

UZBEKISTAN

O'ZBEKISTON

LANGUAGE & CULTURE

TIL VA MADANIYAT

**KOMPYUTER
LINGVISTIKASI**

ISSN 2181-922X

2023 Vol. 1 (6)

www.compling.tsuull.uz

MUNDARIJA

Eşref Adalı

Corpus for what.....6

Victor Zakharov

Functionality of the russian national corpus.....18

Botir Elov, Dilrabo Elova

NLPda koreferens masalasi.....27

Botir Elov, Shahlo Hamroyeva, Oqila Abdullayeva,

Zilola Xusainova, Nizomaddin Xudayberganov

O'zbek, turk va uyg'ur tillarida pos

teglash va stemming.....40

Dilrabo Elova, Sabohat Allanazarova

O'zbek tili matnlarida sentiment tahlil usullari.....65

Oqila Abdullayeva, Sabura Xudayarova

O'zbek tilida so'z birikmalarining lisoniy sintaktik qoliplari va

ularni modellashtirish masalasi77

Xolisa Axmedova

Statistik usullar yordamida turli so'z turkumlari orasidagi

omonimiyani aniqlash.....91

O'ZBEK TILI MATNLARIDA SENTIMENT TAHLIL USULLARI

Dilrabo Elova¹,
Sabohat Allanazarova²

Annotatsiya

Ijtimoiy tarmoqlardagi postlar, so'rovlar yoki sharhlarni tasniflash orqali kompaniyalar o'zlarining strategiyalari qanday ishlayotganligi va qay darajada samarador ekanligini aniqlashlari mumkin. Sentiment tahlil – bu NLPning dolzarb vazifalaridan biri bo'lib, berilgan matnlarda ifodalangan his-tuyg'ular subyektiv bahosini aniqlash imkonini beradi. Bugungi kunda tuyg'ularni tahlil qilishning bir necha usullari mavjud bo'lib, ularning har biri o'zining kuchli va zaif tomonlariga ega. Ushbu maqolada hissiyotlarni tasniflash usullarining zamonaviy yondashuvlari haqida umumiy ma'lumot beriladi. Leksemaning lug'aviy ma'nosiga asoslangan yondashuvni amalga oshirish usullari yoritib beriladi.

Kalit so'zlar: *NLP, hissiy tahlil, sharhlar, mashinali o'qitish, ML, sun'iy intellekt AI.*

Kirish

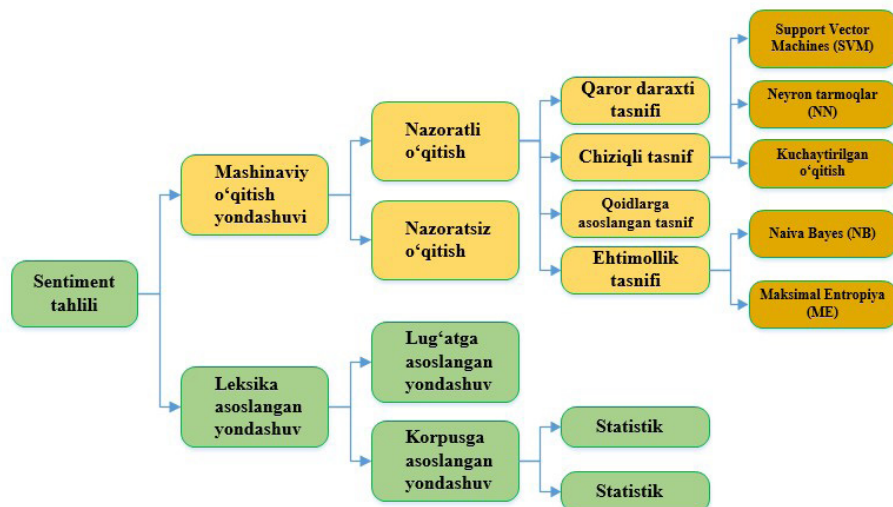
Sentiment (hissiy) tahlil – bu matndagi his-tuyg'ularni aniqlashga yordam beradigan tabiiy tilni qayta ishlash (NLP) usuli [Yilmaz, 2023]. Matnni izohlash NLP modellarini o'rgatish jarayoni orqali amalga oshiriladi. Matn ko'rinishidagi ma'lumotlarni NLP modellari asosida qayta ishlash kontentda ifodalangan mazmunni aniqlash va ulardan muhim xulosalar chiqarishga yordam beradi [Bolshakova, Boronsov, Efremova, 2017].

