



O‘ZBEK TILIDAGI MATNLARNI SENTIMENT TAHLIL QILISH
VA AVTOMATIK TASNIFLASH USULLARINI ISHLAB CHIQISH

Jumagul M.Sulaymonova

O‘zbekiston Milliy universiteti, Toshkent, O‘zbekiston.

Kompyuter lingvistikasi mutaxassisligi 1-kurs magistranti

E-mail: jumagulsulaymonova0003@gmail.com

Toshkent shahri

Annotatsiya: Ushbu tezisda o‘zbek tilidagi matnlarni sentiment tahlil qilish masalasi ko‘rib chiqiladi. Sentiment tahlil matnlardagi hissiy munosabatni aniqlab, ularni ijobiy, salbiy va neytral toifalarga ajratadi. Tadqiqotda turli matnlar lingvistik va statistik yondashuvlar asosida tahlil qilindi. Natijalarga ko‘ra, neytral matnlar ustunlik qilishi, ijobiy va salbiylik esa asosan leksik birliklar orqali ifodalanishi aniqlandi. Shuningdek, o‘zbek tilining morfologik murakkabligi avtomatik tahlil jarayoniga ta’sir ko‘rsatishi kuzatildi.

Kalit so‘zlar: Sentiment tahlil, opinion mining, tabiiy tilni qayta ishlash, NLP, ijobiy matn, salbiy matn, neytral matn, lingvistik tahlil, morfologiya, o‘zbek tili, matn tasnifi.

Abstract: This thesis examines the problem of sentiment analysis of Uzbek-language texts. Sentiment analysis aims to identify the emotional polarity in texts and classify them into positive, negative, and neutral categories. In this study, various texts were analyzed using both linguistic and statistical approaches. The results indicate that neutral texts are predominant, while positive and negative sentiments are mainly expressed through lexical units. It was also observed that the morphological complexity of the Uzbek language significantly affects the automatic text analysis process.



Keywords: Sentiment analysis, opinion mining, natural language processing, NLP, positive text, negative text, neutral text, linguistic analysis, morphology, Uzbek language, text classification.

Kirish. Hozirgi kunda tabiiy tilni qayta ishlash (NLP) sohasida sentiment tahlil muhim yoʻnalishlardan biri hisoblanadi. Ushbu yoʻnalish matnlardagi hissiy munosabatni aniqlashga qaratilgan boʻlib, ijtimoiy tarmoqlar, elektron tijorat, mijozlar fikrini oʻrganish hamda ommaviy axborot vositalarini tahlil qilishda keng qoʻllanilmoqda. Xususan, matnlarning ijobiy, salbiy yoki neytral ekanligini aniqlash orqali katta hajmdagi maʼlumotlarni tez va samarali tahlil qilish imkoniyati yaratiladi. Bugungi kunga qadar ingliz tilida sentiment tahlil boʻyicha koʻplab tadqiqotlar amalga oshirilgan boʻlib, turli algoritmlar va dasturiy yechimlar ishlab chiqilgan. Ushbu yoʻnalishda Machine Learning va Deep Learning usullari keng qoʻllanilib, yuqori aniqlikka ega modellar yaratilgan. Oʻzbek tilida esa bu soha hali toʻliq rivojlanmagan boʻlib, mavjud tadqiqotlar va amaliy yechimlar soni cheklangan. Ayniqsa, sentiment tahlilni amalga oshirish uchun zarur boʻlgan lingvistik resurslar — matnlar bazasi, annotatsiyalangan maʼlumotlar toʻplami, lugʻatlar va tayyor modellar — yetarli darajada emas. Oʻzbek tili agglutinativ xususiyatga ega boʻlgani sababli, soʻzlarning qoʻshimchalar orqali oʻzgarishi avtomatik tahlil jarayonini murakkablashtiradi. Shu sababli, ushbu yoʻnalishda ilmiy tadqiqotlar olib borish va zamonaviy texnologiyalar asosida yangi yechimlar ishlab chiqish dolzarb vazifa hisoblanadi. Ushbu ishning maqsadi — oʻzbek tilidagi matnlarni avtomatik tarzda sentiment tahlil qiluvchi lingvistik platforma yaratishdan iborat. Mazkur platforma matnlarni qayta ishlash, ularni tahlil qilish hamda natijalarni foydalanuvchiga qulay shaklda taqdim etishga xizmat qiladi.

