. Впровадження алгоритмів штучного інтелекту у контексті відтворення нерівності

Оскільки у першому розділі було детально концептуалізовано та типізовано поняття алгоритмів штучного інтелекту й разом з цим пояснено яким чином можуть формуватися упередження й продукуватися нерівність – у даному підрозділі мова піде про дещо інше. Видається доцільним описати процес створення алгоритмів, а також яких дій можна вжити аби змінити потенційно упереджений алгоритм таким чином утилізувавши нерівність. У даному розділі я знову звертатимуся до інтерв'ю з Майком Зайко, оскільки протягом розмови ми зачепили релевантні до цього підрозділу питання щодо алгоритмів штучного інтелекту. Також, я буду спиратися на проведене напівструктуроване глибинне інтерв'ю із Антоном С. – інженером

машинного навчання у одній з українських компаній, бакалавром прикладної математики КНУ, магістром у галузі наук про дані та моделювання складних систем. Даного респондента було обрано в силу практичної обізнаності у темі штучного інтелекту (зокрема, машинного навчання) та позиціонування респондента з боку «творців» алгоритмів.

Перш за все, варто почати із трактування процесу створення алгоритмів. Алгоритми використовуються для створення моделі глибинного чи машинного навчання. В залежності від типу алгоритмів (supervised/unsupervised) вони або навчаються з даних, або ж поміщаються в дані, про які нічого не відомо – усе для побудови якісної моделі (Brownlee, 2020). Тобто варто розуміти, що у буквальному сенсі «створити» алгоритм не можна, це вже існуючий цикл, який потрібно «нагодувати» відповідними до задачі даними. Антон в інтерв'ю наголошує на даній специфіці, пояснюючи той факт, що алгоритми розділяються на різноманітні види, кожен з яких підходить до окремої задачі. У інтересах програміста правильно вибрати алгоритм (його вид) та надати коректні дані:

«...алгоритми, вони не будуються, ну, їх немає. Тобто, вони вже є, їх просто використовують. ... Так, ну, от те, що я все назвав, воно вже існує і його вже використовують.»

Інтерв'юер: *Тобто, це шаблони?*

Антон С.: *Ні, це не шаблони, це реальне, те, що я тільки що проговорив, це якраз от реальні алгоритми, які будуються. Ну, не алгоритми, а методи більше. Але вони під собою мають якраз уже навіть не шаблони, а метод повністю реалізований, який ми можемо просто взяти і використати» (Антон С.)*

Тобто по суті – створюється модель шляхом тренування існуючого алгоритму конкретного виду. Адже алгоритми штучного інтелекту це похідні математичних

алгоритмів, які існують відносно давно. Наприклад, алгоритм кластеризації методом k-середніх був винайдений ще у середині ХХго століття.

Види алгоритмів згадуються у підрозділі 1.3., я лише нагадаю назви деяких: лінійна регресія, логістична регресія, дерево рішень, k-найближчі сусіди, k-середні.

Отже, тепер коли поняття створення алгоритмів було остаточно розтлумачене, можна перейти до розгляду питання упереджень, котрих може набути алгоритм і звідси що можна в діяти аби їх позбутися. Припустімо, що наш алгоритм під час тренування виокремив хибну закономірність, що в подальшому вилилась упередженістю моделі та створенням нерівних умов. Постає логічне запитання: хто ж винен? Можливо, той хто займався тренуванням алгоритму (якщо це supervised)? А може винен той, хто недостатньо оцінив тренувальні дані? Чи ніхто не винен – адже іноді дуже складно зрозуміти за якою логікою конкретний алгоритм дійшов до упередженого висновку, адже *«проблема в тому, що часто ці системи є “чорними скриньками”.* *Ми не знаємо, як вони працюють. Ми не бачимо, як вони працюють.»* (Майк Зайко). А можливо винні усі й одразу? На це питання вкрай важливо знати відповідь, адже це може допомогти відповісти на наступне питання щодо утилізації непотрібних упереджень. Я з інтересом спостерігав за відповідями на запитання «хто винен?» обох респондентів, оскільки вони виявилися кардинально різними і водночас яскраво продемонстрували науковий профіль кожного. Майк Зайко комплексно підійшов до відповіді, пояснюючи, що тут скоріш за все може не прямої упередженості на індивідуальному рівні, оскільки індикатором завжди є система алгоритму:

«...what we're talking about often gets characterized as systemic bias, systemic discrimination, and it's not necessarily about individual bias. It's about institutional discrimination. These are all ways of saying that there doesn't have to be an intention there to discriminate. Someone doesn't have to be explicitly sexist or racist to exclude women or racial minorities» (Майк Зайко)

