ISSN 2181-922X

UZBEKISTAN LANGUAGE & CULTURE

O‘ZBEKISTON TIL VA MADANIYAT

KOMPYUTER LINGVISTIKASI

2023 Vol. 1 (6)

www.compling.tsuull.uz

MUNDARIJA

Statistik usullar yordamida turli so‘z turkumlari orasidagi omonimiyani aniqlash

Xolisa Axmedova[1]

Annotatsiya

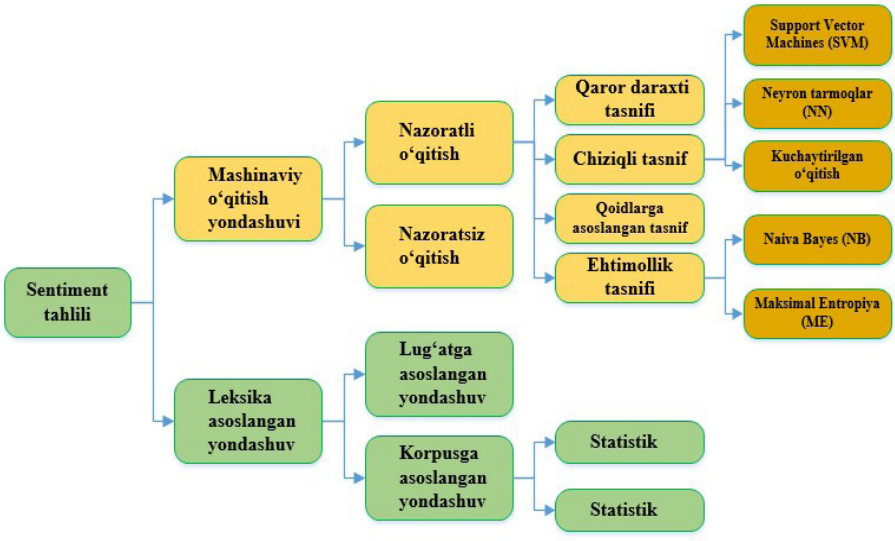
Tabiiy tilni avtomatik qayta ishlash muammosi keyingi yarim asrdan ortiq davr mobaynida dolzarb bo‘lib qolmoqda. NLP sohasining shunday muhim muammolaridan biri semantik analizator yaratish bo‘lib, u ham o‘z navbatida qator qadamlarni bosib o‘tadi. Gaplarni semantik tahlil qilishda omonimiyani aniqlash muhim ahamiyatga ega. Omonimiyani aniqlashda statistik usullardan ham foydalaniladi. Grammatik jihatdan o‘xshash bo‘lgan so‘z turkumlari orasidagi omonimiyani aniqlashda chastotali usuldan foydalaniladi. Ushbu usul omonim so‘zlarni tasniflash paramertlarini ajratib olishni taqazo etadi.

Kalit so‘zlar: NLP, semantik analizator, omonimiya, statistik ma’lumotlarga asoslangan usul, tasniflash parametrlari, chastotali usul, unigram.

Tabiiy tilni qayta ishlash (Natural Language Processing (NLP)) – bu sun’iy intellekt va matematik tilshunoslikning umumiy yo‘nalishi bo‘lib, tabiiy tillardagi gaplarni kompyuterda analiz va sintez qilish muammolarini o‘rganadi. Ushbu muammoni hal qilish inson va kompyuter o‘rtasida o‘zaro aloqaning yanada qulay shaklini yaratishni anglatadi.

Tabiiy tilni avtomatik qayta ishlash muammosi keyingi yarim asrdan ortiq davr mobaynida dolzarb bo‘lib qolmoqda. Muammoning murakkabligi va aniq g‘oyaning yo‘qligi uni hal qilish yo‘llarining qiyinligini bildiradi. Lingvistik analizatorlar esa gapga avtomatik ishlov berish vositalari sifatida alohida ahamiyat kasb etadi. Lingvistik analizatorlar morfologik, sintaktik va semantik analizatorlarga bo‘linadi. Ular ham o‘z navbatida bir nechta guruhlarga bo‘linadi.

