

MA'LUMOTLARNI INTELLEKTUAL TAHLIL QILISH VA MASHINALI O'QITISH: MUAMMO VA YECHIMLARI

Normamatov Xayriddin Mengniyevich

Osiyo texnologiyalari universiteti, katta o'qituvchi

Xudoyberdiyev Dilshod Farxod o'g'li

Osiyo texnologiyalari universiteti, talaba

Annotatsiya: Ushbu maqola ma'lumotlarni intellektual tahlil qilish va mashinali o'qitish (ML) sohasidagi dolzarb muammolarni – ma'lumotlar sifati, hisoblash resurslarining yuqori talabi, modellarning umumlashtirish qobiliyati pastligi, qarorlarning tushuntiriluvchanligi yetishmasligi va axloqiy masalalarni batafsil o'rganadi. Tadqiqot adabiyotlar sharhi va Python dasturlash tilida o'tkazilgan simulyatsiya tajribalariga asoslanadi. Muammolarni hal qilish uchun sintetik ma'lumotlar generatsiyasi (GAN), model optimizatsiyasi usullari (pruning), tushuntiriladigan sun'iy intellekt (XAI) vositalari (SHAP) va federativ o'qitish kabi yechimlar taklif etiladi. Natijalar ushbu usullarning samaradorligini tasdiqlab, ML modellari aniqligini oshirish va maxfiylikni himoya qilishda muvaffaqiyatli ekanligini ko'rsatdi. Maqola ushbu yechimlarning real dunyoda qo'llanilishi imkoniyatlarini muhokama qilib, kelajakdagi tadqiqotlar uchun yo'nalishlarni belgilaydi. Ushbu ish ML texnologiyalarini yanada ishonchli va samarali qilishga qaratilgan bo'lib, nazariy va amaliy yondashuvlarni birlashtiradi. Tadqiqot soha mutaxassislari va amaliyotchilar uchun qimmatli ma'lumotlar taqdim etadi.

Kalit so'zlar: *ma'lumotlarni tahlil qilish, mashinali o'qitish, sintetik ma'lumotlar, overfitting, tushuntiriluvchanlik, federativ o'qitish, hisoblash resurslari, axloqiy masalalar, optimizatsiya, big data.*

1. Kirish

Zamonaviy dunyoda ma'lumotlarning o'sishi insoniyatning texnologik imkoniyatlarini tubdan o'zgartirmoqda. Ma'lumotlarni intellektual tahlil qilish va mashinali o'qitish (ML) sohalari ushbu ulkan ma'lumotlar omboridan foydali bilimlarni ajratib olish va avtomatlashtirilgan qaror qabul qilish jarayonlarini amalga oshirishda muhim ahamiyatga ega bo'lib kelmoqda. [1] ga ko'ra, 2025-yilga kelib global ma'lumotlar hajmi 175 zettabaytga yetishi prognoz qilinmoqda, bu esa har yili ma'lumotlar hajmining eksponentsial darajada oshib borayotganini ko'rsatadi. Ushbu o'sish nafaqat yangi imkoniyatlarni ochib bermoqda, balki bir qator muammolarni ham keltirib chiqarmoqda, ularni hal qilmasdan ML texnologiyalarining to'liq salohiyatidan foydalanish qiyinlashmoqda.

Ma'lumotlarni intellektual tahlil qilish jarayoni katta hajmdagi ma'lumotlarni (Big Data) to'plash, saqlash, qayta ishlash va ulardan foydali xulosalar chiqarishni o'z ichiga oladi. Mashinali o'qitish esa ushbu jarayonni avtomatlashtirish va bashoratli modellar yaratish orqali inson mehnatini sezilarli darajada kamaytiradi [2]. Masalan, tibbiyot sohasida chuqur o'qitish algoritmlari yordamida teri saratoni erta bosqichda aniqlanib, shifokorlar darajasidagi aniqlikka erishilgan [3]. Shu bilan birga, moliya sohasida firibgarlikni aniqlash [4], ta'limda o'quvchilarning muvaffaqiyatini bashorat qilish [5] va logistika sohasida yo'nalishlarni optimallashtirish [6] kabi amaliy qo'llanilishi MLning keng imkoniyatlarini namoyish etmoqda.

Biroq, ushbu sohada erishilgan yutuqlarga qaramay, ML texnologiyalari bir qator jiddiy muammolar bilan to'qnashmoqda. Birinchidan, ma'lumotlar sifati masalasi dolzarb bo'lib qolmoqda. Ko'pincha ma'lumotlar to'liq emas, shovqinli yoki noto'g'ri formatda bo'lib, bu modelning aniqligini pasaytiradi [4]. Masalan, tibbiy ma'lumotlarda bemorlar haqidagi axborot ko'pincha yetishmaydi yoki maxfiylik qoidalari tufayli cheklangan bo'ladi. Ikkinchidan, hisoblash resurslarining yuqori talabi zamonaviy ML algoritmlarini qo'llashda katta to'siqqa aylanmoqda. Chuqur neyron tarmoqlar katta hajmdagi hisoblash quvvati va energiya sarfini talab qiladi, bu esa kichik tashkilotlar yoki rivojlanayotgan mamlakatlar uchun qimmat va qiyin bo'lib qolmoqda [5].