Тобто злого наміру дійсно може не бути на етапі тренування алгоритма даними, проте зважаючи на логіку функціонування алгоритмів та механізм аналізу даних – стає зрозумілим, що з «нічого» алгоритм не може створити упередженість, потрібне підґрунтя у вигляді тренувальних даних, які містять потрібно патерни, що їх згодом відстежить алгоритм:

«If you're doing this thoughtlessly, if you're not checking how these systems work, or if you're implementing a system without understanding what it's doing, and it ends up discriminating, then that can be a way that you are allowing this sort of discrimination to occur, even if you didn't intend it» (Майк Зайко)

Насправді важливо усвідомлювати, що навіть найменша недбалість на етапі оцінки даних для тренування алгоритмів може призвести до негативних наслідків. В такому ключі вина може лежати дійсно на інженерів штучних мереж. Проте, потрібно врахувати також і позицію людей, що займаються імплементацією алгоритмів. З цікавого спостереження після інтерв'ю, проблема яку Зайко описував у своїй статті знайшла пряме відображення та підтвердження у особі другого респондента. Отже, суть проблеми полягає у тому, що часто розробники спираються на тренувальні дані (що є відображенням реального життя відповідно до змінних) як на недоторкану «базову правду» (ground truth). В такому випадку намагання змінити «правду життя» буде вважатися упередженим. Тобто якщо дані з реального життя містять упередженість, намагання внести зміни до цих даних викликає побоювання у цілісності та точності моделі. Коротше кажучи, хоча алгоритмічну модель, яка точно відображає нерівну реальність суспільства, можна вважати неупередженою з точки зору точності відображення, її все одно можна вважати упередженою, якщо розглядати її як таку, що має небажану тенденцію. (Zajko, 2022)

При розмові на тему нерівності та упередження, що містять алгоритми, Антон С. майже з абсолютною точністю відобразив тезис Зайка з позиції розробника: *«...якщо людина зможе побудувати нейронну мережу яка буде неупереджена щодо*

таких питань то це значить ми відходимо від історичних даних, того що існує в світі. І чи це правильно, чи це неправильно, знову-таки, ніхто не може сказати. І ми не можемо сказати чи буде ідеальний світ коли у нас буде нейронна мережа яка буде ідеально розподіляти от там людей брати на роботу, не брати на роботу, куди брати, кого брати...» (Антон С.)

І все ж, якщо розглядати варіанти виправлення алгоритмів в бік утилізації упередження, способи існують. Якщо ми замислюємося над користю соціологів у нівелюванні даних нерівностей на ринку праці, непоганим рішенням було би активне залучення соціологів до розробки моделей штучного інтелекту. Проте, як зазначає Зайко: «... були спроби... залучити як допоміжний засіб... але програмісти усе одно залишаються головними...»

Кращим варіантом виглядає зберігати «певну дистанцію для критики, мати можливість критикувати ці розробки ззовні, а також постфактум» (Майк Зайко). Іншим способом покращення ситуації з алгоритмом, що вже відрекомендував себе як упереджений, як не дивно, є звичайна зміна тренувальних даних, повністю або опосередковано. Адже «...результати роботи нейронної мережі на 70% залежать від даних, тобто у нас є лише 30% для того щоб щось зробити, якимось повпливати як розробники» (Антон С.). Як ми бачимо, можна спробувати виправити й механічним шляхом, «покрутивши» модель та алгоритм, проте, простішим рішенням буде заміна масиву даних. Однак, для того щоб замінити дані потрібна людина, котра вважає за потрібне їх замінити. Тут може згодитися комбінація із залученням соціологів до розробки моделей штучного інтелекту.

2.4. Шляхи мінімізації негативного впливу штучного інтелекту на ринку праці України. Негативні наслідки використання штучного інтелекту на ринку праці

Наостанок, важливою частиною роботи є оцінка та пропозиція шляхів мінімізації негативного впливу штучного інтелекту конкретно щодо ринку праці України. Також необхідною частиною є окреслення потенційних негативних наслідків використання штучного інтелекту на ринку праці. Враховуючи пророблену роботу з літературою та опис великої частки складових штучного інтелекту, виконати поставлене завдання не буде проблемою.