Matnni sentiment tahlil qilish natijasida aniqlangan hissiyot **ijobiy, salbiy** yoki **betaraf** bo'lishi mumkin [Shivanandhan]. Bugungi kunda ingliz, nemis, xitoy, rus va shu kabi xalqaro tillardagi matnlar uchun sentiment tahlil qilish algortimlari va NLP ilovalari

¹ *Elova Dilrabo Qudratillayevna* – filologiya fanlari falsafa doktori (PhD). Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti.
E-pochta: elova@navoiy-uni.uz
ORCID: 0000-0002-2329-1811

² *Allanazarova Sabohat Yusupboyevna* – Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti tayanch doktoranti.
E-pochta: allanazarova.sabosh@gmail.com
ORCID: 0009-0003-5459-6049

ishlab chiqilgan (1-rasm) [Taboado, Brooke, Tofiloski, Voll, 2011; Daud, Zhou, Muhammad, 2010; Yanina, Boronsov, 2016, Olivas, Romero, 2015]. Tasniflashda uch xil bahoni darajalarga (**quyi, o'rta va yuqori**), har bir bahoni hissiyot toifalariga (**salbiy: g'amgin, xafalik, afsuslanish, norozilik; ijobiy: xursandlik, mamnuniyat, ishonch**) ajratish mumkin. Yuqori aniqlikka ega modellarni ishlab chiqish uchun NLPda an'anaviy va mashinali o'qitish algoritmlaridan birgalikda foydalanish maqsadga muvofiq [Matlatipov, Kuryozov, Miguel, 2019].



1-rasm. Eng ko'p qo'llaniladigan hissiyotlarni tasniflash usullarining umumiy ko'rinishi

1. Leksemaning lug'aviy ma'nosiga asoslangan yondashuv

So'zlar nafaqat tushuncha ifodalash vazifasini, balki, shu bilan birga, axborot uzatuvchining axborotga nisbatan hissiy munosabatini ham ifodalaydi. Til birliklari (fonema, morfema, leksema) tarkibida bu hissiy munosabatni ifodalovchi vositalardan biri leksemalardir. Ana shu jihatdan o'zbek tili leksemalari neytral, bo'yoqsiz va bo'yoqdor qatlamga bo'linadi. Leksikaga asoslangan yondashuv asosida sentiment tahlilni amalga oshirishda leksemaning semantik tarkibiga asoslaniladi. Ya'ni, leksik birliklarining lug'aviy ma'nosidagi ijobiy yoki salbiy bo'yoqdorlik ifoda semalariga, neytral qatlamga tegishli leksemalarning salbiy yoki ijobiy ma'nolariga e'tibor qaratiladi.

Ijobiy bo'yoqdor	Neytral	Salbiy bo'yoqdor
chehra	yuz	bashara, turq
sohibjamol	chiroyli	badbashara
	ozg'in	qiltiriq

Leksemaning lug'aviy ma'nosiga asoslangan yondashuv so'zlarni teglash orqali lug'atni shakllantirish va sintaktik belgilarni ko'rib chiqish kabi bosqichlarni o'z ichiga oladi.

Lug'atga asoslangan yondashuvda leksik birliklarni teglash orqali ma'lumotlar bazasi ishlab chiqiladi. Lug'atga bog'liq yo'nalishning o'ziga xos qiyinchiliklari mavjud bo'lib, ba'zi bir so'zlar bir holatda *betaraf*, boshqa birida esa *ijobiy* yoki *salbiy* ma'no ifodalashi mumkin. Masalan,

Bu rusumli raqamli kamerada fokusni sozlash **uzoq vaqt** talab qiladi.

Bu rusumdagi televizorlar **uzoq vaqt** xizmat qiladi, shuning uchun kafolat muddati ham 10 yilga belgilangan.

