



IJTIMOYIY TARMOQ MATNLI YOZISHMALARINI TASNIFLASH TEXNOLOGIYALARI

A.A.Taniberdiyev

"Information Technologies" Department of Gulistan State
University, Doctor of Philosophy in Pedagogical sciences (PhD)

S.F.Saidova

Master's student of Gulistan State University "Information
Technologies" department

<https://doi.org/10.5281/zenodo.14294316>

ARTICLE INFO

Received: 29 th November 2024

Accepted: 1st December 2024

Published: 7 th December 2024

KEYWORDS

*ijtimoiy tarmoq, matnli
yozishmalari, taxlil, internet
tarmoqlarida, messages, comments, nlp,
bert, gpt.*

ABSTRACT

Ushbu maqolada bugungi kunda ijtimoiy tarmoqlardagi axborotlar, global axborot oqimlari va ularni taxlil qilish algoritmlari, bu boradagi tadqiqotlar va ularning samaralari o'rganilgan. So'ngi vaqtlardagi samaraliy algoritmlar keltirilgan va ulardan foydalanish tartibi ko'rsatilgan.

Bugungi kunda ijtimoiy tarmoqlardagi axborotlar, global axborot oqimining asosiy qismini tashkil qilmoqda. Ijtimoiy tarmoq matnli yozishmalari deganda, odatda ijtimoiy tarmoqlar orqali amalga oshiriladigan yozma shakldagi muloqotlar, postlar, izohlar yoki xabarlar tushuniladi. Bu turdagi yozishmalar turli formatlarda bo'lishi mumkin, masalan, qisqa xabarlar, maqolalar, sharhlar yoki javoblar. Bu axborotlarni qayta ishlash keng miqyosida jadallashib bormoqda. Sababi katta axborotlar hajmiga ega internet tarmoqlarida yetarlicha taxlil o'tkazish axborotlarning ishonchiligi va qilmmatliligini belgilashda muhim vositadir.

Ijtimoiy tarmoq matnli yozishmalarining asosiy turlari. Xabarlar (Messages), Shaxsiy muloqot (masalan, Facebook Messenger yoki Instagram Direct orqali), Guruh suhbatlari (masalan, WhatsApp yoki Telegram guruhlari), Postlar, ijtimoiy tarmoqda o'z fikrlarini keng auditoriyaga yetkazish uchun joylashtirilgan yozuvlar (Twitter, Facebook, yoki LinkedIn postlari). Izohlar (Comments), Postlar yoki kontentga bo'lgan javoblar yoki fikrlar. Masalan, biror video ostida qoldirilgan sharhlar. Teglar va muhokamalar (Mentions and Discussions). Boshqa foydalanuvchilarni eslatish yoki muhokamalarda qatnashish. Heshteg va kontekstual yozishmalar (#Hashtags). Ma'lum bir mavzuga tegishli bo'lgan qisqa yozuvlar yoki mavzularni belgilovchi kalit so'zlar. Forum va guruh yozishmalari. Facebook guruhlari, Reddit yoki boshqa forumlardagi kengroq muhokamalar. Ijtimoiy tarmoq yozishmalarining o'ziga xos xususiyatlari, Matnning qisqalik darajasi. Twitter kabi platformalarda postlar qisqa bo'lishi lozim (masalan, 280 ta belgigacha). Shu bilan birga, uzun maqola formatida yozishmalar ham bo'lishi mumkin (masalan, LinkedIn yoki Facebook maqolalari). Norasmiy til va qisqartmalar. Ko'pincha oddiy va norasmiy til ishlatiladi. Ma'lumotning xilma-xilligi. Yozishmalar turli mavzularni o'z ichiga oladi: shaxsiy hayot, siyosat, madaniyat, reklama va boshqalar. Real vaqtda muloqot. Ko'p hollarda yozishmalar real vaqt rejimida sodir bo'ladi (masalan, chatlar yoki translyatsiya davomida izohlar). Ko'pkanallilik. Matn boshqa formatlar bilan aralash bo'lishi mumkin (rasmlar, videolar yoki ovoqli xabarlar bilan birga).

Ijtimoiy tarmoq yozishmalarini o'rganish ko'plab maqsadlarda qo'llaniladi. Foydalanuvchilar brend haqida qanday fikrda ekanligini aniqlash. Sentiment tahlili. Matnning ijobiy, salbiy yoki neytral ekanligini bilish. Trendlarga aniqlik kiritish. Hozirgi zamonda dolzarb bo'lgan mavzularni aniqlash. Spam yoki zararli kontentni filtrlash. Noto'g'ri yoki xavfli yozishmalarni bloklash va boshqalar.

Matnli yozishmalarni tasniflash algoritmi – bu matnni oldindan belgilangan kategoriyalarga yoki sinflarga ajratish jarayonidir. Ushbu algoritmlar tabiiy tilni qayta ishlash (NLP) texnikalari yordamida amalga oshiriladi. Quйда matnli yozishmalarni tasniflash jarayonini tushunish va amalga oshirish uchun zarur bo'lgan asosiy elementlar keltirilgan. O'z navbatida tasniflash jarayoni birqator bosqichlarni o'z ichiga oladi. Buna tasniflash algoritmi bosqichlari mavjud. Ma'lumotni yig'ish va tayyorlash: ijtimoiy tarmoqdan yozishmalarni to'plash, matnlarni tayyorlash, tozalash, normalizatsiya, tokenizatsiya.

Matnni raqamli formatga aylantirish. Matnni raqamlar bilan ifodalash uchun quyidagi usullardan foydalaniladi: Bag of Words (BoW)-matnda uchraydigan so'zlar sonini hisoblash. TF-IDF (Term Frequency-Inverse Document Frequency)-so'zning matn ichidagi va butun ma'lumotlar to'plamidagi ahamiyatini baholash. Embedding-matnni vektor ko'rinishida ifodalash (masalan, Word2Vec, GloVe, yoki Transformer asosli embeddinglar).

