

DEVELOPMENT OF A LINGUISTIC DATABASE FOR SENTIMENT ANALYSIS

Sabokhat Yusupboyevna ALLANAZAROVA

PhD student

Alisher Navo'i Tashkent State University of Uzbek Language and Literature
Tashkent, Uzbekistan

SENTIMENT TAHLILI UCHUN LINGVISTIK TA'MINOTNI ISHLAB CHIQISH

Sabohat Yusupboyevna ALLANAZAROVA

Tayanch doktorant

Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti
Toshkent, O'zbekiston

РАЗРАБОТКА ЛИНГВИСТИЧЕСКОЙ БАЗЫ ДАННЫХ ДЛЯ СЕНТИМЕНТ-АНАЛИЗА

Сабохат Юсупбоевна АЛЛАНАЗАРОВА

Базовый докторант

Ташкентский государственный университет узбекского языка и литературы им. А. Навои
Ташкент, Узбекистан allanazarova.sabosh@gmail.com

For citation (iqtibos keltirish uchun, для цитирования):

Allanazarova S.Y. Sentiment tahlili uchun lingvistik ta'minotni ishlab chiqish // O'zbekistonda xorijiy tillar. — 2024. — 10-jild, № 4. — B. 12-23.

<https://doi.org/10.36078/1729151311>

Received: May 25, 2024

Accepted: August 17, 2024

Published: August 20, 2024

Copyright © 2024 by author(s).

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Abstract. Sentiment analysis is a field of research that uses NLP/Artificial Intelligence techniques to determine thoughts, emotions, and attitudes expressed about an object or event. This means analyzing the tone or mood of the text, feelings and thoughts. Many sentiment analysis studies and linguistic resources have been developed for the English language. However, there are a limited number of studies on the Uzbek language. The reasons for this are insufficient data collection, direct translation of existing English language resources not being compatible with the Uzbek language, differences in the grammatical structure of the two languages, and regional, religious, and mental differences. The study aims to identify emotional-expressiveness units that express emotional evaluation and to develop a linguistic database for emotional analysis. It deals with the issues of developing the linguistic support for sentiment analysis of Uzbek language texts; the study of context-dependent factors in emotional analysis; and determining the best style for the Uzbek language. Attention is paid to the analysis of expressing emotional evaluation means specific to the Uzbek language levels.

Keywords: Sentiment analysis; lexicon-based; Uzbek language; opinion mining; attitude; emotion; linguistic database.

Annotatsiya Sentiment tahlili biror obyekt yoki voqea-hodisa haqida bildirilgan fikr, hissiyot va munosabatni aniqlash uchun NLP/sun'iy intellekt usullaridan foydalanadigan tadqiqot sohasi hisoblanadi. Bu — matn ohangini yoki kayfiyati, tuyg'ularni tahlil qilish, inglizcha fikrlarni tahlil qilish demakdir. Ingliz tili uchun ko'plab hissiyotlarni tahlil qilish borasida tadqiqotlar o'tkazilgan va lingvistik resurslar ishlab chiqilgan. Biroq o'zbek tili bo'yicha cheklangan miqdordagi tadqiqotlar mavjud va bu tadqiqotlar ingliz tilidagi tadqiqotlarga qaraganda kamroq samaradorlikka ega. Buning sabablari ma'lumotlar

to'plamining yetarli emasligi, mavjud ingliz tilidagi resurslarni to'g'ridan-to'g'ri tarjimasi o'zbek tiliga mos kelmasligi, ikki til grammatik qurilishining farqliligi, mintaqaviy, diniy va mentalitetlar turfa xilligi bo'lishi mumkin. Ko'zlangan asosiy maqsad emotsional-ekspressivlik jihatdan hissiy baho ifodalovchi birliklarni aniqlash va hissiyot tahlili uchun lingvistik ta'minot ishlab chiqish. Mazkur maqolada o'zbek tilidagi matnlarning sentiment tahlili uchun lingvistik ta'minot ishlab chiqish; hissiy tahlil qilishda kontekstga bog'liq bo'lgan omillarni o'rganish; o'zbek tili uchun eng yaxshi uslubni aniqlash; o'zbek tili sathlariga xos hissiy baho ifodalovchi vositalarni tahlil qilishga e'tibor qaratiladi.

Kalit so'zlar: tuyg'u tahlili; leksika asosida; o'zbek tili; fikrlar intelluktual tahlili; munosabat; hissiyot; lingvistik ta'minot.

