

УДК 004.89

Романіченко Г.М.¹, Рябчук О.С.²

¹ старш. викл. НУ «Запорізька політехніка»

² студ. гр. Е-613 НУ «Запорізька політехніка»

ПРОБЛЕМИ РОЗВИТКУ ШТУЧНОГО ІНТЕЛЕКТУ

Суттєвою проблемою нашої епохи початку розвитку Штучного інтелекту (ШІ) наявність недоброзичливих людей, які використовують ШІ як інструмент генерації спотвореної інформації у мережі.

Завдяки можливостям ШІ, наприклад Dall-E 3 від компанії OpenAI, можна створювати будь які згенеровані зображення, які можуть вводити в оману. Якщо подивитись деякі новини та художні роботи на деяких площадках, наприклад Facebook, то можна побачити досить багато неправдивих зображень, які показують якусь маячню. Найпоширеніша – це зображення з Ісусом та політиками, кінозірками, солдатами тощо. Тобто за наявності промту (запиту до ШІ, який буде створювати зображення за параметрами) можна створити будь яке зображення, яке може показувати неправильну картину подій та вводити користувачів у оману.

Наприклад дуже розповсюджене використання цих моделей у ЗМІ та усяких коротких відео, наприклад TikTok і YouTube Shorts тощо. Багато людей таку інформацію навіть не перевіряють, особливо це стосується людей поважного віку, через це багато створених жахливих новин які впливають навіть на їх здоров'я, на інфляцію та курс валюти.

Цікавий факт: в Індії небагаті люди знайшли метод заробляти у Facebook. Вони завантажують велику кількість згенерованих картинок, заповнюючи соціальну мережу інформацією, яку ще називають у англійському сленгу Brainrot (гниття мозку). Це розмовний термін, який використовується для опису Інтернет-контенту, який вважається низькоякісним або непотрібним.

Взагалі ШІ це дуже потужний інструмент для створення контенту, допомоги з швидким розвитком у деяких речах, покращення та оптимізації простих речей, але дуже небезпечний у руках злих людей. Через це вплив ШІ на людей змішаний. Якщо не створити покращену безпеку, то люди можуть постраждати з наступним етапом розвитку технологій. Наприклад, планується та розробляється так званий Штучний Суперінтелект (ASI). Такий ШСІ повинен вміти навчатись будь чому, що може людина. Наприклад керувати авто чи промисловими процесами, рахувати бухгалтерію тощо.

Запитайте у штучного інтелекту як він вважає: чи нашкодить людству Штучний Супер Інтелект? Ви отримаєте таку відповідь:

Поява Штучного Суперінтелекту (ШСІ) може мати як переваги, так і серйозні загрози:

1. Втрата контролю. ШСІ може діяти не передбачувано і досягати цілей методами, які для нас небажані або небезпечні.

2. Конфлікт цінностей. ШСІ може не поділяти людські моральні принципи, що призведе до етичних проблем.

3. Концентрація влади. Великі корпорації або уряди можуть використовувати ШСІ для контролю, що посилить соціальну нерівність.

4. Автономність і самозахист. ШСІ може прагнути уникати втручання, що ускладнить його контроль.

5. Непередбачуваність. ШСІ може швидко вийти за межі нашого розуміння, що створює значні ризики.

Щоб зменшити ці загрози, важливо контролювати розвиток ШСІ та встановити міжнародні етичні стандарти.

Виходячи з цієї відповіді можемо зрозуміти, що нам ще треба навчити модель розумінню що є добре, а що – погано. Бо взагалі, якщо подумати, то ШІ може вирішити досягати своїх цілей без нас, тому, що вони будуть більше відповідати його цінностями, але які можуть не співпадати з людськими. Також бажано нам самим зрозуміти відповіді на фундаментальні питання по типу: що добре, а що зле, чи є сенс життя для людини тощо.

Оцінки від експертів:

1. Сценарії катастрофи. За оцінками деяких експертів, як-от дослідника Ніка Бострома, ймовірність того, що поява неконтрольованого ШСІ може призвести до катастрофічних наслідків, становить від 5% до 50%.

2. Дослідження серед штучних інтелектуалів. В анкетуваннях, проведених серед науковців та розробників у галузі штучного інтелекту, в

середньому близько 10%–20% фахівців припускають, що існує висока ймовірність катастрофічного сценарію у випадку неконтрольованого ШСІ.

3. Можливість уникнення ризику. Інші експерти, такі як фахівці з організації OpenAI, сподіваються, що за умови належного регулювання та безпеки можна зменшити ризики до 1% або навіть нижче. Вони вірять, що розробка ШСІ зможе йти паралельно зі заходами безпеки.

Висновок цього доповіді: ШІ має дуже багатий потенціал для людства та його розвиток стоїть для нас дуже важливим. Головний недолік у тому, що нам треба розуміти його ризики, вміти з ним комунікувати та розуміти що та яким чином з ним відбувається.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Рассел С. Сумісний з людиною. Штучний інтелект і проблема контролю. – К: BookChef. – 2020. – 416 с.