

JAMOATCHILIK FIKRINI O'RGANISH YOKI SENTIMENT TAHLIL

Allanazarova Sabohat Yusupboyevna

allanazarovasabohat@gmail.com

Mirzo Ulug'bek nomidagi O'zbekiston Milliy Universiteti magistranti

Annotatsiya. Dunyo bo'ylab ma'lumotlar hajmining keskin oshishi natijasida, ularni mavzuiy tasniflash, guruhlariga ajratish orqali kerakli va keraksiz ma'lumotlarni farqlay olishimiz mumkin. Bunda Sentiment (hissiy) tahlili yordam beradi. Avtomatik o'rgatilgan dastur asosida har qanday sohaga mansub bo'lgan matnlarni oson va tezda tahlil qilish mumkin. Dunyo tajribasidagi yutuqlardan samarali foydalanish hamda o'zbek tilining xususiyatlarini hisobga olgan holda tizim ishlab chiqish orqali ko'pgina sohalar rivojiga erishish mumkin. Ushbu maqolada matnlarni hissiy jihatdan o'rganish masalasi muhokama qilinadi.

Kalit so'zlar: *hissiy tahlil, tabiiy tilni qayta ishlash, ijtimoiy tarmoqlarni tahlil qilish, Twitter.*

Tabiiy tilni qayta ishlash sohasida mashina o'rganish algoritmlari bo'yicha eng so'nggi yutuqlardan foydalanilib, yuqori yutuqlarga erishilmoqda. Twitterda NLTK bazasida minglab - minglab ijtimoiy tarmoq foydalanuvchilarining sharhlari mavjud bo'lib, ijobiy va salbiy ko'rinishda annotatsiyalab chiqilgan. Baza 5000 ta ijobiy va 5000 ta salbiy tvitlarni o'z ichiga olgan bo'lib, asosan sentiment (hissiy) tahlilda korpus sifatida qo'llaniladi. Tuyg'ularni tahlil qilishning tobora ortib borayotgan ahamiyati sharhlar, forum munozaralari, bloglar, mikro-bloglar, Twitter va ijtimoiy tarmoqlarning o'sishiga sabab bo'lmoqda. Mikrobloglar bugungi kunda Internet foydalanuvchilari orasida juda mashhur aloqa vositasiga aylandi [1]. Ijtimoiy tarmoqda kundalik hayotdagi turli xil voqealar va mavzular bo'yicha fikrlar bilan bo'lishishadi va dolzarb masalarni muhokama qilishadi. Shuning uchun mikrobloglar veb-saytlari fikrlarni o'rganish va hissiyotlarni tahlil qilish uchun boy ma'lumotlar manbayi hisoblanadi. Ushbu ma'lumotlarga asoslanib, hissiyotlarni yanada aniqroq bilib olishimiz mumkin. Tuyg'ularni tahlil qilish va fikrlarni o'rganish - bu yozma tildan odamlarning fikrlari, baholari, munosabati va hissiyotlarini tahlil qiladigan o'rganish sohasi bo'lib, dunyo bo'ylab sentiment tahlil deb nomlanadi. Bu tabiiy tilni qayta ishlashning eng faol tadqiqot yo'nalishlaridan biri hisoblanadi [2]. Darhaqiqat, hissiy tahlil tadqiqotlari biznes va

umuman jamiyat uchun muhimligi sababli kompyuter fanidan tashqaridagi sohalar va ijtimoiy fanlarga tarqaldi va nihoyatda keskin darajada o'sib bormoqda. Fikrlar deyarli barcha inson faoliyati uchun markaziy o'rin tutadi va bizning xatti-harakatlarimizning asosiy mezonidir. Bizning e'tiqodimiz va haqiqatni anglashimiz va tanlagan qarorlarimiz asosan boshqalarning dunyoni qanday ko'rishlari va baholashlariga bog'liq [3]. Shu sababli, biz qaror qabul qilishimiz kerak bo'lganida, ko'pincha boshqalarning fikrlarini bilishga qiziqamiz. Bu nafaqat jismoniy shaxslarga, balki tashkilot va korxonalarga ham tegishli. Masalan, ishlab chiqaruvchi kompaniyalarni quyidagi savollar qiziqtirishi mumkin:

- Odamlar bizning mahsulotimiz (xizmat, kompaniya va boshqalar) haqida qanday fikrda?
- Odamlar bizning mahsulotimizga nisbatan qanchalik ijobiy (yoki salbiy) munosabatda?
- Odamlar bizning mahsulotimiz qanday bo'lishini afzal ko'rishadi?

Bunda foydalanuvchilar mahsulot bilan bog'liq ijobiy yoki salbiy axborot berishi yoki shunchaki betaraf munosabat bildirishi mumkin. Aspektga asoslangan hissiyotlarni tahlil qilishda, siz odamlarning qaysi jihatlari yoki xususiyatlarini ijobiy, betaraf yoki salbiy tarzda eslatib o'tayotganini bilishni xohlaysiz. Masalan, ushbu matnda: "Ushbu kamera batareyasining ishlash muddati juda qisqa", aspektga asoslangan klassifikator bu jumlaning batareyaning ishlash muddati haqida salbiy fikr bildirishini aniqlay oladi.