Аннотация. Сентимент-анализ (или анализ тональности текста) — это область исследований, в которой используются методы НЛП/искусственного интеллекта для определения мыслей, чувств и отношений, выраженных по поводу объекта или события. Для английского языка было разработано множество исследований по анализу настроений и лингвистических ресурсов. Однако количество исследований на узбекском языке ограничено, и эти исследования менее эффективны, чем исследования на английском языке. Причинами этого являются недостаточный набор данных, прямой перевод существующих англоязычных ресурсов, несовместимый с узбекским языком, различия в грамматической структуре двух языков, а также региональные, религиозные и ментальные различия. Основная цель статьи — выявление эмоционально-экспрессивных единиц, выражающих эмоциональную оценку, и развитие языкового обеспечения эмоционального анализа. В этой статье мы обсуждаем развитие лингвистической поддержки анализа тональности текстов на узбекском языке; изучение контекстно-зависимых факторов в эмоциональном анализе; определение лучшего стиля узбекского языка; внимание уделено также анализу средств выражения эмоциональной оценки, специфичных для разных уровней узбекского языка.

Ключевые слова: анализ тональности; лексика; узбекский язык; интеллектуальный анализ мнений; отношение; эмоции; лингвистический ресурс.

Kirish

Kompyuter lingvistikasi fanining maqsadi til va matnlar bilan ishlashni avtomatlashtirish, matnlarni tahlil qilish, tarjima qilish, til o'qitish tizimini avtomatlashtirish va shu kabi vazifalarni bajarishdir. Ushbu soha o'zbek tilshunosligida yangi soha bo'lganligi uchun hali ko'p tadqiqotlar bajarilishi kerak. Shu jumladan, morfologik analizator, sintaktik masalalarni hal etish, semantik muammolarni ochish va hal etish, sentiment tahlil analizatorini yaratish kabi masalalar oldimizda turgan ustuvor vazifalar sanaladi. Emotsional-ekspressiv so'zlarning bazasini shakllantirish yuqorida keltirilgan katta maqsadlar sari kichik amaliyot sanaladi. Ushbu lingvistik resursni shakllantirish bizga semantik tahlilni yaxshilashga, sintaktik analizatorni to'g'ri ishlashiga, mashina tarjimasi ravon bo'lishiga va ko'plab qimmatli maqsadlarga xizmat qiladi.

O'zbek tili lug'at boyligidagi so'zlar so'zlovchining turli emotsional-ekspressiv munosabatini ifodalash xususiyatiga ko'ra ikki turga:

emotsional-ekspressiv bo'yoqsiz (betaraf) so'zlar va emotsional-ekspressiv bo'yoqdor so'zlarga ajratiladi. Leksik ma'no ustiga so'zlovchining munosabati kabi turli ma'noni ifodalovchi so'zlar emotsional-ekspressiv bo'yoqdor so'zlar bo'lib, sentiment (hissiy) tahlilda tadqiqot predmeti sifatida qaraladi (9, 14). Shunga ko'ra ikkiga: *ijobiy emotsional bo'yoqdor so'zlar* va *salbiy emotsional bo'yoqdor so'zlarga* ajraladi (1, 3). Salbiylik o'z ichida mensimaslik, norozilik, noxush holat, nafratlanish kabi munosabatlarni, ijobiylik esa xursandlik, mamnunlik kabi holatlarni ifodalash uchun ishlatiladi. Ushbu tadqiqotda biz o'zbek tilining o'ziga xos xususiyatlarini hisobga olgan holda leksikaga asoslangan (LB) usulni taklif qilishni maqsad qildik.

