

HIS-TUYG‘ULARNI TAHLIL QILISH USULLARI

Allaberganova Nasiba Muradovna

TATU doktoranti

abisan_1990@mail.ru,

Abdrimova Feruza Kamiljanovna

UrDU o‘qituvchisi

ANNOTATSIYA

Ushbu maqolada hozirgi kunda his-tuyg‘ularni tahli qilish qanchalik muhim ekanligi, tuyg‘ularni tahlil qilish usullari keltirilgan. His-tuyg‘ularni tahlil qilish usullarining afzalliklari va kamchiliklari tahlil qilingan. Avtomatlashtirilgan va leksikaga asoslangan hissiyotlarni tahlil qilish usullari o‘zaro taqqoslangan.

Kalit so‘zlar: *Sentiment analiz, Leksikaga asoslangan yondashuv, Mashinaga asoslangan yondashuv, Gibrid usul, so‘zlarni teglash, Korpus.*

SENTIMENT ANALYSIS METHODS

ABSTRACT

This article discusses the methods of analyzing emotions and importance of analyzing them in today’s world. Also the research work analyses advantages and disadvantages of sentiment analysis methods as well as comparing automated and lexical-based sentiment analysis methods.

Key words: *Sentiment analysis, Lexicon-based approach, Machine learning approach, Hybrid approach, Tagging words, Corpus.*

МЕТОДЫ АНАЛИЗА ЭМОЦИЙ

АННОТАЦИЯ

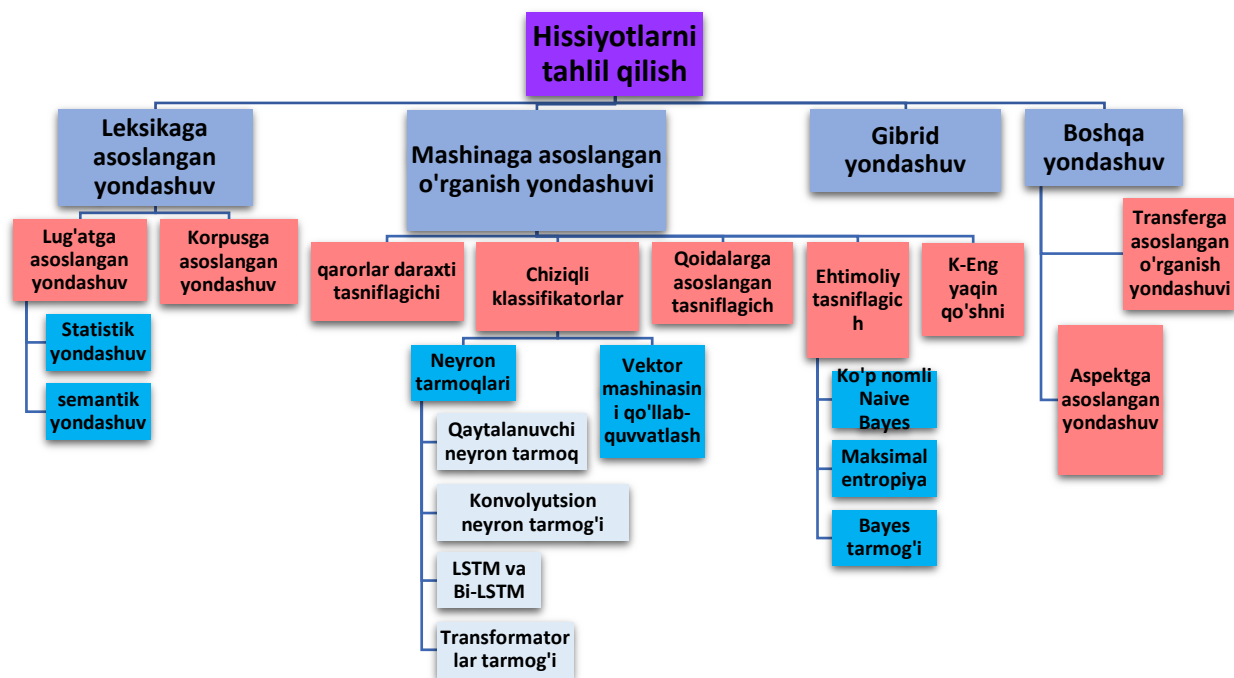
В данной статье приведены важность анализа эмоций в наши дни и методы анализа. Изучены преимущества и недостатки методов анализа эмоций. Сравнены методы анализа автоматизированных и основанных на лексике эмоций.