Під час етапу підготовки необхідної літератури для дослідження, я зіткнувся із проблемою недостатньої висвітленості тематики нерівності, що може бути відтворена штучним інтелектом саме в українському контексті. Переважна більшість статей та робіт були спрямовані на виключно позитивний аспект імплементації штучного інтелекту, зокрема й на ринку праці. Проте, зрозуміло, що негативні наслідки використання штучного інтелекту присутні. Більше того, українці в цілому свідомі щодо загроз, що може нести за собою імплементація штучного інтелекту. За даними дослідження Кантар (Kantar), 57% респондентів остерігаються розвитку штучного інтелекту, зокрема серед найпопулярніших категорій побоювання є страх втрати робочих місць і знецінення фаху (22%), а також загроза неконтрольованої діяльності, повстання машин і можлива агресія до людини (22%) (2023). Отож, в силу браку джерел на тему нерівності та штучного інтелекту саме українського походження, я буду спиратися на доробок іноземних науковців з даної галузі при спробі надати шляхи мінімізації негативного впливу штучного інтелекту на ринок праці України.

Насправді, рішень, які можна здійснити вже зараз, аби не мати невітні наслідки у майбутньому, достатньо. До прикладу, непоганою опцією для середовищ, де штучний інтелект вже імплементований, може бути аудит алгоритму на упередження та відповідне наймання аудиторів алгоритмів. Аудит упередженості алгоритму — це незалежна оцінка алгоритму, яка першочергово фокусує увагу на виході алгоритму у пошуку упереджень, пов'язаних з вибраною категорією людей (Recker, 2023). Очевидно, що найкращим рішенням є наймання незалежних аудиторів задля мінімізації заангажованості. Дане рішення допоможе компаніям зменшити ризик виникнення упередження алгоритмів та подальшої нерівності.

Якщо ж ми говоримо про більш масштабне рішення, звісно, регуляція штучного інтелекту з боку держави може стати корисною. До того ж, такий приклад у світі вже існує – зовсім нещодавно, 13 березня 2024 року Європейський парламент одноставним рішенням ухвалив перший у світі закон про штучний інтелект, який регулює абсолютно усі сектори штучного інтелекту та класифікує деякі як ті, що несуть в собі неприйнятний ризик, з подальшою заборонаю (як-от система соціального рейтингу Китаю), деякі класифікує як високого ризику (інструмент, що сканує резюме та оцінює кандидатів) та зобов'язує відповідати певним правовим вимогам. Даний закон – чудове рішення при умові перебування у демократичних руках: заборона на державному рівні небезпечного та регулювання штучного інтелекту, що має високий ризик продукувати нерівність, є вкрай ефективним інструментом. До того ж, 45% респондентів вважають, що Україні потрібен саме такий закон (Kantar, 2023). Впровадження помірною державного регулювання штучного інтелекту, що не обмежуватиме громадян у правах і свободах допоможе здійснити крок назустріч зупинці циклу відтворення нерівності між реальністю та алгоритмами штучного інтелекту.

На додаток, непоганим рішенням видається використання допоміжного програмного забезпечення, котре дозволяє з більшою точністю знаходити

упередженості в моделях штучного інтелекту (у машинному навчанні, якщо бути точним). Прикладом може слугувати програма «AI Fairness 360 toolkit (AIF360)», випущена компанією IBM. Програма надає розробникам штучного інтелекту та дата-саєнтистам інструментарій для легкого виявлення упереджень на різних етапах робочого процесу машинного навчання, використовуючи відповідні метрики відповідно до контексту. Більше того, воно пропонує передові методи зменшення виявлених упереджень. Ці методи виявлення також піддаються автоматизації, дозволяючи командам штучного інтелекту регулярно перевіряти наявність упереджень так само, як вони перевіряють наявність багів або безпекових проблем у процесі неперервної інтеграції. Такі інструменти можуть стати в нагоді компаніям, що активно використовують штучний інтелект у своїй діяльності.

Якщо ж виокремити негативні наслідки використання штучного інтелекту на ринку праці України, це будуть наступні елементи. Першочергово, неминучість використання штучного інтелекту для автоматизації широкої кількості процесів може призвести до формування упереджень алгоритму. Відтак, як згадувалося у підрозділі 2.2., певні маргіналізовані групи, котрі не потрапляють до масиву даних, так і залишаться у своєму статусі. Наприклад, відсутність ромів у тренувальних масивах даних алгоритму прийому на роботу, призведе до відтворення вже наявної етнічної нерівності а можливо навіть й до збільшення (з ширшою імплементацією штучного інтелекту на ринок праці, все більше компаній та підприємств використовуватимуть алгоритми, в тренувальних масивах даних котрих відсутні представники маргіналізованих груп).