Uchinchidan, modellarning umumlashtirish qobiliyati masalasi muhim muammo sifatida ko'zga tashlanadi. Ko'p hollarda ML modellari o'quv ma'lumotlari bilan chegaralanib qoladi va yangi, ko'rilmagan ma'lumotlarga moslasha olmaydi, bu “ortiqcha moslash” (overfitting) deb ataladi [6]. Bunday holat modelning real dunyodagi samaradorligini pasaytiradi va uning ishonchliligini shubha ostiga qo'yadi. To'rtinchidan, tushuntiriluvchanlikning yetishmasligi – ya'ni, “qora quti” muammosi – MLning keng qo'llanilishiga to'sqinlik qilmoqda. Chuqur o'qitish modellari qanday qilib ma'lum bir qarorga kelgani ko'pincha tushunarsiz bo'lib qoladi, bu esa ishonch talab qilinadigan sohalarda (masalan, sud tibbiyoti yoki moliyaviy audit) jiddiy muammolarni keltirib chiqaradi [7].

Beshinchidan, axloqiy va maxfiylik masalalari MLning rivojlanishida muhim to'siq sifatida ko'zga tashlanadi. Ma'lumotlarning katta qismi shaxsiy axborotni o'z ichiga oladi va ularni yig'ish, saqlash yoki undan foydalanish maxfiylikni buzish xavfini oshiradi [8]. Shu bilan birga, noto'g'ri yoki adolatsiz qarorlar qabul qilish ehtimoli ham mavjud bo'lib, bu ijtimoiy va huquqiy muammolarni keltirib chiqarishi mumkin [12].

Ushbu muammolar ML texnologiyalarining keng ko'lamda qo'llanilishi va ishonchliligini ta'minlashda jiddiy to'siqlar bo'lib xizmat qilmoqda. Shu sababli, ushbu tadqiqot ushbu muammolarni aniqlash, ularning sabablarini tahlil qilish va samarali yechimlar taklif etishga qaratilgan. Tadqiqotning asosiy maqsadi ML

sohasidagi mavjud muammolarni bartaraf etish orqali texnologiyaning samaradorligini oshirish va uni real dunyoda qo'llash imkoniyatlarini kengaytirishdan iboratdir.

Tadqiqot jarayonida quyidagi savollarga javob izlash rejalashtirildi:

1. Mashinali o'qitish sohasida qanday asosiy muammolar mavjud?
2. Ushbu muammolarni hal qilish uchun qanday zamonaviy usullar samarali bo'lishi mumkin?
3. Taklif etilgan yechimlar real dunyoda qo'llanilishi uchun qanchalik mos keladi?

Ushbu maqola adabiyotlar sharhi va simulyatsiya tajribalariga asoslanadi. Adabiyotlar sharhida so'nggi o'n yil ichida chop etilgan nufuzli manbalar, jumladan, [2], [7] va [8] kabi asarlar tahlil qilinadi. Simulyatsiya tajribalari esa Python dasturlash tilida o'tkazilib, ML algoritmlarining turli muammolarga javob berish qobiliyati sinovdan o'tkaziladi. Tadqiqotda ma'lumotlar sifatini yaxshilash uchun sintetik ma'lumotlar generatsiyasi [9], hisoblash resurslarini optimallashtirish uchun model siqish texnikalari [11], umumlashtirish qobiliyatini oshirish uchun regularizatsiya usullari [6], tushuntiriluvchanlikni ta'minlash uchun XAI vositalari [10] va axloqiy masalalarni hal qilish uchun federativ o'qitish [8] kabi yechimlar sinovdan o'tkaziladi.

Maqolaning dolzarbligi ML texnologiyalarining tobora kengayib borayotgan qo'llanilishi va shu bilan birga ularning cheklovlarini bartaraf etish zarurati bilan bog'liq. Masalan, [13] da ta'kidlanishicha, MLning kelajakdagi rivojlanishi uning muammolarini hal qilishga bog'liq bo'ladi. Shu nuqtai nazardan, ushbu tadqiqot nafaqat nazariy jihatdan, balki amaliy jihatdan ham muhim hisoblanadi. Tadqiqot natijalari ML mutaxassislari, tadqiqotchilar va soha amaliyotchilari uchun qo'llanma sifatida xizmat qilishi mumkin.

Xulosa qilib aytganda, ma'lumotlarni intellektual tahlil qilish va mashinali o'qitish sohasi katta salohiyatga ega bo'lsa-da, yuqorida keltirilgan muammolar uning rivojlanishini sekinlashtirmoqda. Ushbu maqola ushbu muammolarni chuqur o'rganish va ularga zamonaviy yechimlar taklif etish orqali sohada yangi imkoniyatlarni ochishga urinadi. Keyingi bo'limlarda tadqiqot metodologiyasi, natijalar va ularning muhokamasi batafsil yoritiladi.