Ijobiy bo'yoqdor leksemalar nutq jarayonida salbiy bo'yoqdor so'z o'rnida qo'llanilishi mumkin. Masalan,

Jamolingizni (turqingizni ma'nosida) boshqa ko'rmay!

Rosaaa *xursand* qildingiz? (xafa qildingiz ma'nosida.)

Leksemalarning ijobiy va salbiy bo'yog'ini ijobiy va salbiy ma'nodan farqlash lozim. Masalan, *[yaxshi]* leksemasi ijobiy ma'noga ega, ammo uning ifoda semasi neytral, ya'ni u bo'yoqdorlikka ega emas. *[xunuk]* salbiy ma'noga ega bo'yoqsiz leksemdir.

2-jadval. Leksemalarning ijobiy va salbiy bo'yog'i

Ijobiy ma'noli neytral leksema	Salbiy ma'noli neytral leksema
chiroyli	xunuk
yaxshi	yomon
vafo	xiyonat

Ijobiy yoki salbiy ma'noga ega leksemalar ham soha yo'nalishiga qarab hissiy bahosini o'zgartirishi mumkin. Misol uchun, *ma'no* va *xiyonat* so'zlari kino sharhlari sohasida baholovchi atamalar emas, chunki ular bu jarayonda film syujetida ifodalangan voqeaga nisbatan ham qo'llanilishi mumkin. Masalan, "*Kesakpolvon Asadbekka bekor xiyonat qildi*" – deb Shaytanat serialining voqealarida keltirilgan holatni tasvirlab sharh qoldirilsa, *[xiyonat]* leksemasi salbiy ma'noga

ega bo'lsa ham, sentiment tahlilda baholovchi xususiyatga ega emas.

Siyosat sohasida *kulgili* so'zi salbiy baho va aksincha, komediyalarga kelganda ijobiy bahoni ifodalashi mumkin:

3-jadval. Ijobiy yoki salbiy ma'noga ega leksemalar

Kino sohasi	Siyosat sohasi
Javlon Shodmonov (komediya yo'nalishida ishlaydigan ijodkor) filmlarini ko'rishni tavsiya qilaman. Hammasi kulgili . Bo'sh vaqtingda, zerikmaysan.	Ba'zi deputatlarimizni fikrlari juda kulgili . Biri – "Ikki yarim million so'm besh nafar a'zoga ega bo'lgan oilaga bemalol yetadi" - desa, boshqasi "Kuniga 5600 so'm pul ishlagan odam kambag'al emas" deydi.

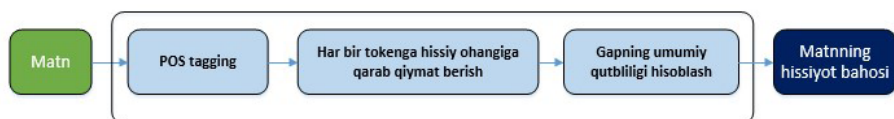
Mavzu doirasida baholovchi so'zlar vositasida ta'riflanayotgan obyektning turli xususiyatiga ko'ra leksema ijobiy yoki salbiy ohangga ega bo'lishi mumkin. Masalan,

– Bu rusumdagi raqamli kameraning batareyasi **uzoq** vaqt xizmat qiladi.

– Ammo fokusni sozlash uchun ham **uzoq** vaqt sarflash kerak.

Demak, lug'atga asoslangan yondashuv orqali sentiment tahlilni amalga oshirish uchun leksemaning lug'aviy tarkibida mavjud bo'lgan semaga asoslaniladi. Leksemalarni emotsional-ekspressivligiga ko'ra, lug'aviy semadagi ijobiy-salbiyligiga ko'ra lug'atlar shakllantiriladi. Kontent tarkibidagi so'zlar ma'lumotlar bazasi bilan qiyoslanadi, olingan natija asosida hissiy baho aniqlanadi.

Matnning hissiyot bahosi quyidagi bosqichlar orqali aniqlandi:

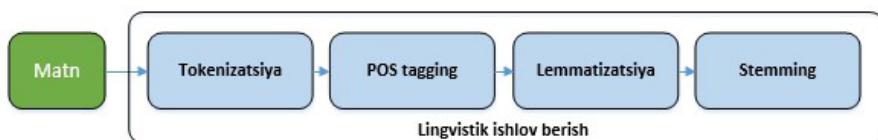


1-rasm. Matnning hissiyot bahosini aniqlash bosqichlari

POS tegging

Lug'atga asoslangan usul yordamida kontent sentiment tahlilini amalga oshirishning dastlabki bosqichida **POS tegging (teglash)** usuli orqali har bir so'z ajratib olinadi.