Ключевые слова: *Сентимент анализ, Лексиконный подход, Подход к машинному обучению, Гибридный подход, Тегирование слов, Корпус.*

Mijozlarning uchdan bir qismi faqat bittagina yomon mahsulot yoki xizmatdan keyin o'zlari yoqtirgan brendlar bilan biznes qilishni to'xtatishlarini aytishadi. Darhaqiqat, mijozlarning mahsulot yoki xizmatlarga nisbatan munosabatini tushunish biznesda muvaffaqiyatga erishishning asosiy omilidir. Korxonalar mijozlarning his-tuyg'ularini tushunish va shunga mos ravishda mahsulot va xizmatlarini yaxshilab borish maqsadida hissiyotlarni tahlil qilish (sentiment analysis) usullaridan foydalanadilar.

Tuyg'ularni tahlil qilish - bu matndagi his-tuyg'ularni aniqlashga yordam beradigan tabiiy tilni qayta ishlash (NLP) usulidir. Ijtimoiy tarmoqlardagi postlar, so'rovlar yoki sharhlardagi his-tuyg'ularni tasniflash orqali korxonalar o'zlarining strategiyalari qanday ishlashini o'lchashlari va o'sish uchun yangilarini aniqlashlari mumkin. Tuyg'ularni tahlil qilishning bir necha usullari mavjud bo'lib, ularning har biri o'zining kuchli va zaif tomonlariga ega.

Quyida biz hissiyotlarni tahlil qilish metodlari va ularning har birining afzalliklari va kamchiliklari haqida umumiy ma'lumot beramiz.



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

1. Leksikaga asoslangan yondashuv

Leksikaga asoslangan yondashuv bu bilimga asoslangan yondashuv bo‘lib, qo‘lda oldindan ishlab chiqilgan sintaktik (ya‘ni grammatikaga oid sintaksisdagi qoliplar) va semantik qoliplarni tahlil qilishga murojaat qiladi. Birinchisi so‘zlarni teglash orqali lug‘at yaratishni nazarda tutsa, ikkinchisi sintaktik qoliplarni ko‘rib chiqishni o‘z ichiga oladi. Lug‘at metodlari so‘zlarni teglash orqali lug‘at yaratadi va korpusga asoslangan metodlar sintaktik qoliplarni ko‘rib chiqishni o‘z ichiga oladi.

Matnning hissiyot bahosi quyidagilar bilan belgilanadi:

- Har bir so‘zlarga hissiy muvofiqlikka asoslangan holda baho berish.
- Gapning qarama qarshi qutbliligini umumiy hisoblash
- Matndagi barcha gaplarning umumiy qutblilik baholarini yig‘ish

Ijobiy tomonlari:

- Leksikaga asoslangan his-tuyg‘ularni tahlil qilish metodlariga osongina kirish mumkin, chunki ko‘plab ochiq manbalar (masalan, SentiWordNet) mavjud.

- Ular biroz arzonroq chunki ular ilg'or hissiyotlarni tahlil qilish algoritmlarini amalga oshirishni talab qilmaydi.

- Kompaniyalar lug'atga asoslangan yondashuvdan foydalansalar o'quv ma'lumotlariga ehtiyoj qolmaydi, chunki teglar qo'lda aniqlanadi va so'zlarning ma'nosiga tezkor kirish imkoniyati mavjud.