[1] *Axmedova Xolisxon Ilxomovna*-tayanch doktorant, Alisher Navoiy nomidagi Toshkent davlat O‘zbek tili va adabiyoti universiteti.
E-pochta: xolisa0629@gmail.com
ORCID: 0000-0002-9828-1650



1-rasm: Lingvistik analizatorlar va ularning elementlari iyerarxiyasi

1-rasmdagi iyerarxiyada tasvirlangan har bir lingvistik analizator o‘z vazifasiga ega.

Morfologik analizator – lug‘atdagi alohida so‘z/so‘z shakllarni taqqoslovchi va so‘zlarning grammatik xususiyatlarini aniqlovchi algoritmlar majmui.

Sintaktik analizator- bu kirish tilini xususiyatlaridan kelib chiqib kiritilgan gapni sintaktik tahlil qiluvchi axborot tizimidir.

Semantik analizator- gapdagi so‘zlarning ma’nolari orqali gap mazmunini aniqlashni amalga oshiruvchi axborot tizimidir. Semantik tahlil orqali semantik qidiruv amalga oshiriladi. Semantik tahlil qanchalik mukammal ishlab chiqilgan bo‘lsa, qidiruv shunchalik samarali natija beradi. Semantik tahlilni amalga oshirish bevosita lingvistik resurslarga bog‘liq. Leksik resurslarga lug‘at, tezaurus va ontologiyalar kiradi. Semantik analiz ham o‘z elementlari ega bo‘lib, ularni alohida o‘rganishni talab qiladi. Semantik analizning muhim elementlaridan biri omonimiya hosidasidir. Omonimiyani aniqlash turli tabiiy tillarda turlicha talqin qilinadi. Jahon kompyuter lingvistikasida gaplarni semantik tahlil qilishda asosan 3 ta usuldan foydalanilgan:

– Qoidalarga asoslangan usul
– Statistik ma’lumotlarga asoslangan usul
– Mashinali o‘qitishga asoslangan usul
– Neyron tarmoqlar metodlari

Ushbu usullar turli tillarda turlicha foydalanilgan. Masalan,

Rus tilshunosligida omonimlarni o‘rganishga bag‘ishlangan tadqiqotlar juda ko‘p. Omonimiya hodisasi A.A.Poroxin [2013], D.N.Gomon [2004], D.A. Mixaylovna [2015], P. Boris Kobritsov, Olga N. Lashevskaja, Olga Ju. Shemanaeva [2011], A.I. Bolshakova [2003], B. A. Bobnev (2010), S.V.Rysakov, E.S.Klyshinskiy [2015] , A.V. Gashkov [2012]va boshqalarning ishlarida maxsus tadqiq etilgan. Rus tilida omonim so‘zlarni farqlashda statistik usullardan foydalanilgan.

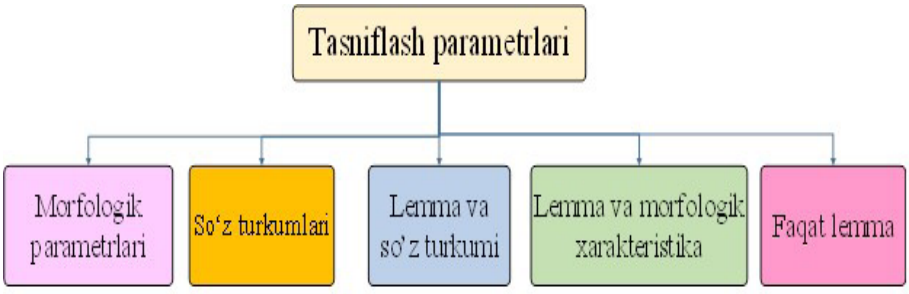
Omonimiya Baltabayeva J.K. va Sulaymonova J.N (qozoq tili) [2019], Ch.A.Davlyatova (tojik tili) [2017], V.V. Kukanova(boshqird tili) [2014], H.Geydarova(ozarbayjon tili) [2017] V.V. Kukanova (qalmoq tili) [2011] kabi turkolog olimlarning tadqiq mavzusi bo‘ldi. Bu turkolog olimlar A.I.Smirniskiy, V.V.Vinogradov, O.S.Axmanova va boshqa tilshunos olimlar ishlab chiqqan omonimiya nazariyasiga tayanadilar. Turkolog olimlarning izlanishlari natijasi shuni ko‘rsatadiki omonimiyani aniqlash yuqorida keltirilgan metodlar muhim ahamiyatga ega.