Що ж стосується гендерної нерівності, можуть траплятися схожі кейси з Амазоном, коли в силу гендерного розподілу працівників певних сфер, представники іншого гендеру виявляться викинутими з моделі. Або ж, можна зробити припущення, що якщо для вибору кандидатів на роботу у масиві даних враховуватимуться змінні статі та, наприклад, середньорічної зарплатні або ж

іншого фінансового показника, жінки котрі в середньому в Україні мають меншу зарплатню ніж чоловіки можуть вважатися алгоритмом вдалим кандидатом в силу менших витрат на зарплатню. В такому випадку, гендерний розрив, а разом із ним гендерна нерівність продовжуватиме відтворюватися.

Отже, згідно з завданнями даного розділу було здійснено наступний перелік дій: у першому підрозділі було окреслено вплив штучного інтелекту на ринок праці у глобальному контексті, надано оцінку потенційним негативним наслідкам імплементації штучного інтелекту, зокрема спираючись на висловлювання Майка Зайко у глибинному інтерв'ю. Було виділено сферу ринку праці, де вкрай часто трапляються ситуації з упередженням штучного інтелекту, пояснено механізм дії та наведено приклади. У другому підрозділі було визначено проблемні типи нерівності на ринку праці України, наведено контекст. Було роз'яснено пов'язаність існуючих нерівностей та імплементації штучного інтелекту на ринку праці України крізь логіку функціонування алгоритмів. У третьому підрозділі було розглянуто поняття алгоритму штучного інтелекту, пояснено зв'язки між алгоритмом та моделлю, уточнено поняття «створення» алгоритмів. Також було наведено рішення з усунення упереджень з алгоритмів, користуючись даними з інтерв'ю Майка Зайка та Антона С. Наостанок, у четвертому підрозділі було описано проблематику українських досліджень на тему нерівності та штучного інтелекту, наведено потенційні шляхи мінімізації негативного впливу штучного інтелекту на ринок праці України. Також було зазначено негативні наслідки використання штучного інтелекту на ринку праці.

ВИСНОВКИ

У результаті даного, головним чином, теоретичного дослідження було визначено наявні та потенційні прояви нерівності у контексті впровадження штучного інтелекту на ринку праці України. Це було здійснено шляхом роботи переважно із закордонними джерелами в силу недостатньої висвітленості тематики штучного інтелекту та нерівності на ринку праці в Україні. Було у повному обсязі концептуалізовано поняття штучного інтелекту у контексті відтворення нерівностей, концептуалізовано поняття алгоритмів штучного інтелекту, механізмів роботи алгоритмів. Здійснено аналіз теоретичного доробку на тему нерівності, обрано базисні концепції авторів, наведено паралель обраних теорій з імплементацією штучного інтелекту. Також була проведена суттєва робота з опрацювання наукових досліджень на тему штучного інтелекту та відтворення нерівності. Було також визначено шляхи подолання нерівності у контексті імплементації штучного інтелекту на ринку праці України, найефективнішим визнано державне регулювання штучного інтелекту. Врешті, було визначено подальший вплив штучного інтелекту на ринок праці України шляхом транспонування світового становища впливу штучного інтелекту на ринок праці України.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- Національний інститут стратегічних досліджень. (2023, June 2). Гендерні диспропорції в Україні під час війни.
- Паніотто, В. (2022, October 20). МІЖЕТНІЧНІ УПЕРЕДЖЕННЯ в УКРАЇНІ, ВЕРЕСЕНЬ 2022. *КМІС*.
<https://www.kiis.com.ua/?lang=ukr&cat=reports&id=1150&page=1>
- Самаєва, Ю. (2023, July 16). Ставлення українців до штучного інтелекту на диво легковажне. *Дарма. Дзеркало Тижня | Mirror Weekly*.
<https://zn.ua/ukr/TECHNOLOGIES/stavlennja-ukrajintsiv-do-shtuchnoho-intelektu-na-divo-lehkovazhne-darma.html>
- Тужанський, Д. (2022, April 15). Ромські сім'ї повертаються до України та скаржаться на дискримінацію в Європі. *Infopost.Media*.
<https://infopost.media/romski-simy-i-povertayutsya-do-ukrayiny-ta-skarzhatsya-na-dyskryminacziyu-v-yevropi/>
- Уповноважений Верховної Ради України з прав людини (2022, September 28). Моніторинг додержання прав і свобод ромської національної меншини в умовах воєнного стану, а також прав і свобод ВПО у Львівській та Закарпатській областях.
https://www.ombudsman.gov.ua/news_details/monitoring-doderzhannya-prav-i-svobod-romskoyi-nacionalnoyi-menshini-v-umovah-voyennogo-stanu-takozh-prav-i-svobod-vpo-u-lvivskij-ta-zakarpatskij-oblastyah
- Acemoglu, D. & Restrepo, P. (2017). Robots and Jobs: Evidence from US Labor Markets. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2940245>
- Anirudh, V. K. (2022, February). What Is the Difference Between Artificial Intelligence, Machine Learning, and Deep Learning? *Spiceworks*.
https://www.spiceworks.com/tech/artificial-intelligence/articles/ai-machine-learning-deep-learning/#_008
- Artificial Intelligence, Automation & Marxism. (2023, May 5). VeroZen.com.
<https://verozen.com/f/artificial-intelligence-automation-%E2%80%98marxism%E2%80%99>
- Baudrillard, J. (1994). *Simulacra and Simulation* (S. Glaser, Trans.). *University of Michigan Press*.
- Baudrillard, J. (1998). *The consumer society: Myths and Structures*. *SAGE Publications Limited*.

