

“EXPLAINABLE AI (TUSHUNTIRIB BERUVCHI SUN’IY INTELLEKT)”**Quvvatova Sanobar Gulbek qizi**

Qarshi Davlat Texnika Universiteti

“Axborot kommunikatsiya texnologiya

Sohasida kasb ta’limi” yo’nalishi

4-bosqich talabasi

E-mail: squvvatova62@gmail.com

Phone number: +99888 339 04 34

Annotatsiya: ushbu maqolada sun’iy intellekt tizimlari tobora ko’proq hayotimizga kirib kelayotgani sari ularning qanday qaror chiqarishini bilish muhim ahamiyatga ega bo’lmoqda. Shu bois tushuntirish imkoniyatiga ega bo’lgan sun’iy intellekt-Explainable AI (XAI) yo’nalishi shakllanmoqda. Ushbu yondashuv foydalanuvchi uchun aniqroq tushuncha yaratadi, shaffoflikni ta’minlaydi va oshiradi. Maqolada XAI tushunchasi, uning qo’llanish sohalari va kelajakdagi imkoniyatlari yoritiladi.

Аннотация: Поскольку системы искусственного интеллекта всё глубже проникают в нашу жизнь, важно понимать, как они принимают решения. Простые модели могут предсказывать результаты, но не могут объяснить, почему они таковы. Поэтому возможна разработка объяснимого ИИ (XAI). Он более точен, прозрачен и заслуживает доверия для пользователя. В статье будет рассмотрена область ИИ, её дополнительные области и качественные аспекты.

Abstract: as artificial intelligence systems become more and more embedded in our lives, it will be important to understand how they make decisions. Simple models can predict results but cannot explain why. Therefore, the form of artificial intelligence that can be worked on is Explainable AI (XAI). It is more accurate, transparent, and trustworthy for the user. The article will cover the field of AI, its additional areas, and its qualitative aspects.

Kalit so’zlar: sun’iy intellekt, tushuntirib beruvchi AI, XAI, shaffoflik, ishonch, qaror qabul qilish.

Ключевые слова: искусственный интеллект, объяснимый ИИ, ИИ, прозрачность, доверие, принятие решений

Keywords: Artificial Intelligence, explainable AI, AI, transparency, trust, decision-making

Explainable AI (XAI) – atamasi ingliz tilidagi “to explain” (tushuntirmoq, izohlab bermoq) fe’lidan kelib chiqqan “explainable” sifatdoshidan va “AI” qisqartmasidan tuzilgan.

Explainable – “izohlab bo’ladigan”, “tushuntirish mumkin bo’lgan” degan ma’noni anglatadi.

AI (Artificial Intelligence) – mashinalarning insondek fikrlash, o’rganish va qaror qabul qilish qobiliyatini bildiradi.

Sun’iy intellektning aksariyat modellarida natija to’g’ri chiqishi mumkin, ammo nega aynan shunday qaror qabul qilingani tushunilmaydi. Bunga “qora quti” muammosi deyiladi. Masalan, AI shifokorga bemorda kasallik borligini aytadi, lekin uning asosida qaysi belgilar turganini izohlab bera olmaydi.

Hozirgi kunda Explainable AI (XAI) texnologiyasi turli sohalarda keng qo’llanilmoqda va qaror qabul qilish jarayonini shaffoflashtirish orqali foydalanuvchi ishonchini oshirishda muhim ro’l o’ynamoqda. Quyida XAI’ ning eng muhim amaliy yo’nalishlarini keltirib o’tamiz:

Tibbiyot : sun’iy intellekt asosida ishlab chiqilgan diagnostik tizimlar kasalliklarni aniqlashda yuqori aniqlikka ega bo’lsada qaror sababini izohlash imkoniyati mavjud emas. XAI esa natijalarni izohlab, shifokorlarga qaror qabul qilish jarayonini tushunishga yordam beradi. Bu, o’z navbatida, bemor ishonchini oshirish va xatoliklarni kamaytirishga xizmat qiladi.