Dunyo miqyosida ko'pgina olimlar Sentiment tahlili bo'yicha turli tadqiqotlar bilan shug'ullanib kelganlar va hozirda ham shug'ullanmoqdalar. Ingliz, turk, qozoq, rus va boshqa tillardagi hissiyotlarni tahlil qilish uchun bir qator manbalar va vositalar ishlab chiqilgan. Xususan, ingliz tilidagi matnlarni fikrlar asosida tahlil qilish juda mashhur va yaxshi o'rganilgan mavzu hisoblanadi. Ingliz tili uchun ko'plab qutblanish manbalari mavjud. Masalan: SentiWordNet (SWN), SenticNet (SN) va NRC Emotion Lexicon kabi. Turkiy tillar oilasiga mansub bo'lgan Turk va Qozoq tillari bu sohada ancha taraqqiy etdi.. O'zbek tilida sentiment tahlili bo'yicha O'zbekiston Milliy Universitetida Sanatbek Matlatipov magistrlik ishida, Samarqand Davlat Universitetida Ilyos Rabbimov boshchiligidagi bir guruh izlanuvchilar o'zbek filmlariga qoldirilgan sharhlarni emojilar asosida o'rganishdi. O'zbek tilidagi matnlar uchun sentiment tahlili dasturini ishlab chiqishda mazkur tilning xususiyati hisobga olinishi zarur. O'zbek tili ko'p tilli hissiy so'zlar to'plamini o'z ichiga oladi. Hissiy jihatdan qutblanish manbalari yaratilmagan tillar uchun eng oddiy yechim ingliz tilidagi manbalarni

tarjima qilish bo'lishi mumkin. Ammo agglyutinatив harakterga ega bo'lgan o'zbek tili uchun bu to'g'ri yechim emas [4]. O'zbek tilida ibora, kinoya va piching, kesatirlardan keng foydalaniladi. Xususan, kinoya haqida gap ketganda, odamlar o'zlarining salbiy his-tuyg'ularini ijobiy so'zlardan foydalangan holda ifoda etadilar. Bu esa mashinaga tahlil qilish jarayonida qiyinchilik tug'diradi. Fikrlarni avtomatik tasniflashning tabiiy tilni qayta ishlash (NLP), mashina o'rganish algoritmlariga asoslangan klassifikasiya, logistik regressiya (LG), qo'llab-quvvatlash vektor mashinalari (SVM), Bayes klassifikasiyasi, neyron tarmoqlari kabi ko'plab metodikalari adabiyotlarda tavsiya etiladi.

Hissiyotlarni avtomatik aniqlash muammosining uchta asosiy yechimi mavjud:

1. Avtomatik yondashuv: mashina o'rganish texnikasiga tayanadi. Maxsus ishlab chiqilgan dastur asosida ishlaydi;

2. Qoidalarga asoslangan yondashuv. Ushbu qoidalar NLP texnikalarini o'z ichiga oladi. Bu tizim juda sodda bo'lib, berilgan matnda paydo bo'lgan ijobiy va salbiy so'zlar sonini sanaydi. Agar ijobiy so'zlarning soni salbiy so'zlarning sonidan ko'p bo'lsa, tizim matnni ijobiy deb topadi. Agar raqamlar juft bo'lsa, tizim neytral deb baholaydi.

3. Gibril: kompyuterlarga yordam tariqasida inson tomonidan bajariladigan jarayon. Bunda qoidalarga asoslangan va avtomatik texnikaning kerakli elementlarini bitta tizimga birlashadi. Ushbu natijaning katta foydasi shundaki, natijalar ko'pincha aniqroq bo'ladi.

Xulosa qilib aytganda ijtimoiy tarmoqlarning tez sur'atlarda o'sib borishi shu qadar kengaydiki, ma'lumotlar hajmining har o'n sakkiz oyda ikki barobar oshishini ko'rsatmoqda. Fikrlarni tahlil qilish deb ham ataladigan sentiment tahlili tabiiy tilni qayta ishlashning eng faol tadqiqot yo'nalishlaridan biriga aylandi. Uning qo'llanilishi, biznes xizmatlaridan tortib, siyosiy kompaniyalargacha keng tarqalgan. Axborot urushlari paytida bu vazifa eng muhim vazifalardan biridir XXI asrda jamoatchilik fikri hamma narsadan ustundir. U bilan hech narsa muvaffaqiyatsiz bo'lmaydi; usiz hech narsa muvaffaqiyatga erisha olmaydi.

Foydalanilgan adabiyotlar ro'yhati:

1. Гедранович, Б.А., Гедранович, А.Б. Отношение К Высшему Образованию: Сентимент-Анализ Данных Микроблогов. Инновационные Образовательные Технологии.-2013.-№ 1 (33).-С. 46-54.

2. Dehkharghani R., Saygin Y., Yanikoglu B. et al. SentiTurkNet: a Turkish



polarity lexicon for sentiment analysis. Language Resources and Evaluation. - 2016. 50. -P. 667–685

3. Bing Liu Sentiment Analysis and Opinion Mining Synthesis Lectures on Human Language Technologies, May 2012, Vol. 5, No. 1 , Pages 1-167 (<https://www.doi.org/10.2200/S00416ED1V01Y201204HLT016>)

4. Matlatipov S., Kuryozov E., Miguel A. A., Corlos-Rodriguez. Deep learning vs. classic models on a new uzbek sentiment analysis dataset. Conference: 9th language & technology conference: Human language technologies as a challenge for computer science and linguistics Poznan, -Poland– 2019. -P. 258-262