POS teglash – bu berilgan gapdagi har bir so'zshaklga uning turkum (*ot, fe'l, sifat, son, ravish yoki olmosh*)ga mansubligini belgilash (teglash) vazifasidir. POS teglash tabiiy tilni qayta ishlash (Natural Language Processing, NLP)ning asosiy vazifalaridan biri bo'lib, pipeline konveyerining muhim bosqichi hisoblanadi (1-rasm) [Elov, Hamroyeva, Abdullayeva, Uzoqova, 2022].



2-rasm. Matnga boshlang'ich ishlov berish bosqichlari

POS teglash lug'at asosida yoki lug'atsiz amalga oshirilishi mumkin. POS teglash bo'yicha amalga oshirilgan ilmiy tadqiqotlarning aksariyati [Kumawat, Jain, 2015] so'zlarga asoslangan (word-based) bo'lib, so'zlarning morfologik segmentatsiyasini (morphological segmentation) amalga oshirmaydi.

4-jadval. O'zbek tillaridagi so'zshaklning lemmasi, stemi va POS tegi

No	So'z	Lemma	POS	Stem	POS	Root	POS
1	muzladi	muzlamoq	VB	muz	N	muz	N
2	issiqroq	issiq	JJ	isi	VB	isi	VB
3	sodda-lashtiriladi	sodda-lashtirmoq	VB	sodda	JJ	sodda	JJ
4	ixtiyoriy	ixtiyoriy	JJ	ixtiyor	N	ixtiyor	N
5	qo'llaniladigan	qo'llamoq	VB	qo'l	N	qo'l	N
6	yo'lakda	yo'lak	N	yo'l	N	yo'l	N
7	qishlog'im	qishloq	N	qishlog'	?	qishloq	N

Har bir tokenga hissiy ohangiga qarab qiymat berish

Ikkinchi bosqichda ajratib olingan har bir token ma'lumotlar bazasi asosida emotsional-ekspressivligiga ko'ra, lug'aviy ma'noning salbiy va ijobiyligiga ko'ra tekshiriladi.

Gapning umumiy qutbliligi hisoblash

Uchinchi bosqichda olingan natijalar asosida kontentda ifodalangan fikrning hissiy bahosi hisoblanadi.

Lug'atga asoslangan yondashuvning afzalligi shundaki, o'rgatilgan resurslarga boy tillar uchun ishlab chiqilgan ko'plab manbalarning mavjudligi va bunday manbalardan erkin foydalanish imkoniyati kengligidir. Kamchilik tomoni, lug'atlar odatda rasmiy so'zlarni o'z ichiga oladi, aksincha sharhlar (tvitlar) jonli so'zlashuv uslubida ifodalanadi.

Korpusga asoslangan usul

Korpusga asoslangan usul *sintaktik belgilarni ko'rib chiqishni* o'z ichiga olib, ushbu usulning samaradorligi korpusning hajmiga

bog'liq. Ushbu yondashuvdan foydalanish tezaurus lug'atlarni, turli sohalar va kontekst mavzusiga xos lug'atlarni shakllantirishga yordam beradi [Park, Kim, 2016; Liu, 2012]. Belgilangan matndagi so'zlar yig'indisini korpusdan foydalanib hissiy bahosi aniqlanadi. Olingan statistik natijalar asosida matnga hissiy baho beriladi. Jahon tajribasida olib borilgan tadqiqotlar shuni ko'rsatadiki, aniqroq natijalarga erishish uchun korpusda maxsus lingvistik qoidalar qo'llaniladi [Hatzivassiloglou, Keown, 1997].

