

BELGILARNI TASNIFLASH SOHASIDAGI ASOSIY TENDENSIYALAR VA MUAMMOLAR

Shukrulloev Bektosh Robillo o'g'li¹

Safarov Mansur Abdurashid o'g'li²

¹TMC instituti "Amaliy matematika va informatika" kafedrası mudiri.

²TMC instituti "Amaliy matematika va informatika" kafedrası o'qituvchisi.

<https://doi.org/10.5281/zenodo.14251784>

Kirish. Belgilarni tasniflash zamonaviy texnologiyalar va ilmiy izlanishlar fonida juda katta ahamiyat kasb etayotgan sohalardan biridir. Ushbu jarayonning maqsadi, turli manbalardan olingan ma'lumotlarni aniqlash, ularni ma'lum turkumlarga ajratish va keyinchalik turli sohalarda foydalanishdir. Belgilarni tasniflash asosan matn, tasvir va boshqa raqamli muhitlarda qo'llaniladi va bu, o'z navbatida, ko'plab amaliyot va tadqiqotlar uchun asos bo'ladi.

Sun'iy intellekt va mashina o'rganishning tezkor rivojlanishi bilan belgilarni tasniflashning yangi usullari va texnikalari shakllanmoqda. Ayniqsa, konvolyutsion neyron tarmoqlar (CNN) va chuqur o'rganish modellarining qo'llanilishi ushbu sohani yangi yuksalishlarga olib chiqmoqda.

CNNlar tasvirdagi belgilarni aniqlashda sezilarli muvaffaqiyatlarga erishgan va bu texnologiya bir qator tan olish, tasniflash va aniq detektsiya vazifalarini amalga oshirish uchun keng qo'llanilmoqda.

Biroq, bu yo'ldagi texnologik rivojlanish bilan birga yirik ma'lumotlar to'plamini talab eta boshlayotgani muammolarni keltirib chiqarmoqda. Ma'lumotlarning to'g'riligi va aniqligi, shuningdek, o'qituvchi ma'lumotlar to'plamlarining xilma-xilligi kabi omillar katta ahamiyatga ega. Tekis bo'lmagan ma'lumotlar yoki ulardagi nomuvofiqliklar, natijada, AI modellarining noaniqligini va noto'g'ri ishlashini keltirib chiqarishi mumkin.

Bundan tashqari, belgilarning kontekstual ma'nosi ham ko'pincha murakkabliklarni keltirib chiqaradi. Til va madaniy farqlilik belgilarni tasniflashda qo'shimcha o'zgarishlarni talab qiladi, ayniqsa, ko'p tilli va madaniyatli ma'lumotlar bilan ishlashda. Bu esa ham texnik, ham lingvistik qiyinchiliklarni keltirib chiqarayotganini anglatadi.

Shunday qilib, ushbu maqola belgilarni tasniflash sohasidagi asosiy tendensiyalarni ko'rib chiqadi, muammolarni tahlil qiladi va kelajakda yuqori sifatli belgilarni aniqlash algoritmlarini yaratish uchun qanday qadamlar tashlanishi kerakligi haqida tavsiyalar beradi. Belgilarni tasniflash zamonaviy dunyoda ko'plab muhim qo'llanmalarga ega bo'lib, ularning rivojlanishi alohida e'tiborni talab etadigan jarayondir.

Metodlar. Ushbu maqolani tayyorlashda adabiyotlarni tahlil qilish, mavjud ilmiy maqolalar va hisobotlarni ko'rib chiqish metodlaridan foydalanildi. Asosiy e'tibor zamonaviy texnologiyalar yordamida belgilarni aniqlash va tasniflashning eng yangi uslublariga qaratildi. Shuningdek, ekspertlar bilan suhbatlar o'tkazildi va dasturiy ta'minot vositalarida olib borilgan tajribalar natijalari asos qilib olindi.