Xorijiy tajribalarni chuqur o‘rgangan holda, o‘zbek tilidagi omonim so‘zlarni farqlashni amalga oshirishda qoidalarga asoslangan, statistic ma’lumotlarga asoslangan va mashinali o‘rganishga asoslangan usullardan foydalanamiz. O‘zbek tilidagi omonim so‘zlarni farqlashda ularni so‘z turkumlari doirasida uchrashiga ko‘ra bir so‘z turkumi, ikki so‘z turkumi, uch so‘z turkumi va to‘rtta so‘z turkumi doirasidagi omonimlar kabi guruhlarga ajratdik. Qoidalarga asoslangan usul yordamida grammatik jihatdan o‘xshash bo‘lmagan so‘z turkumlari doirasidagi omonimiyani aniqlashda foydalandik. Bu haqida [Elov B.B, Axmedova X.I., 2022:150-161], [Gulyamova Sh.K., Axmedova X.I., 2021:103-109], [Axmedova X.I., 2022:455-457], [Elov B.B. Axmedova X.I] keltirilgan ilmiy maqolalarda keltirganmiz.

Ushbu maqolada o‘zbek tilidagi omonim so‘zlarni farqlashda statistik usullardan ya’ni chastotali usuldan foydalanish va uning ahamiyatini yoritishga harakat qilamiz.

Statistik ma’lumotlarga asoslangan usul – masalani yechishda so‘zlarning grammatik parametrlarini tasniflash orqali foydalaniladi. Bu parametrlar turli tabiiy tillarda turlicha tanlab olinadi. Masalan rus tilida morfologik omonimiyani aniqlashda so‘z turkumi, so‘zning qaysi jinsga(род) oidligi, birlik yoki ko‘plik shakli, lemma, lemma va so‘z turkumi, faqat lemmasi, omonimligi kabi parametrlar ajratib olingan [Рысаков С.В. Клышинский Э.С., 2015:555-563]. Rus tilidagi SemSin tizimida morfologik omonimiyani olib tashlash (снятие омонимии) masalasi yechilgan. Rus tilida bu termin снятие омонимии deb nomlangan, chunki semantik analizatorning

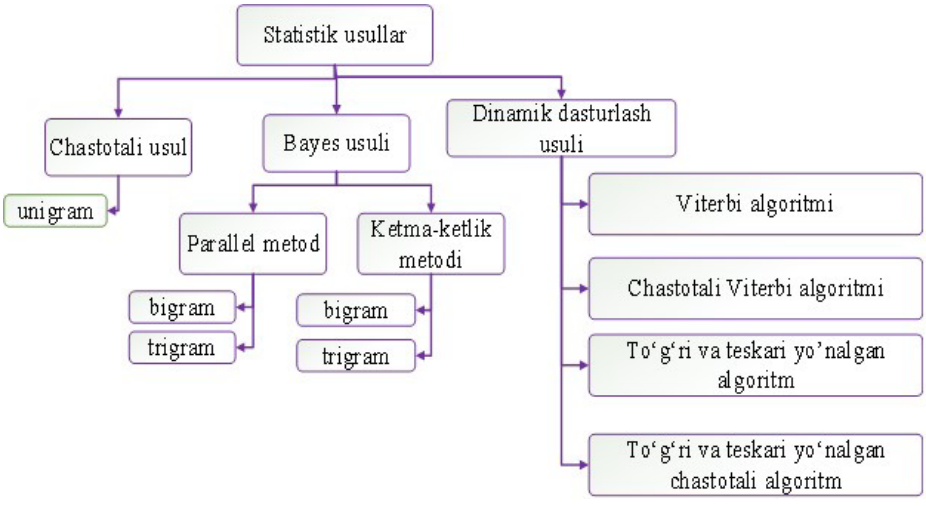
elementlarini aniqlanganini matn ichidan ajratib olib tashlansa, qolgan elementlarni analiz qilish- tizimning ishlash tezligini oshiradi, algoritmlarning aniq va to‘g‘ri ishlashini ta’minlaydi. Tasniflash parametrlarini o‘zbek tili uchun quyidagicha ajratish mumkin.