Asosiy qism

1-jadval

Tadqiqot ishlarining umumiy natijalari

Tadqi-qot	Metodologiya	Cheklovi	Natija
(10, 4)	Lug'atga asoslangan yondashuv hissiyotlarni tahlil qilishni, tvitlardan ma'no chiqarishni, ballarni hisoblashni va R tilidan foydalangan holda sharhlarni ijobiy yoki salbiy deb tasniflashni o'z ichiga oladi.	– Qo'llanish doirasi tvitlarning hissiyot tahlili bilan cheklangan. – his-tuyg'ularni tasniflashni ortiqcha soddalashtirish. – So'zga asoslangan yondashuvning murakkabligi – ML bilan taqqoslanmagan. – Umumlashtirilmagan.	tabiiy tilni qayta ishlashda hissiyotlarni tahlil qilishga qaratilgan gazeta matnlarini turli darajadagi ijobiy, salbiy yoki betaraf deb tasniflangan.
(8, 1)	Metodologiya tabiiy tilni qayta ishlashda hissiyotlarni tahlil qilish, hissiyot so'zlarini aniqlash va gazeta matnlarini turli darajadagi his-tuyg'ularga tasniflashni o'z ichiga oladi.	– Qo'llanish doirasi tvitlarning hissiyot tahlili bilan cheklangan. – his-tuyg'ularni tasniflashni ortiqcha soddalashtirish. – So'zga asoslangan yondashuvning murakkabligi – ML bilan taqqoslanmagan. – Umumlashtirilmagan.	Sharhlardan ma'no chiqarish va hissiyotni aniqlash uchun ballni hisoblash kiradi.

		an.	
(3, 1)	SentiWordNet lugʻatidan foydalangan holda veb-muhokama postlarining hissiyotlarini tasniflash uchun leksik va leksik boʻlmagan kontekstual qiymat oʻzgarishlarini hisobga olishni oʻz ichiga oladi. Yondashuv lugʻatdan oldingi qutblilik koʻrsatkichlarini olish va ularni kontekstga koʻra oʻzgartirish, soʻngra ularni uchta maʼlumotlar toʻplamida baholashni oʻz ichiga oladi.	Baholash faqat uchta maʼlumotlar toʻplamida oʻtkazilgan.	Tadqiqot veb-muhokama muhitida hissiyotlarni tahlil qilish usullarini sinab koʻrish zarurligini taʼkidlaydi va veb-muhokama postlarining hissiyotlarini tasniflash uchun SentiWordNet-dan foydalanish samaradorligini koʻrsatadi.
(13, 87)	Metodologiya maʼlum bir kontseptsiya bilan bogʻliq boʻlgan matn qismlarini tushunish uchun hissiy lugʻatlar, hissiyotlarni hisoblash funktsiyalari va lingvistik qoidalarini birlashtirishni oʻz ichiga oladi.		matnni tushunishni yaxshilaydigan lingvistik qoidalarini aniqlash, his-tuygʻularni tahlil qilish natijalarida shaffoflik va sharhlashga eʼtibor berish va lugʻatga asoslangan hissiyotlarni tahlil qilishda ilgʻor yondashuvlarni osonlashtirgan,
(14,2)	Tuygʻularni aniqlash va ajratib olish uchun ot, sifat, feʼllar va ravishalar keyinchalik tasniflash uchun aniqlanadi. SVM klassifikatori maʼlumotlarni maksimal chegara bilan ijobiy yoki salbiy ikki toifaga ajratuvchi giperplanni topadi. Ushbu texnikaning nomi har bir tekislikning chegarasini topish uchun ishlatiladigan maʼlumotlar nuqtalarining ketma-ketligi boʻlgan qoʻllab-quvvatlash vektoridan foydalanish. Tasniflagich yangi kirishni chegaraning qaysi tomoniga tegishli ekanligini taxmin qilish	Ushbu yondashuvning asosiy muammosi shundaki, u oʻquv korpusidagi atamalarning qutbliligiga bogʻliq. Ushbu cheklovga qaramay, uning soddaligi hissiyotlarni tahlil qilish boʻyicha turli takliflarda foydalanishni oshirdi	