1. Hissiy bahosi kontentga bog'liq bo'lgan so'zlarning semantik yo'nalishini aniqlashda, obyektning turli xususiyatlarini ifodalovchi otli birikmalar, harakatning belgisini ifodalovchi fe'lli birikmalar korpus yordamida aniqlab olinadi va ifodalayotgan hissiy bahosiga ko'ra teglanadi. Soha yo'nalishidan kelib chiqib, ma'lumotlar bazasi shakllantiriladi. Masalan,

Bu rusumli raqamli kamerada **fokusni sozlash uzoq vaqt talab qiladi.**

Bu rusumdagi **televizorlar uzoq vaqt xizmat qiladi**, shuning uchun kafolat muddati 10 yilga belgilangan.

2. Korpus tarkibidan hissiy bahoga ega sifatlar va ravishlar ajratib olinadi. Jumlada bunday so'zlar [va], [hamda], [ham] bog'lovchilar bilan biriksa, ular bir xil hissiy baholi deb qabul qilinadi. Masalan,

Toshkent restoranida tayyorlanadigan taomlar *mazali* va sog'liq uchun **foydali.**

Jumladagi [mazali] va [foydali] leksemalarining lug'aviy ma'nosida ijobiylik xususiyati bor, ular [va] bog'lovchisi bilan birikkan. Bu jumla ijobiy baholanadi.

3. [na ... na....] inkorni ifodalovchi vosita yordamida birikkan belgi bildiruvchi so'zlarning lug'aviy ma'nosida ijobiy sema bo'lsa, hissiy baho natijasi salbiy, aksincha leksemaning lug'aviy ma'nosida salbiy ma'no bo'lsa hissiy baho natijasi ijobiy xarakterga ega bo'ladi. Masalan,

Katta yo'llar bo'yidagi oshxonalarining taomlari **na ishonchli, na mazali.**

Jumladagi [ishonchli] va [mazali] leksemalarining lug'aviy ma'nosida ijobiylik xususiyati bor [na] bog'lovchisi bilan birikkan.

Jumlaning tarkibiy qismlari [ammo], [lekin], [biroq], [-u], [-yu] bog'lovchi vositalari yordamida biriksa, hissiy baho ifodalovchi so'zlar qarama-qarshi qutbli bo'ladi. Masalan,

Bu yerda tayyorlangan baliq menga juda **yoqadi, lekin** baliqqa **allergiyam bor.**

Bundan tashqari, [yoki ..., yoki], [yoki] kabi ayiruv bog'lovchilari ham qo'llanishi mumkin. Bunday hollarda lingvistik qoidalar asosida olingan natija noaniq bo'ladi. Jummalarda hissiy baho ifodalovchi birliklarning ijobiy yoki salbiy qutbga egaligini aniqlash murakkablashadi. Bunday hollarda matematik modellardan foydalaniladi. Misol sifatida so'zlarning [-10, 10] oraliqdagi qiymatlar bilan baholangan ma'lumotlar bazasi ishlab chiqiladi. Yakuniy hissiy bahoni olish uchun matnni alohida so'zlarga ajratish va ularni so'zlar ro'yxati bilan solishtirish talab qilinadi. Mavjud ma'lumotlar bazasidan [yoqadi] va [allergiya] so'zlarini topib, ularning tegishli ballari +3 va -2 aniqlanadi. Qolgan so'zlar e'tiborsiz qoldiriladi. Ushbu ikkita baho umumlashtirilib, +1 umumiy ball chiqadi. Ushbu jumlaning **biroz/qisman** ijobiy deb baholash mumkin.

Yuqorida bildirilgan fikr-mulohazalar asosida leksemaning lug'aviy ma'nosiga asoslangan yondashuvning bir qator afzallik va kamchiliklarini keltiramiz:

5-jadval. Lekseмага asoslangan yondashuvning yutuq va kamchiliklari

Afzalliklari	Kamchiliklari
– Ko'plab ochiq manbalar (masalan, SentiWordNet , Senticnet , Word-Net) mavjud	– Ijtimoiy tarmoqdagi ma'lumotlarni tasniflashda murakkablik
– Tejamkor, chunki ular hissiyotlarni tahlil qilish algoritmilarini talab qilmaydi	– Kinoya va istehzoni farqlamaydi
– O'quv ma'lumotlariga ehtiyoj qolmaydi, chunki so'zlarning ma'nosiga tezkor kirish imkoniyati mavjud	– Grammatik xatolar, noto'g'ri imloni aniqlamaydi
	– Juda qat'iy va domenga bog'liq (so'z kontekstdan ke- lib chiqib baholanmaydi)