Salbiy tomonlari:

- Leksikaga asoslangan his-tuyg'ularni tahlil qilish usullari odatda kinoya, inkor, grammatik xatolar, noto'g'ri imlo yoki kesatiqni aniqlamaydi. Shunday qilib, u ijtimoiy tarmoq platformalaridan to'plangan ma'lumotlarni tahlil qilish uchun mos kelmasligi mumkin.

- Bu metod butunlay teglar va qoidalarga asoslanganligi sababli, kompaniyalar ishonchli lug'at yaratish uchun yetarli ma'lumotlarga ega bo'lishi kerak.

- Gapda strukturasi juda ham qat'iy, dominatsiya bo'lib turgan sohaga qaram va vaziyat qanday bo'lishidan qat'iy nazar so'z bir xil ma'no anglatadi. Masalan, "ajoyib" atamasi vaziyatga qarab ijobiy yoki salbiy bo'lishi mumkin.

- Ular insoniy tarafdorlikka moyil. Misol uchun, agar lug'at tayyorlayotgan odamlar faoliyat sohasi bo'yicha yetarli bilimga ega bo'lmasa, usul aniq natijalarni bermaydi.

- Teglash qo'lda bajarilganligi sababli, ma'lumotlarni tayyorlash ko'p vaqt talab qilishi mumkin.

2. Avtomatlashtirilgan/Mashinani o'rganish usullari (Machine learning)

Avtomatlashtirilgan his-tuyg'ularni tahlil qilish usullari statistik modellar asosida hissiyotlarni tasniflaydigan ML algoritmlarini o'z ichiga oladi. Mashinani o'rganish algoritmlarini amalga oshirish uchun gaplarni vektor maydoniga akstirish kerak. Keyin modellar gapning hissiyotini bashorat qilishga o'rgatilgandek o'zini tutishi mumkin.

Ijobiy tomonlari:

Mashinani o'rganish algoritmlari(Machine learning algorithms):

- Tuygʻularni tahlil qilishda kinoya, kesatiq va negativlikni aniqlashga oʻrgatish mumkin. Bu ijtimoiy tarmoq hissiyotlarini tahlil qilishni osonlashtirishi mumkin.

- Soʻzlarning samarador valentligini oʻrganish, shuningdek ular oldindan belgilangan maʼlumotlar toʻplamini talab qilmaydi.

- Anʼanaviy his-tuygʻularni tahlil qilish usullaridan tezroq.