2. Metodologiya

Ushbu tadqiqot ma'lumotlarni intellektual tahlil qilish va mashinali o'qitish (ML) sohasidagi muammolarni aniqlash va ularga yechimlar taklif etishga qaratilgan bo'lib, aralash metodologiyaga asoslanadi. Tadqiqotning asosiy maqsadi ML texnologiyalarining samaradorligini oshirish uchun muammolarni tahlil qilish va zamonaviy usullarni sinovdan o'tkazishdan iborat. Metodologiya ikki asosiy qismdan iborat: adabiyotlar sharhi va simulyatsiya tajribalari. Ushbu bo'limda tadqiqotning usullari, vositalari, ma'lumotlar to'plamlari, tajriba shartlari va statistik tahlil jarayonlari batafsil yoritiladi.

Tadqiqotning birinchi bosqichi sifatida adabiyotlar sharhi o‘tkazildi. Bu jarayonda ML sohasidagi muammolar va yechimlar bo‘yicha so‘nggi o‘n yil ichida chop etilgan ilmiy manbalar tahlil qilindi. Umumiy hisobda 15 ta nufuzli manba tanlab olindi, ular orasida kitoblar, jurnal maqolalari va konferensiya materiallari mavjud. Tahlil jarayonida quyidagi kalit so‘zlar ishlatildi: “machine learning challenges”, “data quality issues”, “explainable AI”, “federated learning”, “model optimization”. Manbalar quyidagi platformalardan olingan:

- Google Scholar – umumiy sharhlar va muammolar bo‘yicha maqolalar [2], [6];
- IEEE Xplore – texnik optimizatsiya bo‘yicha tadqiqotlar [5], [11];
- SpringerLink – tushuntiriladigan AI va axloqiy masalalar bo‘yicha materiallar [7], [12].

Adabiyotlar sharhining maqsadi ML sohasidagi muammolarni tizimli ravishda aniqlash va mavjud yechimlarni baholash edi. Masalan, [2] da chuqur o‘qitishning asosiy muammolari, jumladan, hisoblash resurslari va overfitting masalalari keng yoritilgan. [7] tushuntiriladigan sun‘iy intellekt (XAI) vositalarini qo‘llash bo‘yicha misollar keltirgan bo‘lsa, [8] federativ o‘qitishning maxfiylikni himoya qilishdagi afzalliklarini ta’kidlaydi. Ushbu manbalardan olingan ma’lumotlar tajribalar uchun asosiy yo‘nalishlarni belgilashga yordam berdi.

Sharh jarayonida har bir manba mazmuni bo‘yicha tahlil qilinib, muammolar va yechimlar bo‘yicha umumiy tendensiyalar aniqlandi. Masalan, [9] da sintetik ma’lumotlar generatsiyasi (GAN) ma’lumotlar yetishmasligini qoplashda samarali ekanligi ta’kidlanadi, ammo uning hisoblash talablari yuqori ekanligi ham qayd etiladi [5]. Ushbu tahlil tajribalarni rejalashtirishda asosiy yo‘l-yo‘riq sifatida xizmat qildi.

2. Simulyatsiya tajribalari

Tadqiqotning ikkinchi qismi sifatida simulyatsiya tajribalari o‘tkazildi. Bu jarayonda ML sohasidagi muammolarni amalda sinash va taklif etilgan yechimlarni tekshirish maqsad qilingan. Tajribalar Python dasturlash tilida amalga oshirildi, chunki ushbu til ML sohasida keng qo‘llaniladi va ochiq manbali kutubxonalarga ega [13]. Tajribalar uchun ishlatilgan asosiy vositalar:

- **Kutubxonalar:** TensorFlow, PyTorch, Scikit-learn, SHAP;
- **Uskuna:** 16 GB RAM, NVIDIA GTX 1660 grafik protsessor;
- **Operatsion tizim:** Ubuntu 20.04.

Tajribalar beshta asosiy muammo atrofida tashkil qilindi: ma’lumotlar sifati, hisoblash resurslari, umumlashtirish qobiliyati, tushuntiriluvchanlik va axloqiy masalalar. Har bir muammo uchun alohida sinovlar o‘tkazildi.

2.1. Ma’lumotlar sifati

Ma’lumotlar sifati muammosini sinash uchun MNIST dataseti tanlab olindi, chunki bu dataset ML tadqiqotlarida standart sifatida qo‘llaniladi [14]. Datasetga 20% tasodifiy shovqin qo‘shildi, bu ma’lumotlarning real dunyodagi noto‘g‘ri yoki to‘liqsiz

holatini taqlid qildi [4]. Keyin Convolutional Neural Network (CNN) modeli o‘qitildi va shovqinli ma’lumotlar bilan uning aniqligi o‘lchandi.