Bu turdagi axborotlar bilan ishlash uchun maxsus algoritmlar talab etiladi. Bu algoritmlar axborotning turiga qarab tanlash lozim bo'ladi. Algoritmni tanlash va o'rgatish.

Oddiy algoritmlar. Bularga Naive Bayes (matnni tasniflashda tez va samarali ishlaydi) va Logistic Regression (tasniflash uchun mos, chiziqli model).

Kuchli mashina o'qitish algoritmlari. Support Vector Machines (SVM)-kichik o'lchamdagi ma'lumotlarda yuqori aniqlikka ega. Random Forest yoki XGBoost: Murakkab tasniflash muammolarida ishlatiladi.

Chuqur o'qitish (Deep Learning) algoritmlari. RNN va LSTM (Uzoq bog'lanishli ketma-ketlikli matnlarni tahlil qilishda yaxshi). Transformer modellar (masalan, BERT, GPT) matn kontekstini chuqur tushunish uchun.

Modelni baholash. Tasniflash modelining samaradorligini quyidagi mezonlar orqali baholash mumkin Aniqlik (Accuracy), To'g'rilik (Precision) va qamrov (Recall).

Ma'lumotlarni qayta ishlashda transformer asosidagi modellar mavju bo'lib bularga BERT, GPT larni keltirish mumkin. BERT, GPT matn tasniflashda yuqori samaradorlikka ega. BERT kontekstual ma'lumotni chuqur tahlil qiladi. Masalan, "Men xursandman" va "Men xursand emasman" iboralaridagi farqni aniqlaydi. GPT Matnlarni oldindan o'rgatish va tushunish uchun qo'llaniladi.

Ijtimoiy tarmoq matnli yozishmalarini tasniflash uchun algoritm tanlash jarayonida ma'lumot hajmi, aniqlik darajasi talablari va resurs imkoniyatlarini hisobga olish muhim. Olingan katta hajmdagi ma'lumotlarni taxlil qilishda to'g'ri algoritmlarni tanlash avvalo ma'lumotlarni taxlil qilish tezligini oshiradi. Bulutliiy hotirada ma'lumotlarni taxlil qilish sifati va tezligi bevosita to'g'ri tanlangan algoritmga bog'liq.

Adabiyotlar.

1. Fundamentals of compilation of electronic tasks for students to test and strengthen their knowledge of database(Article). Toshtemirov, D., Muminov, B., Saidov, J. International Journal of Scientific and Technology Research. Volume 9, Issue 4, April 2020, Pages 3226-3228. <https://www.ijstr.org/paper-references.php?ref=IJSTR-0120-29952>

2. Tools and Methods of Formation of Professional Competence of Future Teachers of Computer Science and Information Technologies. SaidovJasur Doniyor o'g'li, Baxodirov Muzrob Doniyor o'gli, Normatova Malika Norkulovna, Mavlonov Sherzod Hazratqulovich. Psychology and Education Journal. Vol. 58 No. 2 (2021): Volume 58 No. 2 (2021) <https://doi.org/10.17762/pae.v58i2.2282>
3. Baxtiyor Baxriddinovich Ergashev, Jasur Doniyor Ogli Saidov, & Sait Xalilovich Islikov (2021). BO'LAJAK INFORMATIKA VA AXBOROT TEXNOLOGIYALARI O'QITUVCHILARI KASBIY KOMPETENTLIGINI SHAKLLANTIRISH VOSITALARI VA METODLARI. Academic research in educational sciences, 2 (2), 1139-1146. doi: 10.24411/2181-1385-2021-00312
4. N.A.Muslimov, K.M. Abdullayeva, O.A.Kuysinov, N.S.Gaipova, N.N.Karimova, M.Kodirov. - Kasb ta'limi o'qituvchilarining kasbiy kompetentligini shakllantirish texnologiyasi. —T.: «Fan va texnologiya», 2013, 128 bet.
5. 3. Y.Y.Chicherina, D.A.Nurkeldiyeva, D.B.Yakubjanova-Mutaxassislik fanlarini o'qitish metodikasi. O'quv qo'llanma Toshkent, 2013
6. Djurayev, M., Husenov, J., & Qurbonova, C. (2024). GEOGRAFIYA DARSLARIDA ZAMONAVIY O'YIN TEXNOLOGIYALAR. Журнал академических исследований нового Узбекистана, 1(5), 67-70.
7. Djurayev, M., Qurbonova, C., & Husenov, J. (2024). GEOGRAFIYA FANINING MAKTABLARDA O'TILISH SIFATI. Журнал академических исследований нового Узбекистана, 1(5), 71-73.
8. Наримова, Г. А., Сиддикова, Н., & Саматова, Г. (2015). Различные подходы в преподавании иностранного языка студентам в неязыковых вузах. Научный альманах, (11-2), 292-296.
9. Musayeva, S. I. (2024). ADVANTAGES AND DISADVANTAGES OF INTERACTIVE METHODS INTRODUCING THE CLT (COMMUNICATIVE LANGUAGE TEACHING) APPROACH. World Scientific Research Journal, 32(1), 217-219.
10. Musayeva, S. I., & Mengliyeva, S. S. (2022). Kursantlarning madaniyatlararo rivojlantirish.
11. Shaimova, G. A., Shavkieva, D., & Abdukadyrova, N. (2014). The Importance of WebQuest Technologies in Formation of the Professional Speech in Future Specialists. In Young Scientist USA (pp. 160-162).