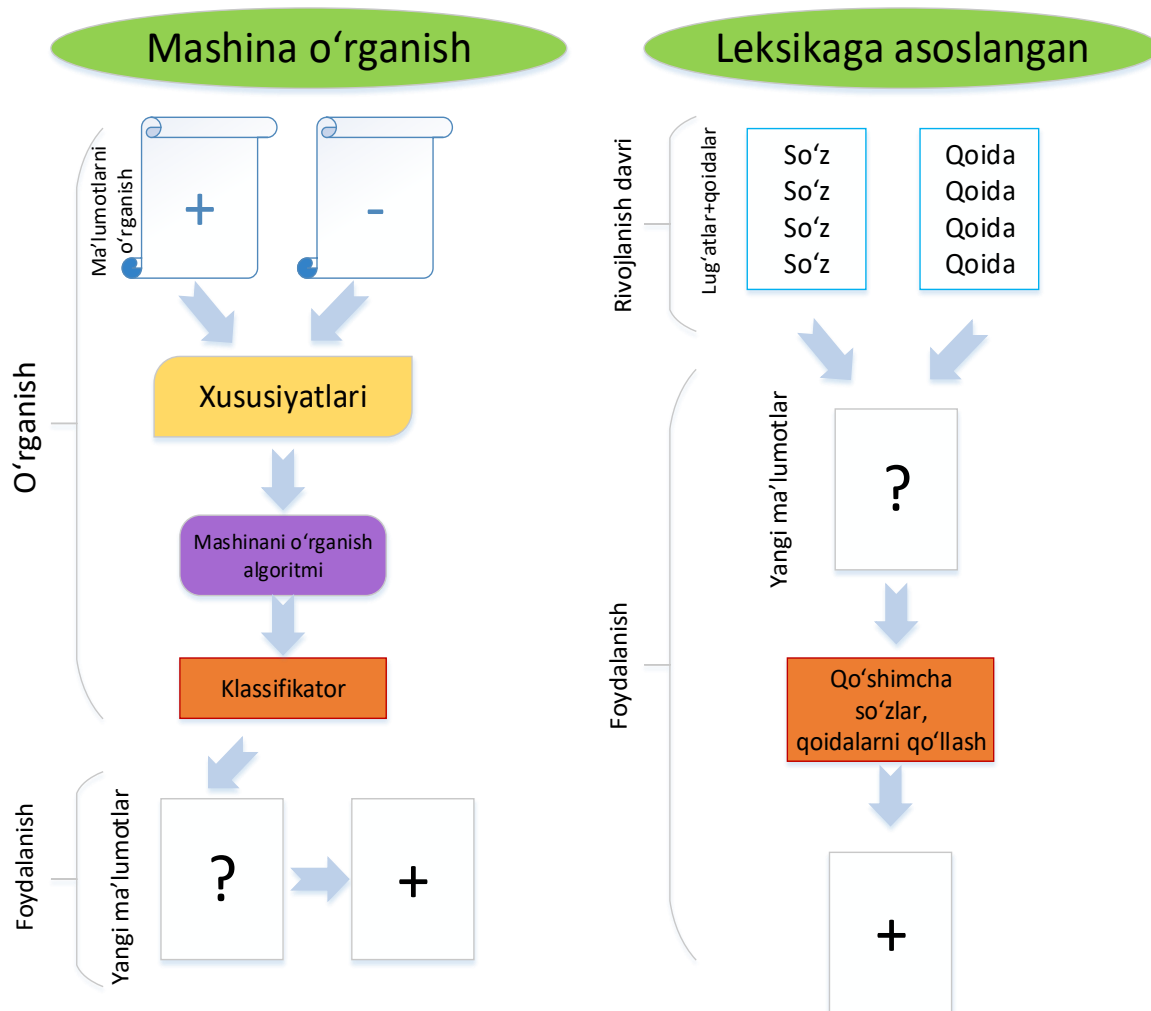
- Aniqroq natijalarni taqdim etadi.

Salbiy tomonlari:

- Toʻgʻri tasniflash uchun kompaniyalar katta yoki yuqori sifatli kichik maʼlumotlar toʻplamiga muhtoj.

- Shovqin (masalan, smayliklar, jargon yoki tinish belgilari) aniqlikni kamaytirishi mumkin.

- Xarajatlar anʼanaviy, qoidalarga asoslangan usullarga nisbatan yuqoriroq.



2-rasm. Avtomatlashtirilgan va leksikaga asoslangan hissiyotlarni tahlil qilish usullarini taqqoslash.

Algoritmarga misollar

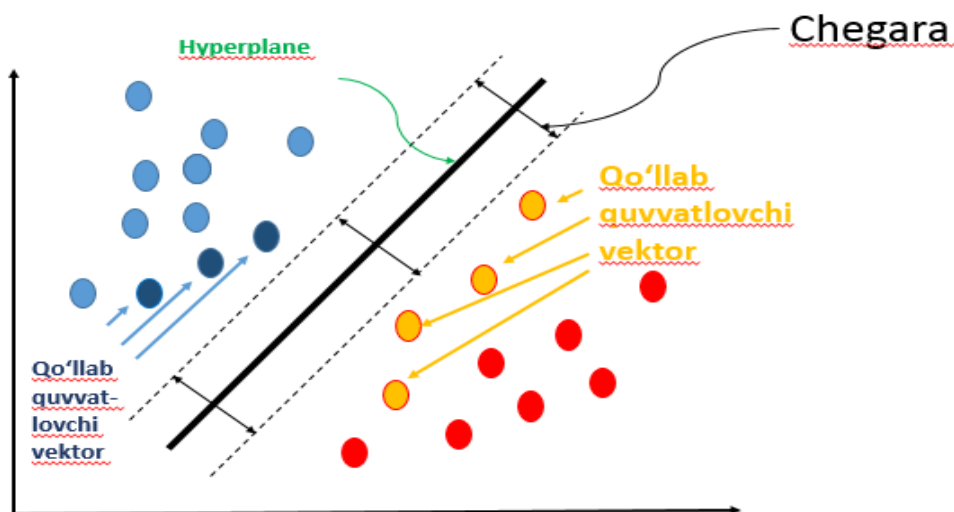
Tuyg'ularni tahlil qilishda ishlatiladigan mashinani o'rganish(ML) algoritmlari quyidagilarni o'z ichiga oladi.