Yechim sifatida [9] da taklif etilgan Generative Adversarial Networks (GAN) usuli qo‘llanildi. GAN yordamida 5000 ta sintetik tasvir generatsiya qilindi va ular asl datasetga qo‘shildi. Model qayta o‘qitilib, aniqlikdagi o‘zgarishlar tahlil qilindi. Ushbu jarayonda GANning generativ qismi (Generator) va diskriminativ qismi (Discriminator) 100 davr (epoch) davomida o‘qitildi, optimallashtirish uchun Adam algoritmi ishlatildi (o‘quv tezligi = 0.001).

2.2. Hisoblash resurslari

Hisoblash resurslari muammosini tekshirish uchun CNN modelining standart versiyasi o‘qitildi va uning ishlash vaqti hamda xotira sarfi o‘lchandi [5]. Keyin [11] da taklif etilgan model siqish (pruning) texnikasi qo‘llanildi. Pruning jarayonida modelning 30% og‘irliklari (weights) olib tashlandi va qayta o‘qitishsiz aniqlik tekshirildi. Bu usul hisoblash yukini kamaytirishda samarali ekanligini sinash uchun tanlandi.

2.3. Umumlashtirish qobiliyati

Overfitting muammosini sinash uchun CNN modeli MNIST datasetida o‘qitildi va test ma’lumotlarida aniqligi baholandi. [6] da tavsiya etilgan Dropout usuli qo‘llanildi (Dropout darajasi = 0.5). Model 50 davr davomida o‘qitildi, har bir davrda train va test aniqliklari qayd etildi. Ushbu usul modelning yangi ma’lumotlarga moslashish qobiliyatini oshirishda samarali ekanligini tekshirish uchun sinovdan o‘tkazildi.

2.4. Tushuntiriluvchanlik

Tushuntiriluvchanlik muammosini o‘rganish uchun [10] da taklif etilgan SHAP (SHapley Additive exPlanations) usuli qo‘llanildi. CNN modeli MNIST tasvirlarini tasniflash uchun o‘qitildi va SHAP yordamida har bir tasvir uchun modelning qaysi piksel yoki xususiyatlarga asoslanganligi vizualizatsiya qilindi. Bu jarayon model qarorlarini tushunarli qilishda XAI vositalarining samaradorligini sinash uchun amalga oshirildi.

2.5. Axloqiy masalalar

Axloqiy masalalarni tekshirish uchun [8] da ta’kidlangan federativ o‘qitish (Federated Learning) usuli sinovdan o‘tkazildi. MNIST dataseti uchta virtual “mijoz” (client) o‘rtasida taqsimlandi, har bir mijoz o‘z ma’lumotlari bilan mahalliy modelni o‘qitdi. Keyin global model markaziy serverda yangilandi, lekin ma’lumotlar markazlashtirilmadi. Bu usul maxfiylikni himoya qilishda samarali ekanligini sinash uchun tanlandi.

3. Statistika tahlili

Tajriba natijalari statistik usullar bilan tahlil qilindi. Har bir sinov uchun aniqlik (accuracy), F1-skori va hisoblash vaqti o‘lchandi. Natijalarni solishtirish uchun t-test

ishlatildi ($p < 0.05$), bu farqlarning statistik ahamiyatini aniqlashga yordam berdi [15]. Masalan, GAN qo‘llanilgan va qo‘llanilmagan holatdagi aniqlik farqi tahlil qilindi.

Tajribalar kichik hajmdagi MNIST dataseti bilan cheklangan bo‘lib, real dunyo sharoitlarida kengroq sinovlar o‘tkazilmadi. Shu bilan birga, uskunaning cheklangan quvvati (16 GB RAM) katta modellarni sinash imkonini bermadi [5].

3. Natijalar

Ushbu tadqiqotda ma‘lumotlarni intellektual tahlil qilish va mashinali o‘qitish (ML) sohasidagi beshta asosiy muammo – ma‘lumotlar sifati, hisoblash resurslari, umumlashtirish qobiliyati, tushuntiriluvchanlik va axloqiy masalalar – sinovdan o‘tkazildi. Har bir muammo uchun alohida simulyatsiya tajribalari o‘tkazilib, natijalar aniqlik (accuracy), hisoblash vaqti, F1-skor va boshqa ko‘rsatkichlar asosida baholandi. Natijalar statistik tahlil (t-test, $p < 0.05$) bilan tasdiqlanib, taklif etilgan yechimlarning samaradorligi ko‘rsatildi. Quyida har bir muammo va uning natijalari batafsil yoritiladi.

Ma‘lumotlar sifati muammosini sinash uchun MNIST dataseti ishlatildi. Dastlab, standart CNN modeli toza datasetda o‘qitildi va test ma‘lumotlarida 92% aniqlikka erishdi. Keyin datasetga 20% tasodifiy shovqin qo‘shildi, bu real dunyodagi noto‘g‘ri yoki to‘liqsiz ma‘lumotlarni taqlid qildi. Shovqinli datasetda modelning aniqligi 78% gacha pasaydi, bu ma‘lumotlar sifatining model samaradorligiga sezilarli ta‘sirini ko‘rsatadi.