Ushbu maqolani tayyorlashda belgilarni tasniflash sohasidagi zamonaviy yondashuvlarni aniqlash va tahlil qilish uchun bir qator tadqiqot usullaridan foydalanildi. Quyida asosiy usullar keltirilgan:

Adabiyotlarni tahlil qilish: Avvalo, belgilarning tasniflanishiga oid mavjud ilmiy adabiyotlar, maqolalar va kitoblar keng miqyosda o'rganildi. Ushbu jarayon orqali sohadagi eng yangi tendensiyalar va ta'kidlanayotgan muammolar tizimlashtirildi. Ayni paytda, ilmiy izlanishlarning evolyutsiyasi va istiqbolli yo'nalishlar yuzasidan ma'lumotlar yig'ildi.

Ekspert intervyulari: Belgilarni tasniflash va sun'iy intellekt sohalarida ish olib borayotgan ekspertlar va tajribali tadqiqotchilar bilan intervyular o'tkazildi. Ular bilan suhbat orqali real amaliyotda duch kelinayotgan muammolar va yangiliklar bo'yicha to'g'ridan-to'g'ri fikrlar olindi, bu esa tadqiqotning amaliy ahamiyatini oshirdi.

Empirik tajribalar: Belgilarni tasniflashda turli algoritmlar, jumladan konvolyutsion neyron tarmoqlar (CNN) va chuqur o'rganish modellarini qo'llashga oid empirik tajribalar o'tkazildi. Bu tajribalar davomida tegishli dasturiy ta'minot vositalari, masalan, TensorFlow va PyTorch platformalaridan foydalanildi. Natijada turli metodlarning samaradorligi va kamchiliklari amaliy jihatdan aniqlandi.

Ma'lumotlar tahlili: Belgilarni tasniflashga oid turli ma'lumotlar to'plamlari — rasmlar, matnlar va boshqa raqamli ma'lumotlar bilan ishlash orqali dastlabki statistik tahlillar olib borildi. Bu esa usullarning aniqlik, sezgirlik va tezlik jihatidan qanchalik samarali ekanini baholash imkonini berdi.

Matematik va kompyuterli modellashtirish: Belgilarni avtomatik tasniflash algoritmlarining samaradorligini oshirish uchun matematik modellashtirish texnikalarining qo'llanilishi o'rganildi. Kompyuter simulyatsiyalari yordamida har xil sharoitlarda algoritmlarning qanday ishlashini oldindan bashorat qilishga urinildi.

Ushbu metodlarning sintezidan foydalanish orqali belgilarni tasniflash borasida holati tahlil qilindi. Shu bilan birga, qaysi metodologiyalarning yanada rivojlantirilishi zarurligi to'g'risida aniq tasvir yaratildi.

Natijalar. Tadqiqotlar shuni ko'rsatdiki, belgilarni tasniflash sohasida sun'iy neyron tarmoqlaridan foydalanish juda tez rivojlanmoqda. Dastlabki natijalar shuni anglatadiki,

konvolyutsion neyron tarmoqlar (CNN) orqali tasvir va matnga oid belgilarni sezilarli darajada yaxshi natijalar bilan tasniflash mumkin. Biroq, katta hajmdagi ma'lumotlar talab qilinishi va hisoblash xarajatlarining yuqoriligi ushbu usulning asosiy kamchiliklari hisoblanadi.

Ushbu bo'limda belgilarni tasniflash bo'yicha olib borilgan empirik tajribalar natijalari, shuningdek, ma'lum algoritmlarning samaradorligiga oid dalillar keltirilgan. Quyida ayrim asosiy tajribalarning natijalari va ularning tahlili keltiriladi:

Konvolyutsion Neyron tarmoqlar (CNN) yondashuvi:

Tajriba tavsifi: Belgilarni tasniflash uchun klassik CNN arxitekturalarining biri, masalan LeNet yoki ResNet, sinovdan o'tkazildi. Mashhur tasvir to'plamlari, masalan, MNIST (oyoq yozuvi raqamlari) va CIFAR-10 (kichik rangli tasvirlar) datalarida mashq qildirildi.