2. Avtomatik/Mashinali o'qitishga asoslangan yondashuv

Bugungi kunda sun'iy intellekt (AI) va mashinali o'qitish (ML) jahon tajribasida ommalashdi. Aksariyat korxonalar o'z faoliyatida Aidan foydalanmoqda. Sun'iy intellektga asoslangan yechimlarni ishlab chiqish va amalga oshirish uchun grammatik qoidalar va katta hajmdagi ma'lumotlar bazasi/ombori shakllantirilishi lozim. Algoritmalar ishlab chiqilgandan so'ng, ma'lumotlar izohining avtomatlashtirish usulidan foydalanish mumkin.

Avtomatlashtirilgan his-tuyg'ularni tahlil qilish usullari statistik modellar asosida hissiyotlarni tasniflaydigan ML algoritmlarini o'z ichiga oladi. Mashinaviy o'qitishga asoslangan

algoritmlar katta hajmdagi matnlar uchun yuqari aniqlikda ishlaydi. Mashinali o'qitish algoritmlarini amalga oshirish uchun jumalarni vektor ko'rinishida tasvirlash kerak. Keyin modellar jumlaning hissiyotini taxmin qilishga o'rgatilishi mumkin. Sentiment tahlilida keng qo'llaniladigan Naive Bayes algoritmi kichik hajmli, Logistik regressiya katta hajmli ma'lumotlar to'plami uchun yaxshi ishlaydi.

6-jadval. **Mashinali o'qitishga** asoslangan yondashuvning yutuq va kamchiliklari

Kamchiliklari	Afzalliklari
– Xarajatlar an'anaviy qoidalarga asoslangan usullarga nisbatan qimmatroq	– Tuyg'ularni tahlil qilishda kinarya, istehzo yoki inkorni aniqlashga o'rgatish mumkin
– Yuqori aniqlik ballariga ega bo'lish uchun katta miqdordagi ma'lumotlar talab qilinadi	– Ijtimoiy tarmoq ma'lumotlarini tahlil qilishni osonlashtiradi
– Eng ilg'or modellar ham inson aralashuvini talab qiladi	– So'zlarning emotsional valentligini biladi, shuning uchun ular oldindan belgilangan ma'lumotlar to'plamini talab qilmaydi

3. Gibrid/Aralash yondashuv

Leksikaga asoslangan va avtomatlashtirilgan usullarning o'ziga xos afzallik va kamchiliklari mavjud. Kamchiliklarni bartaraf qilish maqsadida avtomatlashtirilgan va leksikaga asoslangan usullarni o'z ichiga olgan gibrid usulni amalga oshirish mumkin. Kombinatsiya parallel yoki tahlilning turli bosqichlarida amalga oshiriladi. Afzalliklarga qaramay, avtomatlashtirishda qiyinchiliklarga duch kelish mumkin, shuning uchun mashinaviy o'qitish modellari insonning doimiy aralashuvini va inson tomonidan izohlangan ma'lumotlarni talab qiladi.

Annotatsiyalangan ma'lumotlar

Annotatsiyalangan ma'lumotlar – bu kompyuter tushunishi uchun til birliklarini tegishli ma'lumotlar bilan razmetkalash jarayoni. Ushbu ma'lumotlar lingvistik va ekstralingvistik (rasm, matn, audio yoki video) ko'rinishda bo'lishi mumkin. Ma'lumotlar izohi dastlabki bosqichda qo'lda, keyin ilg'or mashinali o'qitish algoritmlari va vositalaridan foydalangan holda avtomatik ravishda amalga oshirilishi mumkin.