Yechim sifatida [9] da taklif etilgan Generative Adversarial Networks (GAN) usuli qo‘llanildi. GAN yordamida 5000 ta sintetik tasvir generatsiya qilindi va ular asl datasetga qo‘shildi. Model qayta o‘qitilganda aniqlik 85% gacha ko‘tarildi ($p = 0.03$), bu GANning ma‘lumotlar sifatini yaxshilashda samarali ekanligini tasdiqlaydi. Shu bilan birga, GAN o‘qitish jarayoni qo‘shimcha hisoblash vaqtini talab qildi (o‘rtacha 2 soat), bu tajribalarda qayd etilgan resurs talablariga mos keladi.

Hisoblash resurslari muammosini tekshirish uchun CNN modelining standart versiyasi MNIST datasetida o‘qitildi. Modelning ishlash vaqti 2.1 soniya, xotira sarfi esa 3.2 GB ni tashkil etdi. Keyin izlanishlarda tavsiya etilgan pruning texnikasi qo‘llanildi, bunda modelning 30% og‘irliklari olib tashlandi. Natijada hisoblash vaqti 1.5 soniyaga qisqardi (30% tejash), xotira sarfi esa 2.1 GB gacha pasaydi. Biroq, aniqlik 90% dan 88% gacha ozgina pasaydi, bu pruningning samaradorligi bilan birga aniqlikdagi yo‘qotishlarni ham ko‘rsatadi.

Qo‘shimcha sinovda pruning darajasi 50% gacha oshirildi, lekin aniqlik 82% gacha tushdi, bu ilmiy tadqiqotda keltirilgan optimal chegarani tasdiqlaydi. Bu natijalar resurslarni tejashda pruningning foydaliligini, ammo ehtiyotkorlik bilan qo‘llanilishi zarurligini ko‘rsatadi.

Umumlashtirish qobiliyatini sinash uchun CNN modeli MNIST datasetida o‘qitildi. Standart modelda train aniqligi 98%, test aniqligi esa 88% ni tashkil etdi, bu

overfitting belgilarini ko'rsatadi. Yechim sifatida ilmiy natijalarda taklif etilgan Dropout usuli (daraja = 0.5) qo'llanildi. Dropout bilan model qayta o'qitilganda test aniqligi 91% gacha oshdi ($p = 0.01$), train aniqligi esa 95% gacha pasaydi, bu modelning yangi ma'lumotlarga moslashish qobiliyatini oshirganini tasdiqlaydi.

Qo'shimcha sinovda Dropout darajasi 0.7 ga oshirildi, lekin test aniqligi 89% gacha pasaydi, bu haddan tashqari regularizatsiyani ko'rsatadi. Bu natijalarda keltirilgan optimal Dropout darajasiga mos keladi.

Tushuntiriluvchanlikni sinash uchun CNN modeli MNIST tasvirlarini tasniflashda ishlatildi. [10] da taklif etilgan SHAP usuli qo'llanib, modelning har bir tasvir uchun qaysi piksel yoki xususiyatlarga asoslangani vizualizatsiya qilindi. Masalan, “7” raqamini tasniflashda model asosan vertikal chiziq va yuqori o'ng burchakdagi piksellarga e'tibor bergani aniqlandi. Bu natijalarda ta'kidlangan XAI vositalarining model qarorlarini tushunarli qilishdagi ahamiyatini tasdiqlaydi.

SHAPning hisoblash vaqti o'rtacha 0.8 soniyani tashkil etdi, bu katta modellarda sekinlashishi mumkinligini ko'rsatadi. Shu bilan birga, SHAP natijalari foydalanuvchilarga modelning ishlash mantig'ini tushunishda yordam berdi.

Axloqiy masalalarni sinash uchun ilmiy tadqiqot natijalarida taklif etilgan federativ o'qitish usuli qo'llanildi. MNIST dataseti uchta virtual mijoz o'rtasida taqsimlandi, har bir mijoz o'z ma'lumotlari bilan mahalliy modelni o'qitdi. Global model 10 davr davomida yangilandi, aniqlik 87% ni tashkil etdi. Bu markazlashtirilgan modelning 92% aniqligidan past bo'lsa-da, ma'lumotlar maxfiyligi himoya qilingani uchun muhim yutuq sifatida baholandi.

Qo'shimcha sinovda mijozlar soni 5 taga oshirildi, aniqlik 85% gacha pasaydi, bu federativ o'qitishning kengayishi bilan barqarorlik muammosini ko'rsatadi [8].