- Brownlee, J. (2020, August 19). Difference between algorithm and model in machine learning. <https://machinelearningmastery.com/difference-between-algorithm-and-model-in-machine-learning/>
- Darwish, A., Hassanien, A. E., & Das, S. (2019). A survey of swarm and evolutionary computing approaches for deep learning. *Artificial Intelligence Review*, 53(3), 1767–1812. <https://doi.org/10.1007/s10462-019-09719-2>
- Dastin, J. (2018, October 11). INSIGHT - Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G/>
- Delua, J. (2021, March 12). Supervised vs. unsupervised learning: What's the difference? *IBM*. <https://www.ibm.com/think/topics/supervised-vs-unsupervised-learning>
- Ellena, S., & Makszimov, V. (2022, April 14). Faced with discrimination, Ukrainian Roma refugees are going home. *Euractiv*. <https://www.euractiv.com/section/non-discrimination/news/faced-with-discrimination-ukrainian-roma-refugees-are-going-home/>
- European Network on Statelessness. (2022, March 10). BRIEFING: Stateless people and people at risk of statelessness forcibly displaced from Ukraine. https://www.statelessness.eu/sites/default/files/2022-03/ENS%20Briefing%20-%20Stateless%20people%20displaced%20from%20Ukraine%20-%20March%202022_1.pdf
- Florence-Heynard. (2024, April 24). Will artificial intelligence help or hurt the job market and the economy? *Le Blog Du MBA MCI*. <https://mbamci.com/2024/04/will-artificial-intelligence-help-or-hurt-the-job-market-and-the-economy/>
- Georgieff, A. (2024, Січень), « Artificial intelligence and wage inequality », OECD Artificial Intelligence Papers, n° 13, Éditions OCDE, Paris, <https://doi.org/10.1787/bf98a45c-en>.
- Georgieva K. (2024, January 14). AI will transform the global economy. Let's make sure it benefits humanity. *IMF*. <https://www.imf.org/en/Blogs/Articles/2024/01/14/ai-will-transform-the-global-economy-lets-make-sure-it-benefits-humanity>
- Hardesty, L. (2017, April 14). Explained: Neural networks. *MIT News | Massachusetts Institute of Technology*. <https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414>

- IBM Data and AI Team. (2023, October 16). Shedding light on AI bias with real world examples. *IBM*. <https://www.ibm.com/blog/shedding-light-on-ai-bias-with-real-world-examples/>
- Kalluri, P. (2020). Don't ask if artificial intelligence is good or fair, ask how it shifts power. *Nature*, 583(7815). <https://doi.org/10.1038/d41586-020-02003-2>
- Kanade, V. A. (2022, April). What Is Machine Learning? Definition, Types, Applications, and Trends. *Spiceworks*. <https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-ml/>
- Kantar. (2023, December, 7). І вабить, і страшить: рік буму довкола ШІ. <https://www.kantar.com/ua/inspiration/consumers/ai-in-ukraine>
- Krasadakis, G. (2023, October 20). The impact of artificial intelligence on the society and everyday life | 60 leaders. Medium. https://medium.com/60-leaders/the-impact-of-ai-on-society-and-everyday-life-711307e06b87#_ftn2
- Marx, K. (1976). *Das Kapital: A Critique of Political Economy* (Vol. 1, B. Fowkes, Trans.). *Penguin Books*. (Original work published 1867)
- Mba, B. L. (2023, July 30). The role of AI in the Fourth Industrial Revolution (4IR). *Medium*. <https://medium.com/@besniklimaj/the-role-of-ai-in-the-fourth-industrial-revolution-4ir-2b115145804b>
- Nida. (2019, October 18). How algorithms are shaping an alternative reality — a Baudrillardian analysis. *Medium*. <https://towardsdatascience.com/how-algorithms-are-shaping-an-alternative-reality-a-baudrillardian-analysis-76e68f6dad56>
- Recker, J. (2023, February 8). What is an Algorithmic Bias Audit? - Jeffery Recker - Medium. *Medium*. <https://medium.com/@jeffery-recker/what-is-an-algorithmic-bias-audit-ea71252b0ec3>
- Russell, R. (2023, December 7). Artificial intelligence in the job market - Richard Russell - Medium. *Medium*. <https://medium.com/@russelri411/artificial-intelligence-in-the-job-market-d1a3331650f6>
- Russell, S., & Norvig, P. (2022). *Artificial intelligence: A Modern Approach*. *Pearson Higher Education*.
- Shubhendu, S. S., & Vijay, J. (2013, September). Applicability of Artificial Intelligence in Different Fields of Life. *International Journal of Scientific Engineering and Research (IJSER)*.
- Tableau Software. (2021). Artificial intelligence (AI) Algorithms: A complete overview. *Tableau.com*. <https://www.tableau.com/data-insights/ai/algorithms>

