

PAST RESURSLI TILLARDA TIBBIY TRIAJ UCHUN BILIMLARNI SPETSIFIK-YO'NALTIRILGAN INTEGRATSIYALASH

Karimova Maftuna Bahodir qizi

Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti, Robototexnika va intellektual tizimlar kafedrasini, Toshkent

m.karimova@tuit.uz

1. Dolzarblik va muammoning qo'yilishi

Sog'liqni saqlashni raqamlashtirishning jadal rivojlanishi klinik jarayonlarni, jumladan, bemorlarning shikoyatlari asosida birlamchi tibbiy triajni avtomatlashtirish uchun yangi imkoniyatlar ochmoqda. Murojaatlarni tibbiy mutaxassisliklar bo'yicha avtomatik tasniflash registratura xodimlarining yukini sezilarli darajada kamaytiradi va bemorlarni yo'naltirish jarayonini tezlashtiradi.

Biroq, ushbu vazifa past resursli tillar sharoitida, xususan o'zbek tilida, alohida qiyinchilik tug'diradi: izohli klinik korpuslar deyarli mavjud emas va ixtisoslashgan til modellari ishlab chiqilmagan. Katta neyron tarmoq arxitekturalarini (BERT, T5) bunday sharoitda qo'llash maqsadga muvofiq emas, chunki ular katta hisoblash resurslarini va katta hajmdagi o'quv ma'lumotlarini talab qiladi [1, 2].

Ushbu ish o'zbek tilidagi klinik matnlar uchun tibbiy triajni avtomatlashtirish tizimini ishlab chiqish va baholashga bag'ishlangan. Asosiy tadqiqot savollari quyidagilardir: (1) klassik mashinaviy o'qitish usullari ushbu vazifada qanchalik samarali; (2) domen bilimlari integratsiyasi (NER) tasnif sifatiga qanday ta'sir ko'rsatadi; va (3) murakkab neyron arxitekturalarisiz tizim ishonchliligini oshirish mumkinmi.

Ishning amaliy ahamiyati O'zbekistonning raqamli sog'liqni saqlash tizimlariga to'g'ridan-to'g'ri joriy etish imkoniyati bilan belgilanadi - interpretatsiya qilinadigan va resurs tejimkor yechimlarga ehtiyoj ayniqsa yuqori bo'lgan muhitda [3].

2. Metodologiyani tavsifi

Eksperimental baholash 20 ta tibbiy mutaxassislik bo'yicha mutaxassislar tomonidan qo'lda izohlangan 317 ta o'zbek tilidagi bemor shikoyatlaridan iborat korpusda amalga oshirildi. Ma'lumotlar to'plami sezilarli sinf nomutanosibligi bilan tavsiflanadi: eng ko'p uchraydigan uchta sinf (gastroenterologiya, travmatologiya, LOR) umumiy hajmning 31 foizini tashkil etadi, 3 ta mutaxassislik esa 5 dan kam namuna bilan ifodalangan. Ma'lumotlar stratifikatsiyani saqlagan holda 80/20 nisbatida o'quv va sinov to'plamlariga bo'lingan.

Domen xos matn ishlovini qo'llab-quvvatlash uchun o'zbek tilida 65 ta «kasallik» va 192 ta «simptom» turidagi synchoslarni o'z ichiga olgan ixtisoslashtirilgan NER-lug'at tuzildi.

Tadqiqot har biri qo'shimcha bilim darajasini kiritadigan yettita model konfiguratsiyasini (M1–M7) ketma-ket baholashni o'z ichiga oladi:

- M1 - asosiy model: TF-IDF + logistik regressiya (domen bilimlarsiz);
- M2 - tuzilgan metama'lumotlar qo'shilishi (yosh guruhi, og'irlik darajasi);
- M3 - raqamli NER-belgilari (kasallik va simptom eslatmalarining chastotasi);
- M4 - NER-synchoslarni bilan matnni boyitish (semantik augmentatsiya);
- M5 - spetsifik-yo'naltirilgan boyitish: kam informatsiyali tibbiy atamalarni filtrlash, klinik ahamiyatli synchoslarni tanlash;
- M6 - ierarxik triaj: ikki bosqichli tasnif (makroguruh → mutaxassislik);

- M7 - ishonch moslashuvchan triaj: $P(y|x) < \tau$ bo'lganda bashorat qilishdan voz kechish va holatni qo'lda tekshirishga yo'naltirish.

Metodologiyaning asosiy hissasi - spetsifik-yo'naltirilgan bilimlarni integratsiyalash strategiyasi (M5) bo'lib, u noaniq kasallik eslatmalarini (masalan, «kasallik») istisno qiladi va faqat muayyan mutaxassisliklar bilan statistik korrelyatsiyaga ega bo'lgan klinik diskriminativ atamalarni saqlaydi.

3. Asosiy natijalar

Barcha yettita konfiguratsiyaning eksperimental baholash natijalari 1-jadvalda keltirilgan.