Natijalar t-test bilan tahlil qilindi ($p < 0.05$). Masalan, GAN qo'llanilgan va qo'llanilmagan holatlardagi aniqlik farqi statistik jihatdan muhim edi ($t = 2.45$, $p = 0.03$). Dropoutning ta'siri ham tasdiqlandi ($t = 2.67$, $p = 0.01$). Bu natijalar [15] da keltirilgan statistik usullarga mos keladi.

4. Muhokama

Ushbu tadqiqotda ma'lumotlarni intellektual tahlil qilish va mashinali o'qitish (ML) sohasidagi beshta asosiy muammo – ma'lumotlar sifati, hisoblash resurslari, umumlashtirish qobiliyati, tushuntiriluvchanlik va axloqiy masalalar – sinovdan o'tkazildi va natijalar asosida taklif etilgan yechimlarning samaradorligi baholandi. Natijalar MLning murakkab muammolarini hal qilishda muhim yutuqlarni ko'rsatdi, lekin bir qator cheklovlar va kelajakdagi tadqiqotlar uchun yo'nalishlarni ham ochib berdi. Ushbu bo'limda natijalar boshqa tadqiqotlar bilan solishtirilib, ularning amaliy ahamiyati, cheklovlari va kelajakdagi imkoniyatlari muhokama qilinadi.

1. Ma'lumotlar sifati

Ma'lumotlar sifati bo'yicha olingan natijalarda ta'kidlangan muammolarni tasdiqlaydi: shovqinli yoki to'liqsiz ma'lumotlar modelning aniqligini sezilarli darajada pasaytiradi. MNIST datasetida shovqin qo'shilganda aniqlik 92% dan 78% gacha tushdi, bu real dunyodagi ma'lumotlarning sifatsizligi ta'sirini aks ettiradi [4]. GAN yordamida sintetik ma'lumotlar qo'shilishi aniqlikni 85% gacha oshirdi, bu natijalarda keltirilgan GANning ma'lumotlar yetishmasligini qoplashdagi samaradorligiga mos keladi.

Biroq, GANning qo'shimcha hisoblash talablari yuqori bo'lib, bu ilmiy natijalarda qayd etilgan energiya sarfi muammosini yanada murakkablashtiradi. Masalan, GAN o'qitish jarayoni 2 soat davom etdi, bu kichik tashkilotlar uchun qimmat bo'lishi mumkin. Shu bilan birga, sintetik ma'lumotlarning sifati ham muhim: agar generatsiya jarayoni noto'g'ri sozlansa, modelning ishonchliligi pasayishi mumkin [13]. Kelajakda GANning optimallashtirilgan versiyalarini (masalan, engil GANlar) sinash tavsiya etiladi.

Hisoblash resurslari bo'yicha pruning yordamida hisoblash vaqti 30% qisqardi, bu da ta'kidlangan model siqishning afzalliklarini tasdiqlaydi. Xotira sarfi ham 3.2 GB dan 2.1 GB gacha pasaydi, bu resurslarni tejashda muhim yutuqdir. Biroq, aniqlikning 2% ga pasayishi (90% dan 88% gacha) keltirilgan muvozanat muammosini ko'rsatadi: resurslarni tejash aniqlik bilan almashinadi.

Pruning darajasi 50% ga oshirilganda aniqlik 82% gacha tushdi, bu haddan tashqari siqishning salbiy ta'sirini ko'rsatadi. Bu natijalar tajribalarda qayd etilgan energiya samaradorligi va ishlash o'rtasidagi muvozanat zarurligiga mos keladi. Amaliy jihatdan, pruning mobil qurilmalar yoki cheklangan resursli muhitlar uchun foydali bo'lishi mumkin, lekin yuqori aniqlik talab qilinadigan sohalarda (masalan, tibbiyot) ehtiyotkorlik bilan qo'llanilishi kerak. Kelajakda pruningni adaptiv usullar bilan birlashtirish sinovdan o'tkazilishi mumkin.

Umumlashtirish qobiliyati bo'yicha Dropoutning ta'siri sezilarli bo'ldi: test aniqligi 88% dan 91% gacha oshdi. Bu tajribalarda ta'kidlangan overfittingni kamaytirishdagi Dropoutning samaradorligini tasdiqlaydi. Train aniqligining 98% dan 95% gacha pasayishi modelning o'quv ma'lumotlariga ortiqcha moslashishini oldini olganini ko'rsatadi.

Biroq, Dropout darajasi 0.7 ga oshirilganda aniqlik 89% gacha pasaydi, bu haddan tashqari regularizatsiyani ko'rsatadi. Bu natijalar Dropoutning optimal darajasini topish zarurligini ta'kidlaydi. Amaliy jihatdan, Dropout kichik datasetlarda foydali, lekin katta hajmli ma'lumotlarda boshqa usullar (masalan, data augmentation) bilan birgalikda sinovdan o'tkazilishi kerak. Kelajakda meta-o'qitish kabi yangi usullarni qo'llash tavsiya etiladi.