- Thormundsson, B. (2024, February 16). Artificial intelligence (AI) worldwide - statistics & facts. *Statista*. <https://www.statista.com/topics/3104/artificial-intelligence-ai-worldwide/#topicOverview>
- Thormundsson, B. (2024, May 24). Global artificial intelligence market size 2021-2030. *Statista*. <https://www.statista.com/statistics/1365145/artificial-intelligence-market-size/>
- Tiwari, R. (2023). The impact of AI and machine learning on job displacement and employment opportunities. *International Journal of Scientific Research in Engineering and Management*.
https://www.researchgate.net/publication/367092106_The_impact_of_AI_and_machine_learning_on_job_displacement_and_employment_opportunities
- Unesco. (2021). UNESCO Science Report: The race against time for smarter development. *UNESCO Publishing*.
- Wajcman, J. (2017), Automation: is it really different this time?. *The British Journal of Sociology*, 68: 119-127. <https://doi.org/10.1111/1468-4446.12239>
- West, D. M., & Allen, J. R. (2018, April 24). How artificial intelligence is transforming the world. *Brookings*. <https://www.brookings.edu/articles/how-artificial-intelligence-is-transforming-the-world/>
- World Economic Forum. (2023, June). Global Gender Gap Report 2023. https://www3.weforum.org/docs/WEF_GGGR_2023.pdf
- Zajko, Mike. (2022). Artificial intelligence, algorithms, and social inequality: Sociological contributions to contemporary debates. *Sociology Compass*.
<https://doi.org/10.1111/soc4.12962>.
- Zion Market Research. (2022, June 23). Global Artificial Intelligence (AI) market to register an annual growth of 39.4% during forecast period.
<https://www.zionmarketresearch.com/news/global-artificial-intelligence-market>

ДОДАТОК А

Транскрипт інтерв'ю з Майком Зайко

10 травня, 19:00 – 19:23 (23 хвилини). Інтерв'ю з Майком Зайко, PhD Соціології, викладач Факультету мистецтв і соціальних наук University of British Columbia (Zoom)

Інтерв'юер:

All right. Recording has started and yeah, so we can begin. I read your article about the inequality that can be produced by artificial intelligence, about implementing an artificial intelligence by algorithms of artificial intelligence. And I have, I may begin with a question about... maybe you can, maybe you can tell how, how does the popularization and implementation of artificial intelligence may have the impact on a labor market like worldwide. Maybe you... because I, I read in your article that you also mentioned inequalities that can be produced by artificial intelligence on labor market too. And maybe you can tell me more about this.

Респондент:

Sure. It's not my main focus. So the things I have to say here are with not a lot of confidence is that this is a particular topic where there's been a lot of predictions and projections. Most of which have so far proven to be quite inaccurate, right? So, so claims about robots or AI's taking away 50% of jobs or making lots of industries obsolete, you know, those sorts of claims have so far largely been exaggerated. It does seem, though, that there is an impact on, you know, what you call labor market or on different fields of employment. And some jobs are being affected. And the ones that are sort of the most easy to automate are the ones that are primarily being affected, changed, eliminated in some cases, but mostly it's changing the type of work that the people are doing. So, you know, I would say that where a lot of the concerns have been are around this sort of like replacement, which jobs are going to be safe, which ones aren't going to be for automation. And with sort of the current hype around chatbots and generative AI, it seems like a lot of sort of existing customer chatbot type tasks, so customer service is probably seeing a lot of this sort of automation. Anywhere where there's sort of formulaic production of texts,

you know, is another area of the conversation. So, I mean, I guess it's not exactly labor, but particularly on sort of the spammier and the more criminal side of scams and text generation. But also just a lot of like very low quality web content stuff that maybe would have been outsourced to through Mechanical Turk and some other platforms. So, I'm going to say that like at this point it doesn't look like automation is putting a lot of people out of work, except for certain particularly vulnerable sectors. What it is doing is it's changing different aspects in the workforce and in production. And a lot of the concerns have been around hiring. So when it comes to sort of inequality, things that get labeled bias, I think that's where a lot of the concerns have rightfully been because the hiring process is being automated in a lot of places and that determines who gets what jobs.