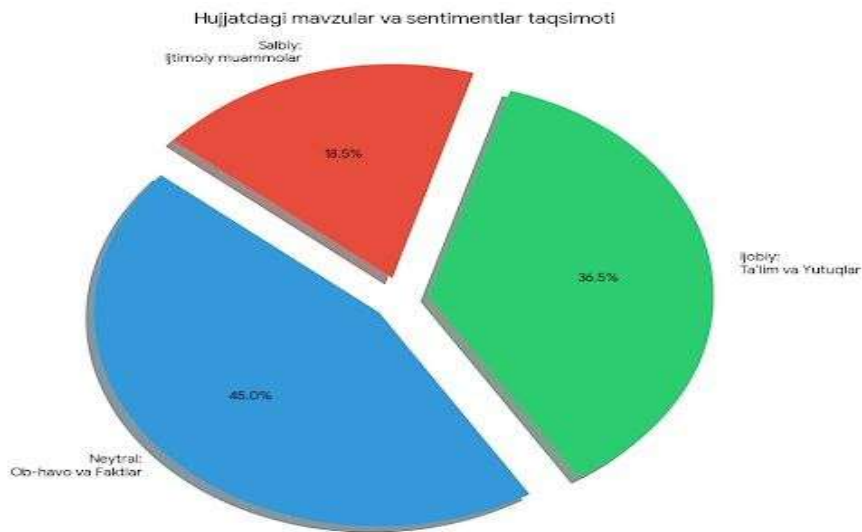
Adabiyotlar tahlili va metodlar. **Sentiment tahlil (yoki opinion mining)** — bu insonlarning turli obyektlarga (mahsulotlar, xizmatlar, tashkilotlar va boshqalar) nisbatan bildirgan fikr, hissiyot va baholarini aniqlash hamda tahlil qilishga qaratilgan muhim ilmiy yoʻnalish hisoblanadi. Ushbu soha 2000-

yillardan boshlab jadal rivojlana boshlagan bo‘lib, hozirgi kunda tabiiy tilni qayta ishlash (NLP), data mining hamda veb-ma‘lumotlar tahlili sohalarining ajralmas qismiga aylangan. Dastlabki tadqiqotlar orasida Sanjiv Das va Mike Chen ishlari alohida ahamiyatga ega bo‘lib, ular internetdagi matnlar asosida moliyaviy bozor kayfiyatini aniqlashga qaratilgan. Shuningdek, Bo Pang, Lillian Lee va Shivakumar Vaithyanathan (2002) tomonidan olib borilgan tadqiqotlar sentiment tahlilni mashinaviy o‘rganish asosida hujjat darajasida tasniflashga asos soldi[3]. “Sentiment analysis” atamasi ilk bor Tetsuya Nasukawa va Jeonghee Yi tomonidan qo‘llanilgan bo‘lsa, “opinion mining” atamasi Kushal Dave, Steve Lawrence va David M. Pennock ishlarida uchraydi[6]. Bugungi kunda ushbu ikki atama deyarli bir xil ma‘noda ishlatiladi. Bing Liu (2012) o‘zining “Sentiment Analysis and Opinion Mining” asarida ushbu sohaning nazariy asoslarini batafsil yoritgan. Unda sentiment tahlil uchta asosiy darajada ko‘rib chiqiladi: **hujjat darajasi (Document level)** — butun matnni ijobiy yoki salbiy deb tasniflash; **gap darajasi (Sentence level)** — har bir gapning sentimentini aniqlash; **aspekt darajasi (Aspect level)** — obyektning alohida xususiyatlariga nisbatan fikrlarni aniqlash[1]. Ayniqsa, Mingqing Hu va Bing Liu (2004) tomonidan taklif etilgan aspekt darajasidagi tahlil yanada chuqur va aniq natijalar berishi bilan ajralib turadi. Bu yondashuvda fikrlar nafaqat obyektga, balki uning alohida jihatlariga (masalan, sifat, narx, xizmat) nisbatan ham aniqlanadi. Bundan tashqari, Nitin Jindal and Bing Liu tadqiqotlarida fikrlar ikki turga — oddiy va taqqoslovchi fikrlarga ajratilgan. Bu yondashuv sentiment tahlilni yanada murakkab va real hayotga yaqinlashtiradi[7]. Umuman olganda, sentiment tahlil nafaqat NLP, balki iqtisodiyot, siyosatshunoslik va marketing kabi sohalarda ham muhim ahamiyat kasb etadi[1]. Ushbu tadqiqotda sentiment tahlilni amalga oshirish uchun bir nechta asosiy metodlardan foydalanilgan. Birinchidan, matnlarni tahlil qilishda **lingvistik yondashuv** qo‘llanilgan. Bu yondashuv doirasida matnlar sintaktik va semantik jihatdan tahlil qilinib, sentiment ifodalar aniqlanadi[1]. Ayniqsa, sentiment ifodalar va subyektlar