Moliya sohasi : kreditlash jarayonida yoki firibgarlikni aniqlashda oddiy AI faqatgina natija beradi. XAI esa qarorning sabablarini izohlash orqali mijozlar uchun shaffoflikni ta’minlaydi hamda moliyaviy institutlarning ishonchliligini oshiradi.

Huquq tizimi: jinoyatchilik xavfini baholash yoki huquqiy qarorlarni qo’llab- quvvatlashda XAI texnologiyasi qo’llanishi mumkin. Qaror sabablarini tushuntirish orqali tizim adolat va javobgarlikni kuchaytiradi.

Sanoat va ishlab chiqarish : avtomatlashtirilgan ishlab chiqarish tizimlarida XAI nosozlikni aniqlash bilan birga, ularning kelib chiqish sabablarini ham ko’rsatadi. Bu muhandislar uchun muammoni tezkor bartaraf etish imkonini beradi va ishlab chiqarish samaradorligini oshiradi.

Transport va avtonom vositalar: avtonom transport vositalari qabul qilgan qarorlarning izohlab berilishi xavfsizlik nuqtai nazaridan alohida ahamiyatga ega. XAI qarorlarni tushuntirish orqali foydalanuvchi ishonchini mustahkamlaydi va transport boshqaruvining ishonchliligini oshiradi.

Explainable AI texnologiyasi turli sohalarda qarorlarning shaffofligi, ishonchliligi va xavfsizligini ta’minlash orqali zamonaviy texnologik rivojlanishda muhim o’rin tutadi.

Explainable AI (XAI) ning afzalliklari. Explainable AI (XAI) texnologiyasining joriy etilishi sun’iy intellekt tizimlarining ishonchliligini va samaradorligini sezilarli darajada oshiradi. Uning asosiy afzalliklari quyidagilardan iborat.

Shaffoflikni ta'minlash : XAI qaror qabul qilish jarayonini izohlab beradi. Bu esa foydalanuvchiga natijani tushunish va tahlil qilish imkonini yaratadi.

Ishonchni oshirish: qarorlarning tushuntirilishi natijada foydalanuvchi va sun'iy intellekt o'rtasida ishonchni mustahkamlaydi. Bu ayniqsa sog'liqni saqlash, moliya va huquq kabi mas'uliyatli sohalarda muhimdir.

Xatoliklarni kamaytirish: natijalarni izohlab berish orqali noto'g'ri qarorlar tezda aniqlanadi va ularni tuzatish imkoniyati paydo bo'ladi.

Mas'uliyat va javobgarlik: XAI qarorlar uchun asosiy sabablarni ochib berishi tufayli huquqiy va etik javobgarlikni ta'minlashda yordam beradi.

O'quv jarayoniga hissa qo'shish: qaror sabablarini tushuntirish insonlar uchun qo'shimcha bilim manbai bo'lib xizmat qiladi, bu esa tizimdan nafaqat foydalanishni, balki o'rganishni ham osonlashtiradi.

Regulyator talablarga moslik: ba'zi sohalarda qarorlarni tushuntirish qonuniy talab hisoblanadi. XAI bunday taalablarni bajarishda muhim vosita bo'lib xizmat qiladi.

Explainable AI oddiy sun'iy intellekt tizimlariga qaraganda ancha shaffof, ishonchli va foydalanuvchi uchun tushunarli bo'lib, uni turli sohalarda joriy etish imkoniyatini kengaytiradi

Explainable AI (XAI) sun'iy intellektning shaffofligini ta'minlashda muhim yo'nalish bo'lsada, uni joriy etishda bir qator texnik va amaliy muammolar mavjud. Ushbu qiyinchiliklar XAI' ni to'liq amalga oshirish va keng qo'llash imkoniyatlarini cheklab turadi. Modelning murakkabligiga to'xtaladigan bo'lsak, zamonaviy chuqur o'rhanishga asoslangan modellar milliardlab parametrlarni o'z ichiga oladi. Ularning qarorlarini oddiy va tushunarli qilib izohlash ko'pincha mumkin emas. Tushuntirish jarayonida soddalashtirish kiritilishi esa natijaning aniqligiga zarar yetkazishi ehtimoli mavjud. Hisoblash resurslariga talab yuqoriligi. XAI metodlari odatiy AI modellariga nisbatan ko'proq vaqt va hisoblash quvvatini talab qiladi. Bu esa katta hajmdagi ma'lumotlarni qayta ishlashda tizimning sekinlashishiga, energiya sarfining oshishiga va qo'shimcha xarajatlarning paydo bo'lishiga olib keladi.

