

O'ZBEK TILIDAGI HISSIY SO'ZLAR LUG'ATI UCHUN KORPUS ASOSIDA BILIMLAR BAZASI TUZISH TAMOYILLARI

Рахимов Хасанбой Комилжонович

ADU 3 - bosqich tayanch doktoranti

Namangan davlat chet tillar instituti stajyor o'qituvchisi

<https://doi.org/10.5281/zenodo.15774239>

Annotatsiya: Ushbu maqolada o'zbek tilidagi hissiy so'zlarning hissiy tahlilga qo'shgan hissasi va ular asosida korpus lingvistikasi orqali ma'lumotlar bazasi yaratishning nazariy-amaliy jihatlarini bayon etilgan. Tadqiqotda hissiy so'zlarni aniqlash, tasniflash va semantik-kontekstual tahlil qilish orqali lug'at shaklida tizimlashtirish muammosi ko'rib chiqiladi. Bundan tashqari, zamonaviy korpus texnologiyalari yordamida tez-tez uchraydigan hissiy birliklarni ajratib olish, ularning qutbiga aniqlik kiritish va ma'lumotlar bazasida (masalan, JSON, XML) saqlash mexanizmlari tasvirlanadi.

Kalit so'zlar: sentiment analiz, hissiy birliklar, korpus lingvistikasi, bilimlar bazasi, emotsional leksika, leksik-semantik tahlil.

Kirish

So'nggi yillarda tabiiy tilni qayta ishlash (NLP) sohasining rivojlanishi tilga oid bir qator tahliliy vazifalarni avtomatlashtirish imkonini berdi. Ayniqsa, sentiment analiz, ya'ni matnlardagi ijobiy, salbiy yoki neytral hissiy munosabatlarni aniqlash yo'nalishi dolzarb tarmoqqa aylandi. Biroq sentiment tahlil tizimlarining samarali ishlashi, eng avvalo, ma'lumotlar bazasida mavjud bo'lgan boy va tizimlashtirilgan leksik resurslarga bog'liqdir. Ingliz, rus, frantsuz tillari uchun bunday hissiy so'zlar lug'atlari mavjud bo'lsa-da, o'zbek tili uchun bu yo'nalishda izchil lug'aviy resurslar mavjud emas.

Shu bois o'zbek tilidagi hissiy birliklarni aniqlash, ularni kontekstual jihatdan tahlil qilish hamda bilimlar bazasiga joylashtirish masalasi nazariy va amaliy nuqtai nazardan dolzarb hisoblanadi. Ushbu maqolada korpus lingvistikasi vositalari yordamida hissiy birliklarni aniqlash va ular asosida tizimli bilimlar bazasi yaratish tamoyillari asoslab beriladi.

1. Hissiy birliklarning turlari va xususiyatlari

Sentimentli birliklar mazmun jihatidan quyidagicha tasniflanadi:

Ijobiy birliklar: yaxshi, chiroyli, go'zal, mehribon;

Salbiy birliklar: yomon, jirkanch, qo'rqinchli, alamli;

Neytral birliklar: rasmiy, befarq, odatiy.

Ushbu birliklar leksik birlik sifatida turli kategoriyalarda namoyon bo'ladi:

Baholovchi sifatlar: yaxshi, yomon, go'zal, jirkanch;

Emotsional otlar: quvonch, g'am, qo'rquv, nafrat;

Ma'no kuchaytiruvchi so'zlar: juda, nihoyatda, anchagina;

Turg'un iboralar: dili xira bo'ldi, osmonga uchdi.

Bu birliklar ko'pincha kontekst asosida ma'no o'zgarishiga ega bo'ladi. Masalan, "juda yaxshi" va "juda xavfli" iboralari bir xil kuchaytiruvchi vositani o'z ichiga olgan bo'lsa-da, turli polaritetga ega. Shu bois hissiy birliklarni tahlil qilishda kontekstual tahlil alohida ahamiyat kasb etadi.