Інтерв'юер:

Yeah, thank you. That's the question I wanted to ask you. Of course, when we're talking about AI, we consider job replacement, but also because AI is really useful in HR context, it can be... the inequalities, except the job replacement can be also producing biases and different inequalities like gender or racial, etc. So can you, maybe you have some thoughts on how artificial intelligence affects the production of inequality in labor market in context of helping assisting the HR department or something like that?

Респондент:

Sure. So, you know, the problem is what that's faced primarily by large employers who have big HR operations. They're dealing with lots and lots of applications and from an efficiency standpoint, it makes sense for them to automate a lot of this process. So my understanding is that where you see AI and automation in HR is primarily with large firms. And it's used to triage applications for human review, to determine which applications even make it to the point where a human might make a decision, which means that a lot of people are going to be excluded on some basis, which if we're talking about sort of machine learning-based systems, what they're going to do is they are going to learn or be trained on some kind of pattern in terms of what they are looking for for employees. So the risk is, the danger is that the sort of pattern they're going to be trained on is going to be some sort of historical set of inequalities that then becomes perpetuated. So there's the famous example, which is easy to find and is often cited, of Amazon trying to use this automated hiring system which they found discriminated against women. And even if you

don't have a clear indication of someone's gender, it's very easy for one of these pattern recognition systems to pick up on lots of proxies, lots of indicators that someone is a woman or someone is an ethnic minority. And if it seems that in the past, whatever training data the system was trained on, those people were disadvantaged, those categories were excluded, then it's just going to repeat those sorts of patterns. So, you know, that's kind of one of the more well-known dangers and harms that these systems can perpetuate, especially if we're not sort of examining and auditing them and seeing how they're treating different categories. Then there's, you know, one of the things I find interesting is in order to get past these sort of algorithmic systems, you have to somehow figure out what they're looking for, what they're picking up on, which creates this sort of incentive to find ways to use automation to sort of game the automation. So especially now where language models are sort of easier to use to compose text, lots of people are using ChatGPT or generative AI to help draft resumes and put out lots of applications. So you have, on the one hand, employees using AI to apply to many more jobs than they would have previously and to try to tailor those applications in a way that will get past the AIs that are screening them on the other side. And in that sort of situation, you know, whoever is best able to use the technologies, whoever has these sorts of digital skills to get past the automation has an advantage over people that don't.

Интерв'юер:

Okay, thank you. So, as far as I see, the problem may not be necessarily about the ones who write algorithms. So, or is it? So, is the reason why artificial intelligence, why neural networks are, for example, neural networks producing biases and inequalities, is the reason in the code, in the algorithms, in the people who write those algorithms? Or the reason may not be, or it's not their guilt. It may not be their guilt of these inequalities occurring?

Респондент:

Yeah, so, I mean, what we're talking about often gets characterized as systemic bias, systemic discrimination, and it's not necessarily about individual bias. It's about institutional discrimination. These are all ways of saying that there doesn't have to be an intention there to discriminate. Someone doesn't have to be explicitly sexist or racist to exclude women or racial minorities. It can be a kind of a negligence, though, especially

since these risks and harms are now fairly well documented. If you're doing this thoughtlessly, if you're not checking how these systems work, or if you're implementing a system without understanding what it's doing, and it ends up discriminating, then that can be a way that you are allowing this sort of discrimination to occur, even if you didn't intend it. So, I mean, the problem is a lot of the times these systems are black boxes. We don't know how they work. We can't see how they work. The code might be proprietary, so it belongs to a third party that is hired to run the HR system or provide the software. Or if it's one of these more complicated machine learning-based systems, even if you could look at the algorithm, you wouldn't be able to really understand it. The best way to do that would be to sort of see what comes out the other end. So, yeah, the short answer is, you know, the fault is in the code rather than someone's intentions, but that I don't think should allow people to get off the hook. There are human managers who are ultimately responsible for these things.