	orqali tasniflaydi.		
(11, 1337)	Tadqiqotda qo'llanilgan metodologiya hissiyotlarni tahlil qilish uchun Sentiment Convolutional Neural Network (SentiCNN) ni taklif qilishni o'z ichiga olgan bo'lib, his-tuyg'ularni bashorat qilishni yaxshilash uchun Highway Network va uchta lug'atga asoslangan mexanizmni (LBAM) birlashtirgan.	Tuyg'ularni tahlil qilishda mavjud yondashuvlar jumalarning kontekstual ma'lumotlarini va hissiyot so'zlariga kiritilgan hissiyot ma'lumotlarini e'tiborsiz qoldiradi.	Jumalarning his-tuyg'ularini tahlil qilish uchun kontekstual va hissiyot ma'lumotlarini o'z ichiga olgan Emotion Convolutional Neural Network (SentiCNN) modelini, shuningdek, Highway Network va LBAM mexanizmini taqdim etadi.
(4, 287)	Metodologiya lug'atga asoslangan hissiyotlarni tahlil qilish bo'yicha oldingi tadqiqotlarni ko'rib chiqishni, qutblanish tasnifi uchun zamonaviy lug'atga asoslangan yechim berish va hissiy tasnif uchun turli lug'atlarni tahlil qiladi.		Ushbu maqola asosan qutblanish tasnifi uchun ishlatiladigan turli lug'atlarni qiyoslashga qaratilgan.
(5, 200)	Metodologiya Amazon ma'lumotlar to'plamidan mashina o'rganishga asoslangan Senti-so'zli lug'atni (MLBSL) ishlab chiqishni, sharhlarni tasniflash uchun "Term Count" hisoblash usullaridan foydalanishni va "Term Score Aggregation" yordamida ijobiy va salbiy sharhlarni o'z ichiga olgan muqobil ma'lumotlar to'plamidan foydalanishni o'z ichiga oladi.	– Qo'lda ishlab chiqilgan tuyg'u lug'atlari ko'p vaqt talab etadi va tarafkashlikka moyil – Ko'rish lug'atlarini kengaytirish uchun leksik induktiv yondashuvlardan foydalaniladi – Tadqiqotda taklif qilingan yaxshilanishlar MLBSL ish faoliyatini yaxshilashga qaratilgan – 15 000 tokenda kuzatilgan eng yaxshi MLBSL ishlashi – Ta'lim sohasidagi	Amazon korpusi uchun 79,36% Filmlar korpusi uchun 77,8% aniqlik. MLBSL oldingi lug'atlardan ustun va yuqori aniqlikka ega.

		ma'lumotlar 0,33 % ga kichik yaxshilanishga olib keldi – Agregatlash usulini baholash tasnifi 5,87 % ga yaxshilandi – string atributini ildiz atributiga almashtirish aniqlikni 0,83 % ga oshirdi	
(6, 468)	Tadqiqotda qo'llaniladigan metodologiya hissiyotlarni tahlil qilish uchun lug'atga asoslangan yondashuvning keng qamrovli sharhini o'z ichiga oladi, uning cheklovlari va ma'lumotlarining sifati va miqdorining ahamiyatini ta'kidlaydi.	– leksik obyektlarning turli sohalarida o'zgaruvchanligi tufayli leksik asosli tahlil yuqori aniqlikka ega emas. – So'z darajasidagi tahlilni optimallashtirish.	Maqolada lug'atga asoslangan tahlilning cheklovlari haqida tushuncha berilgan va ikkilik va ko'p sinfli tasniflash uchun mualliflarning taqqoslashlari keltirilgan.

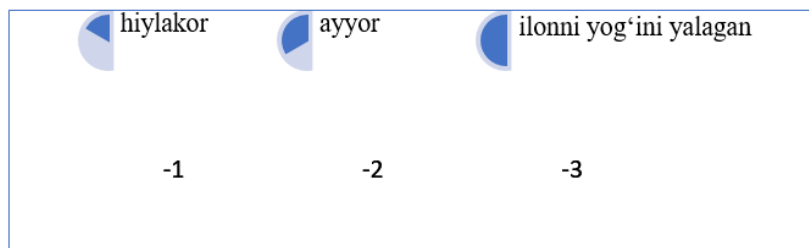
Nutq jarayonida so'zning faqat leksik ma'nosigina nazarda tutilmaydi: so'zlovchi o'zi aytayotgan narsa-hodisalarga o'z munosabatini ham bildiradi. Demak, so'zlar nutq jarayonida leksik ma'nodan tashqari, so'zlovchining atalayotgan narsa-hodisalarga ijobiy yoki salbiy munosabatini ifodalovchi qo'shimcha ma'noga — uslubiy bo'yoqqa ham ega bo'ladi. Uslubiy bo'yoq leksik ma'noga tayangan holda uni yanada boyitadi.

So'z ba'zan matnda bir ma'noni, alohida olinganda boshqa ma'noni anglatishi mumkin, so'zlarning bir necha ma'no anglatib kelishi ko'p ma'nolilik (polisemantizm) sanaladi. So'zning bunday ko'p ma'noli bo'lishi til taraqqiyotining natijasidir. Tilda biror narsa yoki hodisani atash sababli paydo bo'lgan so'z keyinchalik boshqa narsa, hodisalarni ham ifodalashda qo'llana boshlaydi. Shu sababli ko'p ma'noli so'zga aylanadi. Masalan:

1. Hayvon — obyektiv ma'no — betaraf baho.
2. Hayvon — haqorat so'zi — salbiy baho.

Ko'p ma'noli so'zlarda ikki ma'no turi: bosh (tub) ma'no va yasama ma'no mavjud. Yasama ma'no so'zning bosh ma'nosidan o'sib chiqqan ma'nosidir. Yasama ma'no jumla talabiga qarab ko'chma ma'no, majoziy ma'no deb ham yuritiladi.