Naive Bayes

Bu Bayes teoremasiga asoslangan nazorat ostida, ehtimolli tasniflash usuli bo'lib, xususiyatlarni ajratib olish uchun ishlatiladi. Ushbu yondashuv har bir token yoki xususiyat boshqasidan mustaqil ekanligini nazarda tutadi, shuning uchun u har qanday noto'g'ri imlo yoki grammatik xatolarni bartaraf etishi mumkin.

Vektor mashinasini qo'llab-quvvatlash

Bu turli vektorlar orasidagi eng yaxshi giperplanni aniqlaydigan, maksimal chegara masofasidan foydalangan holda toifalar o'rtasida qat'iy chegaralarni o'rnatishga imkon beruvchi, ehtimollik bo'lmagan tasniflash usuli.



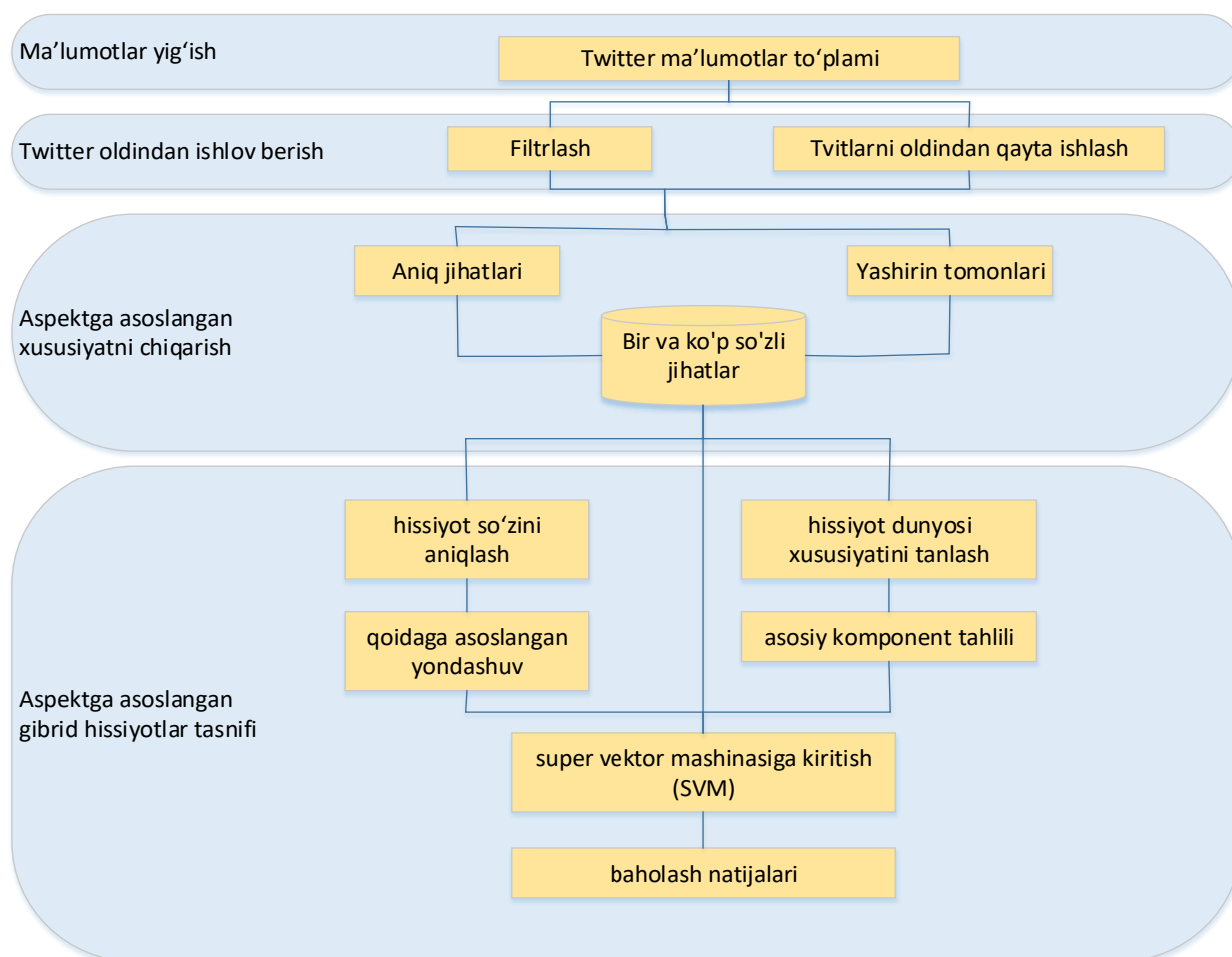
3-rasm. Qo'llab-quvvatlash vektor mashinasi qanday ishlashining ko'rinishi So'zni vektorga akslantirish.