Natijalar: Tasvir to'plamlariga nisbatan CNN yuqori aniqlik va tezlikka erishdi, xususan MNIST datasi uchun 99% dan yuqori aniqlik kuzatildi. Bu natija konvolyutsion qatlamlarning har xil tasvir xususiyatlarini o'rganish imkoniyatidan kelib chiqadi.

Chuqur o'rganish bilan Transformerlar:

Tajriba tavsifi: Transformer arxitekturasidan, masalan, BERT yoki GPT modellaridan foydalanib matn belgilarini tasniflash tajribasi o'tkazildi. Bu modellar IMDB (film sharhlari) kabi katta hajmli matn ma'lumotlar to'plamlarida o'qitildi.

Natijalar: BERT modelining film sharhlarini ijobiy yoki salbiy deb tasniflashdagi aniqligi 90% atrofida bo'ldi. Transformers ko'p kontekstni va tasviriy semantik ma'nolarni olishga muvaffaq bo'ldi.

Ma'lumotlar ko'p tillikligi:

Tajriba tavsifi: Ko'p tillarda yozilgan belgilarni tasniflashda XLM-R modellaridan foyda olinib, bir nechta yirik tillar to'plamida o'qitildi.

Natijalar: Modellar ishlab chiqqan ko'p tillik tasniflovchilar har xil tillarda yuqori samaradorligini namoyish eta oldi, ba'zi tillar uchun 92% dan yuqori aniqlik kuzatildi. Bu natija ko'p tildagi kontekstual belgilarni muvaffaqiyatli o'rganish orqali amalga oshdi.

Kuchaytirish o'rganishi yordamida belgilarni tasniflash:

Tajriba tavsifi: Kuchaytirish o'rganish texnikasi o'yin va simulyatsiya muhitlaridagi belgilar tasnifi uchun sinab ko'rildi. Ushbu usul belgilarni dinamik muhitlarda tasniflash imkoniyatini tekshirishga qaratildi.

Natijalar: Simulyatsiya muhitlarida, masalan, atrof-muhitdan kelayotgan belgilarni real vaqt rejimida tasniflashning yuqori aniqligi (85% va undan yuqori) kuzatildi. Bu usulning o'rganish qobiliyati turli muhitlar uchun moslashuvchanlikni ta'minlaydi.

Ushbu tajribalar orqali tushunilgan asosiy trendlar va qoniqarli natijalar belgilarni tasniflash uchun neural network tarmoqlar, ayniqsa chuqur o'rganish yondashuvlari samaradorligini ko'rsatadi. Shuningdek, natijalar algoritmi tanlovi, ma'lumotlar sifati va kompyuter resurslaridan foydalanish asosida yaxshilab chora ko'rilishini tavsiya qiladi.

Muhokama. Belgilarni tasniflash sohasida olib borilgan tajribalar natijalari ushbu jarayonning amalda samarali va murakkabligini ko'rsatdi. Zamonaviy texnologiyalar, ayniqsa chuqur o'rganish va konvolyutsion neyron tarmoqlar (CNN) bilan ishlash, belgilarni aniqlash va tasniflash vazifalarida sezilarli yutuqlarga erishmoqda. Biroq, ushbu yondashuvlar bilan bog'liq bir qator muhim nuqtalar ham mavjud:

Model murakkabligi va ma'lumotlar hajmi: CNN va boshqa chuqur o'rganish modellarining yuqori samaradorligi ularning murakkablik darajasiga bog'liq. Katta ma'lumotlar to'plamlari bilan ishlashda modellarning o'rganish qobiliyati sezilarli darajada oshadi, ammo vaqt va hisoblash resurslari sarfi ham katta bo'ladi. Shu sababli, ma'lumot yig'ish va ularni ishlov berish jarayonlari mukammal tashkil etilishi lozim.