Aniqlik va tushuntirish o'rtasidagi muvozanat. Amaliyotda aniqlik eng yuqori bo'lgan modellar ko'pincha eng murakkab bo'ladi va ularni izohlash deyarli imkonsiz hisoblanadi. Tushuntirishga qulay modellar esa, aksincha, aniqlik jihatidan yutqazishi mumkin. Bu muvozanatni topish – XAI oldidagi eng katta muammolardan biridir.

Standartlarning yo'qligi ya'ni hozirda XAI uchun yagona universal metodologiya yoki xalqaro standart mavjud emas. Har bir sohada turli xil yondashuvlardan foydalanishi natijasida umumiy qabul qilingan baholash mezonlarini ishlab chiqish qiyinlashmoqda.

Foydalanish uchun murakkablik tomonlari bu tizim taqdim etgan izohlar ba'zan juda murakkab va texnik bo'lib, oddiy foydalanuvchi yoki qaror qabul qiluvchi

shaxs uchun tushunarsiz bo'lib qolishi mumkin. Demak, tushuntirishning o'zi ham inson uchun qulay shaklda taqdim etilishi zarur.

Explainable AI qarorlarni shaffoflashtirish va foydalanuvchi ishonchini oshirishda katta imkoniyatlarga ega bo'lsa-da, u bilan bog'liq texnik, huquqiy, iqtisodiy va amaliy qiyinchiliklar mavjud. Ushbu cheklovlarni bartaraf etish uchun yanada samarali metodologiyalar, xalqaro standartlar va foydalanuvchiga yo'naltirilgan yondashuvlar ishlab chiqarilishi zarur.

Sun'iy intellekt tizimlari inson faoliyatining deyarli barcha jabhalariga kirib kelmoqda, ammo ularning qaror qabul qilish jarayonini tushunish muammosi hanuz dolzarbligicha qolmoqda. Shu nuqtai nazardan Explainable AI (XAI) texnologiyasi nafaqat texnik yutuq, balki inson va mashina o'rtasidagi ishonchni mustahkamlashga xizmat qiladigan muhim vositadir. XAI qarorlarni izohlash, shaffoflikni ta'minlash, xatoliklarni kamaytirish va foydalanuvchi uchun tushunarli mexanizm yaratish orqali oddiy sun'iy intellekt tizimlaridan tubdan farq qiladi.

Biroq, ushbu texnologiyaning keng joriy etilishi hali ham qator qiyinchiliklarga duch kelmoqda: murakkab modellarni izohlashning qiyinligi, hisoblash resurslariga yuqori talab, universal standartlarning yo'qligi va etik-huquqiy masalalar. Shunday bo'lsa-da, XAI'ni rivojlantirish insoniyat uchun katta ahamiyat kasb etadi, chunki u qarorlarni yanada adolatli, xavfsiz va ishonchli qilish imkonini beradi.

Demak, Explainable AI kelajakda nafaqat sun'iy intellekt tizimlarining samaradorligini, balki jamiyatning unga bo'lgan munosabati va ishonchini ham belgilovchi asosiy omillardan biriga aylanishi kutilmoqda.

Explainable AI qarorlarni tushunarli va shaffof qilish orqali inson va sun'iy intellekt o'rtasida ishonchni mustahkamlovchi muhim texnologiya hisoblanib kelmoqda.

Foydalanilgan adabiyotlar ro'yxati

1. Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). IEEE Access, 6, 52138–52160. <https://doi.org/10.1109/ACCESS.2018.2870052>
2. Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608. <https://arxiv.org/abs/1702.08609>
3. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. ACM Computing Surveys, 51(5), 1–42. <https://doi.org/10.1145/32360009>
4. Molnar, C. (2022). Interpretable machine learning: A guide for making black box models explainable. Leanpub. <https://christophm.github.io/interpretable-ml-book>