O'zbek tili lug'at tarkibi jihatdan juda boy va rang-barangdir. Tilimizda mustaqil leksik ma'noga ega bo'lgan so'zlardan tashqari, ma'no jihatdan shunday so'zga teng, ammo ikki va undan ortiq so'zning turg'un birikuvidan tashkil topgan frazeologik birliklar ham mavjud. Frazeologik birlik ifodaydigan ma'no ko'p jihatdan so'z anglatadigan leksik ma'nodan farq qiladi. Frazeologik ma'no so'zga nisbatan, avvalo, ma'noni o'ta kuchli darajada va muhimi, o'tkir obrazlilik, ta'sirchanlik belgisi bilan ifodalaydi. Shunga ko'ra, ular so'zning boy obrazlilikka ega bo'lgan ma'nodoshi sanaladi. Masalan:



1-rasm. Leksema va frazemalarning darajalanish ko'rsatkichlari

Ijobiy istak birliklari ularning mazmuniy xususiyatlariga, nutqiy sharoitga hamda real voqelik bilan bog'lanishga ko'ra duo va olqish shaklida tasnif qilinadi. Ko'plab ijobiy istak birliklari real voqelik bilan bevosita bog'lanmay, metaforik mazmunni yuzaga keltiradi:

Tuproq olsang, oltin bo'lsin!

Hamyoningga Xizr nazar solsin.

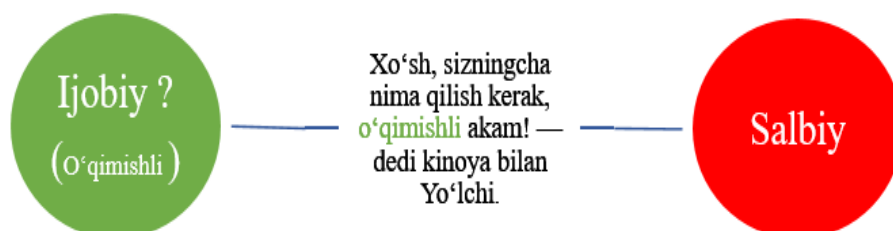
*Olgin-u, soldirma, ko'zlaganing kumush, tishlaganing oltin bo'lsin,
dushmanlaring lol bo'lsin, Allohu akbar!*

(Gazetadan)

O'zbek xalq olqishlari hayotning turli jarayonida turli sohalar bo'yicha xilma-xil shakl va ma'noda qo'llanadi. Duo va olqish shaklidagi ko'plab ijobiy istak birliklari real voqelik bilan bevosita bog'lanmay, metaforik mazmunni yuzaga keltiradi.

Nutqda so'z va iboralarning kesatiq, masxara, piching bilan o'z ma'nosiga zid holda qo'llanilishi kinoya (ironiya) sanaladi.

Sarkazm muallifning tasvirlanayotgan voqea-hodisaga salbiy munosabatini achchiq kulgi, ta'na asosida ifodalashga xizmat qiladi.



2-rasm. O'zbek tilidagi kinoyaviy gap namunasi

Sarkazm va kinoyani avtomatik aniqlashning yondashuvlari mavjud emasligi tasniflash aniqligini kamaytiradi. Lingvistik qoidalarga asoslangan ba'zi vaziyatni ifodalovchi (Amerikani ochgandek; uzoq kun/kuni bo'yi shifokor qabulini kutish; parvozning kechikishi) iboralardan foydalanish ham mumkin. Kinoyaviy birliklar ijobiy baho ifodalovchi so'zning salbiy holatga nisbatan ishlatilishi kuzatiladi (mazza qilib kutish). Matnni semantik qayta ishlashning bir turi bo'lgan kinoyaviy birliklarni aniqlashning avtomatik usullarini yechish zamonaviy kompyuter lingvistikasi uchun hali ham muammo hisoblanadi (7, 250). Buning uchun yashirin his-tuyg'ularni modellashdirish kerak. O'zbek tili uchun kichikroq hajmda bo'lsa-da yuqoridagi kabi bilimlar bazasini ishlab chiqishdek amaliy ishlar mutaxassislarimiz oldida turibdi.