1.1 – jadval. Tibbiy triaj tizimi uchun model konfiguratsiyalarining aniqlik bo'yicha taqqoslanishi

Model	Tavsif	Bilimlar integratsiyasi	Aniqlik
M1	TF-IDF + logistik regressiya	Mavjud emas	0.68
M2	Matn + metama'lumotlar	Demografik belgilar	0.60
M3	Matn + NER-belgilari	Raqamli hisoblagichlar	0.57
M4	NER bilan boyitilgan matn	Kontekstual in'ektsiya	0.68
M5	Spetsifik-yo'naltirilgan boyitish	Filtrlangan sushnostlar	0.74
M6	Ierarxik triaj	Klinik guruhlash	0.78
M7	Iшонch moslashuvchan triaj	Chegara $\tau = 0.3$	0.80*

* Aniqlik ishonch chegarasi $\tau = 0.3$ da o'lchangan; holatlarning taxminan 6 foizi qo'lda tekshirishga yo'naltiriladi.

M1 asosiy modeli 0.68 aniqlikka erishdi va klassik ML usullarining ushbu vazifaga qo'llanilishini tasdiqladi. Metama'lumotlar (M2) va raqamli NER-belgilarini (M3) qo'shish aniqlikni mos ravishda 0.60 va 0.57 gacha pasaytirdi, bu esa domen bilimlarini naivona integratsiyalash semantik shovqin kiritishini ko'rsatadi.

Muhim natija - kasallik eslatmalari bo'lgan va bo'lmagan namunalarni qiyosiy tahlil qilishdir: kasallik eslatmalarisiz namunalar uchun tasnif aniqligi ≈ 0.77 ni tashkil etdi, noaniq atamali namunalar uchun esa atigi ≈ 0.42 . Bu esa modelning samaradorligini bilimlar miqdori emas, balki sifati belgilashini tasdiqlaydi.

Spetsifik-yo'naltirilgan strategiya (M5) kam informatsiyali atamalarni filtrlash orqali aniqlikni 0.74 gacha oshirdi. Ierarxik model M6 0.78 ga erishdi, xatolar tahlili esa ularning tasodifiy tarqalmasdan klinik jihatdan bog'liq mutaxassislik guruhlari ichida (kardiologiya–nevrologiya, pulmonologiya–kardiologiya) to'planishini ko'rsatdi. M7 ishonch mexanizmi $\tau = 0.3$ chegara qiymatida holatlarning atigi 6 foizini rad etib, 0.80 aniqlikni ta'minladi.

4. Xulosalar va amaliy ahamiyat

O'tkazilgan tadqiqot asosida quyidagi asosiy xulosalar shakllantirildi:

- AI-triajning samaradorligi arxitektura murakkabligiga emas, balki domen bilimlarini integratsiyalash usuliga bog'liq. Puxta ishlab chiqilgan belgilar muhandisligi bo'lgan interpretatsiya qilinadigan klassik modellar past resursli sharoitda neyron tarmoq yondashuvlari bilan raqobatlasha oladi.

- Tibbiy sushnostlarni naivona kiritish tasnif sifatini pasaytiradi. Klinik diskriminativ atamalarni tanlaydigan spetsifik-yo'naltirilgan filtrlash asosiy modelga nisbatan aniqlikni 9 foiz punktga oshiradi.
- Ierarxik tasnif xatolarni klinik asoslangan sohalarda tuziladi, bu esa real sharoitda xavfsiz qo'llash uchun printsiptial ahamiyatga ega.
- Ishonchni boshqarish mexanizmi minimal holatlarni rad etish bilan qamrov va aniqlik o'rtasida amaliy ahamiyatli muvozanatni ta'minlaydi.

Amaliy nuqtai nazardan, taklif etilayotgan yondashuv O'zbekiston raqamli sog'liqni saqlash tizimlariga bevosita integratsiya qilish uchun yaroqli yengil, interpretatsiya qilinadigan va kengaytiriladigan yechimni ifodalaydi. Katta til modellariga bog'liqlikning yo'qligi tizimni cheklangan hisoblash resurslari sharoitida - rivojlanayotgan mamlakatlardagi tibbiyot muassasalarining xarakterli belgisi - qo'llash imkonini beradi.

Keyingi tadqiqotlar korpusni kengaytirish, o'zbek tili uchun kontekstga bog'liq NER-modellarini ishlab chiqish va tizimni ko'p tilli klinik ma'lumotlar to'plamlari uchun moslashtirishni nazarda tutadi.

Adabiyotlar ro'yxati

1. Musaev M., Ochilov M., Abdullaeva M., Nasimov R. Restoring punctuation in Uzbek texts using LLM's fine-tuning approaches // Proceedings of ICFNDS'24. - New York: ACM, 2024. - P. 97–102.
2. Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding // Proceedings of NAACL-HLT. - 2019. - P. 4171–4186.
3. Abdusalomov A.B., Nasimov R. et al. Evaluating synthetic medical images using artificial intelligence with GAN algorithm // Sensors. - 2023. - Vol. 23. - 3440.
4. Stewart J. et al. Applications of NLP at emergency department triage: a narrative review // PLoS ONE. - 2023. - Vol. 18(12). - e0279953.
5. Névéol A. et al. Clinical NLP in languages other than English // Journal of Biomedical Semantics. - 2018. - Vol. 9(1). - P. 12.
6. Mengliev D., Barakhnin V., Abdurakhmonova N. Developing NER algorithms for Uzbek // Data in Brief. - 2024. - Vol. 54. - 110413.