o'rtasidagi bog'liqlikni aniqlash muhim ahamiyat kasb etadi[10]. Ikkinchidan, **leksikon asosidagi metod**dan foydalanildi. Bu metodda oldindan tuzilgan sentiment lug'ati yordamida matndagi ijobiy va salbiy so'zlar aniqlanadi va ularning polariteti belgilanadi[10]. Ushbu yondashuv sodda va tushunarli bo'lishiga qaramay, kontekstni hisobga olishda ayrim cheklovlarga ega[3]. Uchinchidan, **mashinali o'rganish usullari** (machine learning) ham muhim rol o'ynaydi. Bu usullarda matnlar oldindan belgilangan (annotatsiya qilingan) ma'lumotlar asosida o'rgatiladi va yangi matnlar uchun sentimentni avtomatik aniqlash imkonini beradi[3]. Shuningdek, ushbu ishda sentiment tahlilning **aspekt darajasidagi modeli** asos qilib olindi. Bu modelda matndan nafaqat umumiy sentiment, balki aniq obyekt va uning xususiyatlariga nisbatan fikrlar ajratib olinadi. Masalan, bitta jumlada bir vaqtning o'zida ijobiy va salbiy fikrlar mavjud bo'lishi mumkinligi inobatga olinadi[10]. Natijada, tanlangan metodlar kombinatsiyasi orqali matnlarni chuqur va aniq tahlil qilish, sentimentlarni to'g'ri aniqlash hamda ularni strukturalangan ko'rinishda taqdim etish imkoniyati yaratildi.

Natijalar. Ushbu tadqiqotda matnlarni sentiment tahlil qilish orqali ularning ijobiy, salbiy va neytral hissiy yo'nalishlari aniqlandi. Tahlil jarayonida turli turdagi yangilik matnlari, ijtimoiy va axborot mazmunidagi gaplar hamda murakkab lingvistik tuzilishga ega o'zbek tilidagi matnlar ko'rib chiqildi. Natijalar shuni ko'rsatdiki, matnlarning sezilarli qismi neytral yoki axborot xarakteriga ega bo'lib, ularda aniq hissiy yuklama (emotsional ifoda) kuchsiz yoki umuman mavjud emas. Bunday matnlar asosan yangilik sarlavhalari, rasmiy xabarlar va statistik ma'lumotlardan iborat bo'ldi. Ijobiy sentimentga ega matnlarda ko'pincha maqtov, rivojlanish, ijobiy o'zgarishlar yoki muvaffaqiyatni ifodalovchi leksik birliklar uchradi. Bu turdagi matnlar odatda optimistik ohangga ega bo'lib, ijtimoiy yoki iqtisodiy yutuqlarni ifodalashda ko'proq kuzatildi. Salbiy sentimentli matnlarda esa muammo, tanqid, xavotir, yo'qotish yoki salbiy holatlarni bildiruvchi so'z va iboralar

ustunlik qildi. Ayniqsa, ijtimoiy muammolar, iqtisodiy qiyinchiliklar va siyosiy mavzulardagi matnlarda salbiy ohang ko‘proq sezildi.



Tahlil jarayonida o‘zbek tilining lingvistik xususiyatlari muhim rol o‘ynashi aniqlandi. Xususan, morfologik murakkablik, so‘z shakllarining ko‘pligi hamda affikslarning turli kombinatsiyalari avtomatik sentiment aniqlash jarayonini murakkablashtiradi. Bundan tashqari, sinonimlar va kontekstga bog‘liq ma’no o‘zgarishi ham tahlil aniqligiga ta’sir ko‘rsatadi. Shuningdek, ayrim hollarda kontekstni to‘liq hisobga olmaslik natijasida neytral va ijobiy yoki salbiy matnlar orasida kichik xatoliklar kuzatildi. Bu esa sentiment tahlil tizimlarida semantik tushunish darajasini yanada rivojlantirish zarurligini ko‘rsatadi. Umuman olganda, o‘tkazilgan tahlil natijalari sentiment analiz usullari o‘zbek tilidagi matnlar uchun ham samarali qo‘llanilishi mumkinligini tasdiqladi, biroq tilning murakkab tuzilishi qo‘shimcha lingvistik resurslar va takomillashtirilgan algoritmlarni talab qiladi.