Daraja ravishi: daraja tushunchasini anglatib, harakat-holatning bajarilish darajasini ko'rsatadi. Ba'zan daraja ravishlari sifat va ravishning o'ziga bog'lanib kelib, belgi ma'nosini kuchaytirib keladi. Bu ravishlar qay darajada? so'rog'iga javob berib, ularga tez, g'oyat, nihoyatda, obdon, yana, o'ta, tag'in, aslo, sira kabilar kiradi. Daraja ravishlari belgi ma'nosini kuchaytirishiga ko'ra (juda, g'oyat) sifat va ravishlarning ortirma darajasini hosil qiladigan element sifatida ham qatnashadi (1; 60). Ko'plik va kamlikni ifodalovchi ravishlar qutblilik qiymatlarini shakllantirishi mumkin. Misol uchun, "qimmat" so'zi salbiy qutblilik qiymatiga ega. U "juda" bilan qo'shilganda, uning qutblilik qiymati juda manfiyga (-5) o'zgaradi, bu qutbning kuchayishini anglatadi. Aksincha, agar "sal" bilan qo'shilsa, u biroz salbiy qiymatni oladi, bu qutbning qisqarishini anglatadi (oldindan hisoblangan qutblilik qiymatini 1 ga kamaytiradi).

Ammo/lekin/biroq (shunga qaramay/garchi): bu bog'lovchilar odatda qarama-qarshi qutbli bo'lishi kerak bo'lgan ikkita ibora o'rtasida bog'lovchi sifatida ishlatiladi. Bu qo'shma gaplarga nisbatan qo'llaniladigan umumiy qoida ikkinchi iboraning qutbligini hisobga olgan holda butun qutblilikni kuchaytirishdan iborat. Yuqorida sanab o'tilgan bog'lovchilar gap ikkinchi qismining ijobiy yoki inkorligiga qarab umumiy qutblilikning 1 ga ortishi yoki kamayishiga sabab bo'ladi (4; 535). Quyida keltirilgan misolda ikkita gapning qutblari alohida hisoblab chiqiladi, keyin esa ikkinchi iboraning manfiy qutbligini hisobga olgan holda bog'lanish tufayli —1 qo'shiladi.

2-jadval.

Tahlil qilish namunasi

Trigram	Kechki ovqat shiringa o'xshaydi,	biroq	qornim og'riyapti	= +1
	trigram	bog'lovchi	trigram	
Qiymati:	2	-1	-2	

O'zbek tilida inkor sifatida qo'llanuvchi qo'shimchalar so'z ma'nosini ijobiydan salbiyga yoki aksincha o'zgartira oladi. Ularga "emas", "be-", "-siz" va belgiga egalikni ko'rsatuvchi "li" kiradi

- sharaf (tub) + li (qo‘shimcha) / sharaflī
- nomus (tub) + siz (qo‘shimcha) / nomussiz

Ma’lumotlar bazasida tub so‘z betaraf sifatida ta’riflangan bo‘lsa-da, “li”, “-siz” qo‘shimchasi tufayli uning qutbliligi manfiy/musbatga aylanadi. 1-jadvalda tub so‘zga qo‘shilganda so‘z ma’nosini o‘zgartiruvchi qo‘shimcha bilan birlashganda qutblarga qanday ta’sir qilishini ko‘rsatadi.

3-jadval.

So‘z darajasidagi tahlil

ASOS+YASOVCHI	YANGI QUTBLANISH
IJOBIY ASOS+INKOR	salbiy
SALBIY ASOS+INKOR	ijobiy
BETARAF ASOS+INKOR	salbiy

Ko‘pgina tadqiqotchilarning hissasi tufayli bir nechta umumiy maqsadli subyektivlik, his-tuyg‘u va hissiyot leksikalari tuzilgan va ularning ba’zilari ham ommaga ochiq, masalan:

- General Inquirer lexicon
- Sentiment lexicon
- MPQA subjectivity lexicon
- SentiWordNet
- EmoBank
- Emotion lexicon
- NRC Emotion Lexicon

4-jadval.