Bu usul tez-tez kontekstda qo'llangan so'zlarni ko'rib chiqish usuli. Vektorlar neyron tarmoqlarga o'xshash tarzda o'rganiladi, shuning uchun so'zlarni joylashtirish chuqur o'rganish usuli hisoblanadi.

Birinchiidan o'lchamlar soni aniqlanadi, har bir so'z pozitsiyasining vektor qiymati fazoda akslantiriladi. Tez-tez ishlatiladigan so'zlar bir-biriga yaqinroq bo'lsa-da, kam birga qo'llaniladigan so'zlar ular orasida uzoq masofaga ega. So'zni akslantirish modellari eng zamonaviy natijalarni taqdim etish imkoniyatiga ega, ammo resurslarni talab qiladigan mexanizm tufayli ular boshqa usullarga qaraganda qiyinroq va qimmatroq.

3. Gibridd yondashuvlar.

Leksikaga asoslangan va avtomatlashtirilgan usullarning afzalliklari va kamchiliklari mavjud. Shunday qilib, kompaniyalar turli yondashuvlar bir-birining kamchiliklarini qoplashi uchun avtomatlashtirilgan va leksikaga asoslangan usullarni o'z ichiga olgan gibridd usullarni amalga oshirishlari mumkin. Kombinatsiya parallel yoki tahlilning turli bosqichlarida bo'lishi mumkin.



4-rasm. Gibridd hissiyotlar tasnifiga misol.

XULOSA

Xulosa qilib aytadigan bo‘lsak hissiyotlarni tahlil qilishning bir qancha usullari mavjud bo‘lib, ularning har birining o‘ziga xos kamchiliklari va yutuqlari mavjud. Demak biz hissiyotlarni tahlil qilishda gibrid hissiyotlarni tahlil qilish usulidan foydalansak maqsadga muvofiq bo‘ladi. Chunki bu usulda bir nechta usul bir vaqta qo‘llaniladi va ular bir-birining kamchiliklarini to‘ldirib, aniq natijalarga ega bo‘lishimizga olib keladi.

FOYDALANILGAN ADABIYOTLAR

1. <https://link.springer.com/article/10.1007/s10462-022-10144-1>
2. <https://research.aimultiple.com/sentiment-analysis/>
3. https://www.researchgate.net/publication/283954600_Sentiment_Analysis_An_Overview_from_Linguistics
4. <https://www.callbridge.com/uz/blog/how-to-use-a-sentiment-analysis-tool/>