Yuqorida ko'rib o'tilgan yondashuvlarning yutuqlariga qaramay, kamchiliklari mavjudligi sababli, ko'plab ishlab chiqaruvchilar mukammal yechim sifatida kraudsorsing xizmatini

Kraudsorsing hissiyot tahlili nima va uning afzalliklari

Kraudsorsing, ayniqsa, so'nggi yillarda mashhur bo'lgan usul bo'lib, u muayyan muammolarni hal qilish uchun bir qator vazifalarni bajaradigan pullik xizmat turi. Unga ko'ngillilarni o'z ichiga oladigan sentiment tahlili bo'yicha mutaxassislar jamlanadi. Inson ishtiroki mavjudligi sababli, kraudsorsing iste'molchilarning sharhlari tahlilida yuqori aniqlikdagi natijalarni beradi.

Afzalliklari:

- Yuqori aniqlik;
- Hissiyotlarni tasniflash kabi subyektiv qarorlarni talab qiladigan sohalarida avtomatlashtirilgan usullardan ko'ra aniqroq natijalar beradi;
- Bu noaniqlikni kamaytiradi, sababi bir xil matnni turli kishilar baholashadi.

Hozirgi kunda kraudsorsing xizmatlarini taklif qilayotgan bir necha guruhlar mavjud. Masalan: **Clickworker, COGITO, iMerit, BUNCH, Shaip, SunTec.AI**. Bunday xizmat turida ma'lumotlar sir tutilishi kafolatlanadi.

Xulosa

Iste'molchi ma'lumotlari ulkan hajmda o'sib borishi bilan ularning barchasini tahlil qilish tobora qiyinlashib bormoqda. Bugungi globallashuv jarayonida axborotlarning bahosi o'z vaqtidagina qadrlidir. Ma'lumotlarni inson tomonidan tahlil qilinishi ko'p vaqt talab qiladi. Kompaniyalar mijozlarining munosabatlarini aniqlash va mavjud vaziyatni tahlil qilish uchun mashinali o'qitishga asoslangan algoritmlardan foydalanadilar. Yuqorida leksikonga va mashinali o'qitishga asoslangan yondashuvlarni, ularning o'ziga xos yutuq va kamchilik tomonlari ko'rib chiqildi.

Foydalanilgan adabiyotlar

Begum Yilmaz. Sentiment analysis methods in 2023: Overview, pros & cons <https://research.aimultiple.com/sentiment-analysis-methods/>

Большакова Е.И., Воронцов К.В., Ефремова Н.Э., Клышинский Э.С., Лукашевич Н.В., Сапин А.С. /Автоматическая обработка текстов на естественном языке и анализ данных : учеб. пособие / — М.: Изд-во НИУ ВШЭ, 2017. — 127-177с. <https://www.hse.ru/>

- Manish Shivanandhan. What is Sentiment Analysis? A Complete Guide for Beginners. <https://www.freecodecamp.org/news/what-is-sentiment-analysis-a-complete-guide-to-for-beginners/>
- Maite Taboada, Julian Brooke, Milan Tofiloski, Kimberly Voll, Manfred Stede. Lexicon-Based Methods for Sentiment Analysis. // Computational Linguistics (2011) 37 (2): 267–307.
- Daud A., Li J., Zhou L., Muhammad F. Knowledge discovery through directed probabilistic topic models: a survey // Frontiers of Computer Science in China. 2010. Vol. 4 (2) - P. 280–301.
- Янина А. О., Воронцов К. В. Мультимодальные тематические модели для разведочного поиска в коллективном блоге // Машинное обучение и анализ данных. — 2016. — Т. 2, № 2. — С. 173–186.
- Jesus Serrana-Guerrero, Jose A. Olivas, Francisco P. Romero, and Enrique Herrera-Viedma (2015), "Sentiment analysis: A review and comparative analysis of web services", Information Sciences, 311, pp.18-38
- 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
- B.Elov, Sh.Hamroyeva, O.Abdullayeva, M.Uzoqova. O'zbek tilida pos tagging masalasi: muammo va takliflar. O'zbekiston: til va madaniyat (Amaliy filologiya), 2022, 5(4). 45-63 b.
- Kumawat, D., & Jain, V. (2015). POSTagging Approaches: A Comparison. International Journal of Computer Applications, 118(6). <https://doi.org/10.5120/20752-3148>
- Park S., & Kim Y. (2016, June). Building thesaurus lexicon using dictionary-based approach for sentiment classification. In Software Engineering Research, Management and Applications (SERA), 2016 IEEE 14th International Conference on (pp. 39-44). IEEE.
- Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 1-167.
- Hatzivassiloglou V., Mc Keown K. R. Predicting the semantic orientation of adjectives. / In Proceedings of the eighth