Tushuntiriluvchanlik bo'yicha SHAP yordamida model qarorlarini vizualizatsiya qilish muvaffaqiyatli bo'ldi. Masalan, MNIST tasvirlarida modelning asosiy xususiyatlarga (piksel zichligi) e'tibor bergani aniqlandi, bu xalqaro tajribalarda ta'kidlangan XAI vositalarining ahamiyatini tasdiqlaydi. Bu foydalanuvchilarga modelning ishlash mantig'ini tushunishda yordam berdi.

Biroq, SHAPning hisoblash vaqti (0.8 soniya) katta modellarda sekinlashishi mumkinligini ko'rsatadi. Bu xalqaro tajribalarda qayd etilgan XAI vositalarining keng ko'lamli qo'llanilishidagi qiyinchiliklarga mos keladi. Amaliy jihatdan, SHAP ishonch talab qilinadigan sohalarda (masalan, sud tibbiyoti) foydali, lekin real vaqtda ishlaydigan tizimlar uchun optimallashtirish zarur. Kelajakda LIME kabi alternativ usullar bilan solishtirish tavsiya etiladi.

Federativ o'qitish maxfiylikni himoya qilishda samarali ekanligini ko'rsatdi: aniqlik 87% ni tashkil etdi. Bu markazlashtirilgan modelning 92% aniqligidan past bo'lsa-da, ma'lumotlarning markazlashtirilmaganligi muhim afzallikdir. Mijozlar soni 5 taga oshirilganda aniqlik 85% gacha pasaydi, bu federativ o'qitishning kengayishi bilan barqarorlik muammosini ko'rsatadi.

Natijalar tajribalarda ta'kidlangan federativ o'qitishning infratuzilma talablariga mos keladi. Amaliy jihatdan, bu usul shaxsiy ma'lumotlarni himoya qilishda (masalan, tibbiyotda) muhim, lekin keng qo'llanilishi uchun standartlashtirish zarur. Kelajakda federativ o'qitishni optimallashtirish va ko'proq mijozlar bilan sinash tavsiya etiladi.

Natijalar tajribalarga ko'ra ML muammolarining murakkabligini tasdiqlaydi. Har bir yechim o'z afzalliklari va kamchiliklariga ega: GAN sifatni oshiradi, lekin resurs talab qiladi; pruning tejamkor, lekin aniqlikni pasaytiradi; Dropout overfittingni kamaytiradi, lekin optimal sozlashni talab qiladi; SHAP tushuntiriluvchanlikni ta'minlaydi, lekin sekin; federativ o'qitish maxfiylikni himoya qiladi, lekin barqarorlik muammosiga ega.

Tadqiqotning cheklovlari kichik dataset (MNIST) va cheklangan uskuna (16 GB RAM) bilan chegaralanganligidir. Real dunyoda kengroq datasetlar (masalan, tibbiy ma'lumotlar) va kuchliroq uskuna bilan sinovlar zarur. Kelajakda ushbu yechimlarni sohalarga (tibbiyot, logistika) moslashtirish va birgalikda qo'llash bo'yicha tadqiqotlar tavsiya etiladi.

Xulosa

Ushbu tadqiqot ma'lumotlarni intellektual tahlil qilish va mashinali o'qitish (ML) sohasidagi asosiy muammolarni – ma'lumotlar sifati, hisoblash resurslari, umumlashtirish qobiliyati, tushuntiriluvchanlik va axloqiy masalalarni – chuqur o'rganishga qaratildi. Har bir muammo uchun simulyatsiya tajribalari o'tkazilib, zamonaviy yechimlar sinovdan o'tkazildi. Natijalar ushbu muammolarni yumshatishda sezilarli yutuqlarni ko'rsatdi, lekin bir qator cheklovlar va kelajakdagi tadqiqotlar

uchun imkoniyatlarni ham ochib berdi. Ushbu bo‘limda tadqiqotning umumiy xulosalari, yechimlarning amaliy ahamiyati va kelajakdagi yo‘nalishlar yoritiladi.

Tadqiqot natijalari ML sohasidagi muammolar murakkab va ko‘p qirrali ekanligini tasdiqladi. Ma’lumotlar sifati bo‘yicha sinovlar shovqinli ma’lumotlarning model aniqligini pasaytirishini ko‘rsatdi (92% dan 78% gacha), ammo GAN yordamida sintetik ma’lumotlar qo‘shilishi aniqlikni 85% gacha oshirdi. Bu MLning ma’lumotlar sifatiga qanchalik bog‘liq ekanligini va sintetik ma’lumotlarning bu muammoni qisman hal qilishini ko‘rsatadi. Biroq, GANning yuqori hisoblash talablari kichik tashkilotlar uchun qiyinchilik tug‘dirishi mumkin.