Интерв'юер:

Okay, thank you. Another question. We were talking about neural networks that use deep learning and using big data for generating responses, for making decisions, etc. And maybe you can tell me about artificial intelligence that is not based on deep learning and maybe these kind of artificial intelligence also may be involved in production of inequality, just if you know the cases or so.

Респондент:

Yeah, I mean, one of the problems here is the language is very imprecise and inconsistently used. So when I hear the words artificial intelligence, I'm typically thinking of machine learning, deep learning, things that evolved from these neural networks. But a lot of the time, people are just talking about different forms of automation. And these algorithms can be a lot more straightforward, simplistic, primitive, you know, if it comes to something like hiring, and you're using an algorithm to screen applicants, very often, all it's doing is it's checking certain boxes. It's seeing, you know, are these qualifications met? Are these conditions met? Does this person have the required degree and expertise? You don't need machine learning for that, right? What machine learning is useful for are these more complex patterns that especially arise out of looking at lots and lots of data, and then trying to make classifications or predictions on the basis of that. So, you know, very often, you

might use a sort of a simpler algorithm that isn't based on machine learning, just to automate a process that requires seeing whether or not something exists. And for maybe more complex questions, like how sort of can we predict whether certain employees are going to stick around for a long time or be successful in the long term, that's when people are tempted to look more towards machine learning. If it's not something that you can simply read off the resume and the danger there is that, yeah, we we might not really understand what is being picked up on and there's the danger that that those patterns being reproduced are just going to cement and entrench particular inequalities. Just a heads up, I'm going to have to go in about three or four minutes, so one or two more questions would be good.

Интерв'юер:

Okay, last question then. As far as I understood, one of the ways to face this, these problems, the inequalities that are produced by artificial intelligence, by neural networks is to have to let those processes being socially monitored by social scientists, I guess, to involve sociologists, for example, into the process of writing an algorithm. Or in the process of changing that properly to reduce the probability of inequality to happen, yeah? And maybe you can for these two, three minutes, you can tell me about your view on how to prevent these things to happen. Like not about how to change the existing fact of inequalities, but how to prevent future neural networks to be biased. If you understood.

Респондент:

Yeah. So there's ways that various kinds of people can be more involved in the process of developing and implementing these algorithms and sociologists are not special or privileged. In some way, you know, there have been efforts to sort of bring social scientists in at different points, but often as a kind of a secondary capacity, right, the the technologists are still in charge, ultimately, there's there's a product that's being developed and we're going to hire some social scientists to see if we can make it less bad. Where I see kind of more of a role for social scientists, for academics is with a bit more sort of critical distance, being able to critique these developments from the outside and then also after the fact. So, there needs to be regulation that is effective. I mean, regulation exists in the labor market, in hiring, but it requires us to have a good understanding of what is going on and a variety of people can try to help with that. There's a variety of ways that you can audit these systems, so algorithmic audits are kind of seen as the main approach to addressing some of these harms. And there's there's not a ton of agreement about what that sort of audit should look like, what it should entail. Do you need to see the code? Do you

just need to see the sorts of outputs that it produces? How far should we look at these sort of social impacts? Hiring decisions are kind of one of the the easier cases to talk about whether somebody gets a job or doesn't get a job. But when your actual workflow is being automated, when aspects of your job are are now being run through artificial intelligence, if you're trying to sort of figure out, like, well, what are the impacts on a worker? What are the impacts on their time, on the sort of their quality of life, that sort of thing that requires maybe some more qualitative and in-depth sort of methods, you need to actually understand the experience of these workers who are interacting with these systems. So, you know, there is a role for independent auditors or ones that are sort of hired by the company, it's now a business model, you can hire algorithmic auditors to to check your algorithms for bias and evaluate these these social impacts. And I think, you know, a lot of this stuff can only go so far if it isn't required by government and regulation. So those are kind of the things I would highlight this need for for algorithmic auditing. But it will only be so effective if if you leave it up to companies to sort of decide to what extent and how they want to conduct those audits.

Интерв'юер:

Thank you so much for I guess we run out of time. So thank you very much for this interview. And would you like to remain confidential or can I use your name in my diploma? Like when I will cite this interview or in the analysis or in some parts of my diploma?

Респондент:

I think it's fine if you use my name. Yeah, feel free to put that in there. And I wish you all the best in pulling this together. If you want to be back in touch about anything over email, just send me a line and I'll see if I can be of any further help, yeah. Good night for you, I guess!

Интерв'юер:

Ok, thank you very much. Yeah, and good day for you, too. And have a nice day. Bye. And major thanks.

Респондент:

Ok, no worries. Bye bye.