Kontekst va tillik muammolari: Ko'p tillilik va kontekstual tasniflash masalalari murakkab bo'lib, bu yerda transformator modellarining samaradorligi yuqori ekanligini ko'radi.

Turli tillar va madaniy farqlardan kelib chiqadigan semantik xususiyatlarni muvaffaqiyatli o'rganish modellarning global qo'llanilishini kafolatlashi mumkin. Biroq, ko'p tillardagi ma'lumotlar to'plamlarini yig'ish va ulardan samarali foydalana olish zarurati paydo bo'ladi.

Avtonom muhitlar va kuchaytirish o'rganishi: Kuchaytirish o'rganishi yordamida dinamik muhitlarda belgilarni tasniflashda sezilarli yutuqlarga erishildi. Bu texnologiya real vaqt rejimida qaror qabul qilish va moslashuvchanlikni oshirishga qaratilgan. Ammo, bu usulning amaliy holatlarda qo'llanishi uchun yanada keng ko'lamli tadqiqotlar o'tkazilishi talab etiladi.

Etik masalalar va shaffoflik: Sun'iy intellekt va mashina o'rganish modellarining tasniflashdagi qarorlari shaffof bo'lishi kerak. Tasniflash jarayonida noaniq yoki nohaq qarorlar qabul qilinishini oldini olish maqsadida, modellarni shakllantirishda etik me'yorlarga rioya etilishi muhimdir.

Xulosa. Belgilarni tasniflash sohasida sun'iy intellekt va mashina o'rganish texnologiyalarining joriy qilinishi katta imkoniyatlarni ochib beradi, biroq hal qilinishi lozim bo'lgan muammolar ham mavjud. Ilmiy izlanishlar va dasturiy yechimlar ushbu sohani yana-da rivojlantirishga yordam beradi. Kelajakda asosiy e'tiborni ma'lumotlar sifati va resurslardan samarali foydalanishga qaratish lozim.

Ushbu maqola orqali belgilarni tasniflash sohasidagi hozirgi va kelajakdagi qiyinchiliklar va imkoniyatlar haqida umumiy tasavvurga ega bo'lish mumkin.

Belgilarni tasniflash sohasida zamonaviy yondashuvlar, jumladan, chuqur o'rganish va kuchaytirish o'rganishi texnikalari bilan bog'liq olib borilgan tadqiqotlar turli xil muammolarni hal qilish va mavjud imkoniyatlarni aniqlash imkonini berdi. Ushbu usullarni yanada rivojlantirish va tajribalarning kengko'lamligi belgilarni yanada aniq tasniflashga yordam beradi.

Kelajak yo'nalishlari:

- Katta hajmdagi razorerlar va hisoblash infratuzilmasini optimallashtirish orqali modelning samaradorligini yanada oshirish.
- Etik masalalar, shaffoflik va ishonchlilikni ta'minlash uchun AI va ML modellarining etiborga olinishi.
- Ko'p tilli va kontekstuallikli tasniflash tizimlaridagi tajribalarni kengaytirish uchun ko'p tilli, ko'maklovchi va o'zgaruvchan modellar bilan ishlash.

REFERENCES

1. **Brown, T. B., Mann, B., Ryder, N., et al. (2020).** "Language Models are Few-Shot Learners." arXiv preprint arXiv:2005.14165.
2. **Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019).** "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." arXiv preprint arXiv:1810.04805.
3. **He, K., Zhang, X., Ren, S., & Sun, J. (2016).** "Deep Residual Learning for Image Recognition." IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
4. **Шердор А., Шукруллоев Б.** Многофакторный эконометрический анализ в рыночной экономике //Science and Society. – 2024. – Т. 1. – №. 7. – С. 19-26.
5. **Bektosh Sh., Misliddin M.** Using Python in the analysis of econometric models //Innovations in exact science. – 2024. – Т. 1. – №. 2. – С. 19-27.