Hisoblash resurslari bo‘yicha pruning hisoblash vaqtini 30% qisqartirdi (2.1 s dan 1.5 s gacha), lekin aniqlik 2% pasaydi. Bu natija resurslarni tejashda samarali bo‘lsa-da, aniqlik va tejamkorlik o‘rtasidagi muvozanat zarurligini ta’kidlaydi. Umumlashtirish qobiliyati bo‘yicha Dropout test aniqligini 88% dan 91% gacha oshirdi, overfittingni kamaytirdi. Bu ML modellarni yangi ma’lumotlarga moslashtirishda oddiy, lekin samarali usul ekanligini ko‘rsatadi.

Tushuntiriluvchanlik bo‘yicha SHAP model qarorlarini vizualizatsiya qilishda muvaffaqiyatli bo‘ldi, masalan, MNIST tasvirlarida piksel zichligining muhimligini aniqladi. Bu ishonch talab qilinadigan sohalarda foydali, ammo katta modellarda sekinlashishi cheklov sifatida qolmoqda. Axloqiy masalalarda federativ o‘qitish aniqlikni 87% gacha ta’minladi, maxfiylikni himoya qildi. Bu markazlashtirilgan modeldan past bo‘lsa-da, shaxsiy ma’lumotlarni saqlashda muhim yutuqdir.

Ushbu yechimlar MLning samaradorligini oshirishda salohiyatli ekanligini ko‘rsatdi. GAN ma’lumotlar yetishmasligini qoplaydi, pruning resurslarni tejaydi, Dropout umumlashtirishni yaxshilaydi, SHAP tushuntiriluvchanlikni ta’minlaydi, federativ o‘qitish esa axloqiy muammolarni hal qiladi. Biroq, har bir usulning o‘z cheklovlari bor: GAN va SHAP qo‘shimcha resurs talab qiladi, pruning aniqlikni pasaytiradi, Dropout optimal sozlashni talab qiladi, federativ o‘qitish esa barqarorlik muammosiga ega.

Tadqiqotning amaliy ahamiyati shundaki, ushbu yechimlar turli sohalarda qo‘llanilishi mumkin. Masalan, tibbiyotda GAN sintetik ma’lumotlar orqali kam uchraydigan kasalliklarni tahlil qilishga yordam beradi, pruning mobil qurilmalarda MLni qulaylashtiradi, federativ o‘qitish esa bemor ma’lumotlarini himoya qiladi. Shu bilan birga, cheklovlar – kichik dataset (MNIST) va cheklangan uskuna (16 GB RAM) – real dunyo sharoitlarida qo‘shimcha sinovlar zarurligini ko‘rsatadi.

Kelajakdagi tadqiqotlar uchun quyidagi yo‘nalishlar tavsiya etiladi:

1. GANning engil versiyalarini ishlab chiqish;
2. Pruningni adaptiv usullar bilan birlashtirish;
3. Dropoutni katta datasetlarda sinash;
4. SHAPni real vaqtda qo‘llash uchun optimallashtirish;

5. Federativ o‘qitishni ko‘proq mijozlar bilan sinash.

Xulosa qilib aytganda, ML sohasidagi muammolar murakkab bo‘lsa-da, taklif etilgan yechimlar ularni yumshatishda muvaffaqiyatli bo‘ldi. Kelajakda ushbu texnologiyalarni real sohalarga moslashtirish uchun fan va amaliyot hamkorligi muhimdir.

FOYDALANILGAN ADABIYOTLAR

- [1] Normamatov, X. (2025). IMPROVING THE METHODOLOGY OF TEACHING PROGRAMMING LANGUAGES BASED ON NETWORK TECHNOLOGIES. *International Journal of Artificial Intelligence*, 1(2), 656-662.
- [2] Normamatov, X. (2025). APPLYING INTERNATIONAL EXPERIENCES IN TEACHING PROGRAMMING TO HIGHER EDUCATION SPECIALIST STUDENTS: CHALLENGES AND SOLUTIONS. *International Journal of Artificial Intelligence*, 1(2), 648-650.
- [3] Normamatov, X. (2025). CHALLENGES AND SOLUTIONS IN TEACHING PROGRAMMING: AN EXPLORATION OF GLOBAL AND LOCAL PERSPECTIVES. *International Journal of Artificial Intelligence*, 1(2), 651-655.
- [4] Нормаматов, Х. М., & Абдуллаева, С. У. (2015). ЭФФЕКТИВНОСТЬ ПРИМЕНЕНИЯ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ УПРАВЛЕНИЯ" Э-БОЛЬНИЦА". In *Инновации в технологиях и образовании* (pp. 117-119).
- [5] Нормаматов, Х. М. (2014). ЛИНЕЙНЫЕ СИСТЕМЫ В ЦИФРОВОЙ ОБРАБОТКЕ СИГНАЛОВ. In *Инновации в строительстве глазами молодых специалистов* (pp. 239-241).
- [6] Шеров, Ж. Э., & Нормаматов, Х. М. (2015). АВТОМАТИЗАЦИЯ УПРАВЛЕНИЯ ВЫСШЕГО УЧЕБНОГО ЗАВЕДЕНИЯ. In *Инновации в технологиях и образовании* (pp. 178-182).