Muhokama. O‘tkazilgan sentiment tahlil natijalari shuni ko‘rsatdiki, matnlarning hissiy yo‘nalishini aniqlashda nafaqat leksik birliklar, balki kontekst ham muhim ahamiyatga ega. Ayniqsa, o‘zbek tilidagi matnlarda so‘zlarning ko‘p ma’noliligi va morfologik o‘zgaruvchanligi sentiment aniqlash jarayonini murakkablashtiradi. Tahlil davomida neytral matnlarning ustunligi kuzatildi. Bu holat asosan axborot beruvchi matnlar — yangiliklar, rasmiy

xabarlar va statistik ma'lumotlar ko'pligi bilan izohlanadi. Bunday matnlarda emotsional baho deyarli mavjud bo'lmagani sababli, ular avtomatik tizimlar tomonidan ko'pincha neytral sifatida belgilandi. Ijobiy va salbiy sentimentlarni aniqlashda esa aniq emotsional yuklama beruvchi so'zlar muhim rol o'ynadi. Biroq ayrim hollarda ironiya, kontekstga bog'liq ma'no yoki yashirin baholash mavjud bo'lgan matnlarda xatoliklar kuzatildi. Bu esa oddiy lug'atga asoslangan yondashuvlarning cheklanganligini ko'rsatadi. Shuningdek, o'zbek tilining lingvistik xususiyatlari — affikslar tizimi, so'z yasash modellari va erkin so'z tartibi — sentiment tahlil algoritmlarini moslashtirish zarurligini ko'rsatdi. Bu holat NLP tizimlarida maxsus o'zbek tiliga mos resurslar yaratish muhimligini tasdiqlaydi. Umuman olganda, natijalar sentiment tahlilning samarali ishlashi uchun faqat algoritmik yondashuv emas, balki lingvistik bilimlar va kontekstual tahlil ham zarurligini ko'rsatadi.

Xulosa. Ushbu tadqiqotda o'zbek tilidagi matnlar asosida sentiment tahlil o'tkazildi va ularning hissiy yo'nalishlari aniqlashga harakat qilindi. Olingan natijalar sentiment analiz usullarining amaliy jihatdan qo'llanilishi mumkinligini tasdiqladi. Tahlil jarayonida matnlar uch asosiy guruhga — ijobiy, salbiy va neytral sentimentga ajratildi. Eng ko'p uchragan toifa neytral matnlar bo'lib, bu axborot xarakteridagi matnlarning ustunligi bilan bog'liq ekanligi aniqlandi. Shuningdek, tadqiqot davomida o'zbek tilining morfologik murakkabligi va kontekstga bog'liq ma'no o'zgarishlari sentiment tahlil aniqligiga sezilarli ta'sir ko'rsatishi aniqlandi. Bu esa kelgusida yanada rivojlangan, kontekstni hisobga oladigan va tilga moslashtirilgan algoritmlarni ishlab chiqish zarurligini ko'rsatadi. Xulosa qilib aytganda, sentiment tahlil o'zbek tilidagi matnlarni avtomatik qayta ishlashda muhim yo'nalishlardan biri hisoblanadi. Ushbu sohani rivojlantirish axborot oqimini tahlil qilish, ijtimoiy fikrni o'rganish va matnlarni intellektual qayta ishlashda katta amaliy ahamiyatga ega.

ADABIYOTLAR RO‘YXATI (REFERENCES)

1. Liu B. *Sentiment Analysis and Opinion Mining*. — Morgan & Claypool Publishers, 2012.
2. Jurafsky D., Martin J. H. *Speech and Language Processing*. — 3-nashr (draft), 2023.
3. Pang B., Lee L. Opinion mining and sentiment analysis // *Foundations and Trends in Information Retrieval*. — 2008. — Vol. 2, № 1–2. — P. 1–135.
4. Feldman R. Techniques and applications for sentiment analysis // *Communications of the ACM*. — 2013. — Vol. 56, № 4. — P. 82–89.
5. Mohammad S. M., Turney P. D. Crowdsourcing a word–emotion association lexicon // *Computational Intelligence*. — 2013. — Vol. 29, № 3. — P. 436–465.
6. Nasukawa T., Yi J. Sentiment analysis: Capturing favorability using natural language processing // *Proceedings of K-CAP*. — 2003.
7. Turney P. D. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews // *Proceedings of ACL*. — 2002.
8. Jindal N., Liu B. Mining comparative sentences and relations // *Proceedings of AAAI*. — 2006.
9. Jindal N., Liu B. Identifying comparative sentences in text documents // *Proceedings of SIGIR*. — 2006.
10. Hu M., Liu B. Mining and summarizing customer reviews // *Proceedings of KDD*. — 2004.
11. <https://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf>
12. <https://www.scirp.org/reference/referencespapers?referenceid=1830388>
13. <https://dl.acm.org/doi/abs/10.1145/945645.945658>