2. Korpus texnologiyasi asosida hissiy birliklarni aniqlash

Korpus lingvistikasi — til birliklarining real matnlarda uchrash chastotasi, kontekstlari va stilistik funksiyalarini aniqlash imkonini beruvchi metodologiyadir. Hissiy birliklarni

aniqlashda quyidagi asosiy bosqichlarga e'tibor qaratiladi:

1. Korpus tanlash: badiiy, ommaviy axborot, ilmiy va ijtimoiy tarmoq matnlari asosida tanlab olinadi;
 2. Frekvent tahlil: eng ko'p uchraydigan hissiy birliklar statistik usullar bilan ajratiladi;
 3. TF-IDF tahlil: korpusdagi ayrim birliklarning hissiy ahamiyat darajasi aniqlanadi;
 4. Kontekstual tahlil: so'zlarning vazifasi, hissiy qiymati va stilistik yuklamalari aniqlanadi;
 5. Crowdsourcing: foydalanuvchilar ishtirokida so'zlarning polaritetini (ijobiy, salbiy, neytral) baholash orqali ishonchli natijalar olinadi.
3. Bilimlar bazasini tuzish tamoyillari
- Aniqlangan hissiy birliklarni tizimli shaklda saqlash uchun ular bilimlar bazasiga joylashtiriladi. Strukturaviy shaklda bu baza quyidagicha ko'rinadi:

So'z	Turkumi	Polaritet	Misol	Sinonim	Antonim	Uslubiy belgi
go'zal	Sifat	ijobiy	go'zal tabiat, go'zal yuz	chiroyli	xunuk	badiiy
jirkanch	sifat	salbiy	jirkanch harakat	yomon	yaxshi	og'zaki
Baxtli	sifat	ijobiy	baxtli hayot, baxtli onalar	xursand	baxtsiz	adabiy
g'azablangan	sifat	salbiy	g'azablangan odam	jahli chiqqan	tinch	og'zaki
maftunkor	sifat	ijobiy	maftunkor qiyofa	jozibakor	oddiy	badiiy
Qabih	sifat	salbiy	qabih so'z	jirkanch	go'zal	adabiy
quvonchli	sifat	ijobiy	quvonchli xabar	sevinchli	qayg'uli	adabiy
Xavfli	sifat	salbiy	xavfli kasallik	tahlikali	xavfsiz	rasmiy
muloyim	sifat	ijobiy	muloyim gap	yumshoq	qo'pol	Adabiy
jahldor	sifat	salbiy	jahldor bola	tez achchiqlanadigan	xotirjam	og'zaki
mehribon	sifat	ijobiy	mehribon ona	g'amxo'r	beparvo	adabiy

Bunday formatda tuzilgan baza JSON, XML yoki SQL formatlarida NLP tizimlarida foydalanish uchun qulay hisoblanadi. Har bir so'zga leksik toifa, polaritet, stilistik xossa, chastota, kontekst kabi atributlar biriktiriladi.

4. Ilmiy va amaliy ahamiyati

Ushbu tadqiqot asosida yaratiladigan hissiy birliklar lug'ati va bilimlar bazasi quyidagi yo'nalishlar uchun muhim resurs bo'lib xizmat qiladi:

1. Sentiment analiz dasturlari;
2. Ijtimoiy tarmoq monitoring tizimlari;
3. Chatbotlar va dialog tizimlari;

4. Avtomatik sharh baholash vositalari;
5. Tarjima va ovozli muloqot tizimlari.

Mazkur resurslar yordamida o'zbek tilidagi NLP tizimlarining rivojlanishi va sun'iy intellekt asosidagi lingvistik modellarni yaratish uchun zarur infratuzilma shakllanadi.

Xulosa

O'zbek tilidagi hissiy birliklarni korpus lingvistikasi asosida aniqlash va ularni tizimlashtirib bilimlar bazasiga joylashtirish sentiment analiz yo'nalishida mustahkam lingvistik asos yaratadi. Tadqiqotda qo'llangan metodlar – frekvent tahlil, TF-IDF, kontekstual yondashuv va crowdsourcing usullari – hissiy so'zlarning ilmiy asoslangan tasnifini ta'minlaydi. Ushbu lug'at va bilimlar bazasi NLP modellarni yaratish, testlash va ularni amaliy tizimlarga joriy qilishda muhim o'rin tutadi.

References:

Используемая литература:

Foydalanilgan adabiyotlar:

1. Pang B., Lee L. (2008). Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval.
2. Liu B. (2012). Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers.
3. Jurafsky D., Martin J. H. (2021). Speech and Language Processing. Pearson.
4. Raximov X. "O'zbek tilidagi adabiy matnlarda sentiment analiz imkoniyatlari". Kompyuter lingvistikasining zamonaviy texnologiyalari. CTCL -2025.
5. Raximov X. " Sentiment tahlil uchun matnlardagi kinoyalarni aniqlash" Xorijiy lingvistika va lingvodidaktika- Foreign Linguistics and Linguodidactics. Special Issue – 4(2024).