Hissiy so‘zlar ta’minotidan namuna

Emotsional ekspressiv so‘zlar	Betaraf ma’nosi	Ijobiyligi	Salbiyligi
<i>badbashara</i>	yuz		salbiy
<i>aft</i>	yuz		salbiy
<i>turq</i>	yuz		salbiy
<i>husn</i>	yuz	ijobiy	
<i>chehra</i>	yuz	ijobiy	
<i>jamol</i>	yuz		
<i>vallomat</i>	avliyo		salbiy
<i>valiy</i>	avliyo	ijobiy	
<i>shpion</i>	josus		salbiy
<i>ayg‘oqchi</i>	josus		salbiy

Xulosa

Hozirgi kunda biz hayotimizni ijtimoiy tarmoqlarsiz tasavvur qila olmaymiz. Har kim internetdan turli maqsadlarda, ya'ni ma'lumot qidirish yoki biror narsa joylashtirish uchun foydalanadi.

Yuqorida sentiment tahlil qilishning lug'atga va korpusga asoslangan yondashuvlari, ularning o'ziga xos yutuq va kamchilik tomonlari ko'rib chiqildi, lingvistik ta'minotni ishlab chiqishda e'tibor qaratish lozim bo'lgan jihatlar haqida so'z bordi. Xulosa qilib aytganda, korpusga asoslangan yondashuv his-tuyg'ularni aniqlash uchun matnning katta to'plamidagi belgilar va kontekstni tahlil qiladi, lug'atga asoslangan yondashuv esa matn qismining hissiyotini baholash uchun oldindan belgilangan so'zlar ro'yxati va ularning hissiyot ballaridan foydalanadi. Ikkala yondashuv ham o'ziga xos kuchli va cheklovlarga ega va hissiyotlarni tahlil qilishning aniqligini oshirish uchun birgalikda ishlatilishi mumkin.

Foydalanilgan adabiyotlar

1. Abdullayev A. O'zbek tilida ekspressivlikning ifodalanishi (Expression of expressiveness (in Uzbek). — Toshkent: Fan, 1983. — 88 b.
2. Abdurahmonov X., Rafiyev A., Shodmonqulova D. O'zbek tilining amaliy grammatikasi (Practical Grammar of the Uzbek language). — Toshkent : O'qituvchi, 1992. — 214 b.
3. Muhammad A.B., Dahiru A.A.. Lexicon-based sentiment analysis of web discussion posts using SentiWordNet // Journal of Computer Science and Its Application. — 2019. — Vol. 26, No. 2. — URL: https://www.researchgate.net/publication/339279069_Lexicon-based_sentiment_analysis_of_web_discussion_posts_using_SentiWordNet
4. Sharma A., Ghose U. "Lexicon A Linguistic Approach for Sentiment Classification" // 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence). — Noida, India, 2021. — P. 887-893. DOI: 10.1109/Confluence51648.2021.9377057.
5. Hamouda A., Marei M., Rohaim M. "Building Machine Learning Based Senti-word Lexicon for Sentiment Analysis" // Journal of Advances in Information Technology. — 2011. — Vol. 2, No. 4. — Pp. 199-203. DOI:10.4304/jait.2.4.199-203.
6. Sadia, A. Khan F., Bashir F.. An Overview of Lexicon-Based Approach For Sentiment Analysis // Computer Science, Linguistics (3 rd International Electrical Engineering Conference (IEEC 2018)) — Pakistan, 2018. — P. 468-473.
7. Lochter J.V, Zanetti R.F, Reller D. Short text opinion detection using ensemble of classifiers and semantic indexing // Expert Systems with Applications. — 2016. — № 62. — pp. 243-249.
8. Kour K., Kour J., Singh P. (2021). Lexicon-Based Sentiment Analysis. In: Hura, G.S., Singh, A.K., Siong Hoe, L. (eds) Advances in Communication and Computational Technology. — ICACCT, 2019. — P. 1421–1430. https://doi.org/10.1007/978-981-15-5341-7_108