METHODS OF SENTIMENT ANALYSIS IN UZBEK LANGUAGE TEXTS

Dilrabo Elova¹,
Sabohat Allanazarova²

Abstract

Companies use sentiment analysis in social media posts, surveys or reviews can measure how their strategies work and efficiency for development. Sentiment analysis is a natural language processing (NLP) method that helps identify the emotions in given texts. There are several methods to conduct sentiment analysis, each with strengths and weaknesses. In this article, we will give on overview of sentiment analysis methods of advantages and disadvantages.

Key words: *NLP, sentiment analysis, tweets, machine learning, artificial intelligence.*

References

- Begum Yilmaz. Sentiment analysis methods in 2023: Overview, pros & cons
<https://research.aimultiple.com/sentiment-analysis-methods/>
- Bolshakova Ye.I., Voronsov K.V., Yefryemova N.E., Klishinskiy E.S.,
Lukashyevich N.V., Sapin A.S. /Avtomaticheskaya obrabotka
tyekstov na yestiestvyennom yazykye i analiz dannykh : uchyeb.
posobiye / — M.: Izd-vo NIU VShE, 2017. — 127-177s. [https://
www.hse.ru/data/2017/07/22/1173852769/NLP_and_DA.pdf](https://www.hse.ru/data/2017/07/22/1173852769/NLP_and_DA.pdf)
- Manish Shivanandhan. What is Sentiment Analysis? A Complete Guide
for Beginners. [https://www.freecodecamp.org/news/what-is-
sentiment-analysis-a-complete-guide-to-for-beginners/](https://www.freecodecamp.org/news/what-is-sentiment-analysis-a-complete-guide-to-for-beginners/)

¹ *Elova Dilrabo Qudratillayevna* – doctor of philosophy in philology. Alisher Navoi
Tashkent State University of Uzbek Language and Literature.

E-mail: elova@navoiy-uni.uz

ORCID: 0000-0002-2329-1811

² *Allanazarova Sabohat Yusupboyevna* – PhD student of Tashkent State University
of Uzbek Language and Literature named after Alisher Navoi.

E-mail: allanazarova.sabosh@gmail.com

ORCID: 0009-0003-5459-6049

- Maite Taboada, Julian Brooke, Milan Tofiloski, Kimberly Voll, Manfred Stede. Lexicon-Based Methods for Sentiment Analysis. // Computational Linguistics (2011) 37 (2): 267–307.
- Daud A., Li J., Zhou L., Muhammad F. Knowledge discovery through directed probabilistic topic models: a survey // Frontiers of Computer Science in China. 2010. Vol. 4 (2) - P. 280–301.
- Yanina A. O., Voronsov K. V. Multimodalnyye tyematchyeskiye modyeli dlya razvyedoch'nogo poiska v kollyektivnom blogye // Mashinnoye obucheniye i analiz dannykh. — 2016. — T. 2, № 2. — S. 173–186.
- Jesus Serrana-Guerrero, Jose A. Olivas, Francisco P. Romero, and Enrique Herrera-Viedma (2015), "Sentiment analysis: A review and comparative analysis of web services", Information Sciences, 311, pp.18-38
- 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
- B.Elov, Sh.Hamroyeva, O.Abdullayeva, M.Uzoqova. O'zbek tilida pos tegging masalasi: muammo va takliflar. O'zbekiston: til va madaniyat (Amaliy filologiya), 2022, 5(4). 45-63 b.
- Kumawat, D., & Jain, V. (2015). POS Tagging Approaches: A Comparison. International Journal of Computer Applications, 118(6). <https://doi.org/10.5120/20752-3148>
- Park S., & Kim Y. (2016, June). Building thesaurus lexicon using dictionary-based approach for sentiment classification. In Software Engineering Research, Management and Applications (SERA), 2016 IEEE 14th International Conference on (pp. 39-44). IEEE.
- Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 1-167.
- Hatzivassiloglou V., Mc Keown K. R. Predicting the semantic orientation of adjectives. / In Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics (1997 July. pp. 174-181).