9. K. Ravi, V. Ravi. A survey on opinion mining and sentiment analysis: tasks, approaches and applications // Knowledge-Based Systems. — 2015. — № 89. — pp. 14-46.
<http://dx.doi.org/10.1016/j.knosys.2015.06.015>
10. N. Nigam, D. Yadav. Lexicon-Based Approach to Sentiment Analysis of Tweets Using R Language. In: Singh, M., Gupta, P., Tyagi, V., Flusser, J., Ören, T. (eds) Advances in Computing and Data Sciences. — ICACDS 2018. — pp. 154–164.
https://doi.org/10.1007/978-981-13-1810-8_16
11. Huang M., Xie H., Rao Y., Liu Y., Poon L. K. M. and Wang F. L. "Lexicon-Based Sentiment Convolutional Neural Networks for Online Review Analysis // IEEE Transactions on Affective Computing. — vol. 13, № 3. — pp. 1337-1348, 1 July-Sept. 2022. DOI: 10.1109/TAFFC.2020.2997769.
12. DEHKHARGHANI R., YANIKOGLU B., SAYGIN Y., OFLAZER K. (2016). Sentiment analysis in Turkish at different granularity levels // Natural Language Engineering. — 2016. — № 23(04). — pp. 535–559. DOI:10.1017/s1351324916000309.
13. Consoli S., Barbaglia L., Manzan S. Explaining Sentiment from Lexicon // Conference: X-sentiment@eswc2021. — volume 2918 of CEUR-WS. — pp. 87-95.
14. Jardim S., Mora C., Santana T. "A Multilingual Lexicon-based Approach for Sentiment Analysis in Social and Cultural Information System Data," // 2021 16th Iberian Conference on Information Systems and Technologies (CISTI). — Chaves, Portugal, 2021. — pp. 1-6. DOI:10.23919/CISTI52073.2021.9476631.

References

1. Abdullayev A. *Uzbek tilida ekspressivlikning ifodalanishi* (Expression of expressiveness), Tashkent: Fan, 1983, 88 p. (in Uzbek)
2. Abdurahmonov Kh., Rafiyev A., Shodmonkulova D. *Uzbek tilining amaliy grammatikasi* (Practical Grammar of the Uzbek language), Tashkent : Ukituvchi, 1992, 214 p.
3. Muhammad A.B., Dahiru A.A.. *Journal of Computer Science and Its Application*, 2019, Vol. 26, No. 2, available at: https://www.researchgate.net/publication/339279069_Lexicon-based_sentiment_analysis_of_web_discussion_posts_using_SentiWordNet
4. Sharma A., Ghose U. 2021 *11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, Noida, India, 2021, P. 887-893. DOI: 10.1109/Confluence51648.2021.9377057.
5. Hamouda A., Marei M., Rohaim M. *Journal of Advances in Information Technology*, 2011, Vol. 2, No. 4, pp. 199-203. DOI:10.4304/jait.2.4.199-203.
6. Sadia, A. Khan F., Bashir F. *Computer Science, Linguistics (3 rd International Electrical Engineering Conference (IEEC 2018))*, Pakistan, 2018, pp. 468-473.
7. Lochter J.V, Zanetti R.F, Reller D. *Expert Systems with Applications*, 2016, No 62, pp. 243-249.

8. Kour K., Kour J., Singh P. *Lexicon-Based Sentiment Analysis*. In: Hura, G.S., Singh, A.K., Siong Hoe, L. (eds) *Advances in Communication and Computational Technology*, ICACCT, 2019, pp. 1421–1430. https://doi.org/10.1007/978-981-15-5341-7_108
9. K. Ravi, V. Ravi. *Knowledge-Based Systems*, 2015, No 89, pp. 14-46. DOI: <http://dx.doi.org/10.1016/j.knosys.2015.06.015>
10. N. Nigam, D. Yadav. *Lexicon-Based Approach to Sentiment Analysis of Tweets Using R Language*. In: Singh, M., Gupta, P., Tyagi, V., Flusser, J., Ören, T. (eds) *Advances in Computing and Data Sciences*, ICACDS 2018, pp. 154–164. DOI: https://doi.org/10.1007/978-981-13-1810-8_16
11. Huang M., Xie H., Rao Y., Liu Y., Poon L. K. M. and Wang F. L. *IEEE Transactions on Affective Computing*, vol. 13, No 3, pp. 1337-1348, 1 July-Sept. 2022. DOI: 10.1109/TAFFC.2020.2997769.
12. DEHKHARGHANI R., YANIKOGLU B., SAYGIN Y., OFLAZER K. *Natural Language Engineering*, 2016, No 23 (04), pp. 535–559. DOI:10.1017/s1351324916000309.
13. Consoli S., Barbaglia L., Manzan S. *Conference: X-sentiment@eswc2021*, volume 2918 of CEUR-WS, pp. 87-95.
14. Jardim S., Mora C., Santana T. *2021 16th Iberian Conference on Information Systems and Technologies (CISTI)*, Chaves, Portugal, 2021, pp. 1-6. DOI:10.23919/CISTI52073.2021.9476631.