2-rasm: Tasniflash parametrlari

So‘zning semantik tahlilining aniqlik darajasi yuqori bo‘lmasa, tasniflanish parametrlarini kengaytirish mumkin. Ajratib olingan parametrlar yordamida statistik usullardan foydalanish mumkin. Ushbu usulning asosiy vazifasi kontekst tarkibini n-grammlarga ajratish, ya’ni kontekstdagi kirish so‘zining birikuvchilari aniqlab, baholash metodlari yordamida baholanadi [[Рысаков С.В. Клышинский Э.С., 2015:559].]. Tatar tilida ham boshqa agglutinative turkiy tillar guruhlarida bo‘lgani kabi morfema tilshunoslikning ham semantik, ham sintaktik ma’lumot tashuvchi eng muhim ma’noli birligidir. TuganTel - tatar tili korpusida morfologik polisemiyani aniqlash algoritmi, ma’lumotlar bazasi ishlab chiqilgan. Gataullin Ramil Raisovich o‘zining nomzodlik dissertatsiyasida tatar tilidagi omonimiya va polisemiyani olib tashlash bo‘yicha model va algoritmlar haqida mulohaza yuritgan. Tatar tilida ham omonimiya va polisemiyani farqlashda statistic ma’lumotlarga asoslanga Markov modellaridan foydalanish tavsiya qilingan [2].

Ba’zi statistik usulda so‘z va uning qo‘shimchalari orqali qaror qabul qilinsa, ba’zilarida so‘zning kontekstdagi semantik valentliklari yordamida xulosa qilinadi. Bundan kelib chiqadiki, statistik ma’lumotlarga asoslangan usullar qaror qabul qilish parametrlariga ko‘ra bir necha guruhlarga bo‘linadi.



3-rasm: Statistik usullar iyerarxiyasi

3-rasmda tasvirlanganidek, Chastotali usul unigram, ya’ni so‘zning o‘zi orqali semantik tahlil qilinadi, Bayes usulida esa so‘zning bigram va trigrammalaridan foydalaniladi. Dinamik dasturlash usulida esa n-grammalardan foydalaniladi. Ushbu metodlar orqali baholashda kontekstni tokenizatsiya va teglashtirish jarayonlari olib boriladi. Statistic usullar yordamida turli so‘z turkumlari doirasidagi omonim so‘zlarni ma’nolarini farqlash jarayonini ko‘rib chiqamiz. Masalan, ot yoki sifat so‘z turkumlari orasidagi omonimiyani chastotali usul yordamida aniqlash ketma-ketligini keltiraylik.

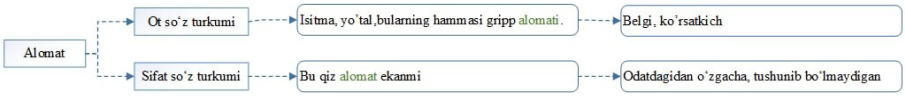
Ot va sifat so‘z turkumi doirasidagi omonim so‘zlar quyidagi para- metrlar bo‘yicha tasniflanadi:

1. So‘z turkumi;
2. O‘zak va lemma;
3. Faqat lemma
4. Faqat o‘zak

Ot va *sifat* so‘z turkumlari doirasidagi omonim so‘z ishtirokidagi gap berilgan bo‘lsin.

*Bu bemordagi **alomatlar** covid-19 kasalligini eslatyapti.*

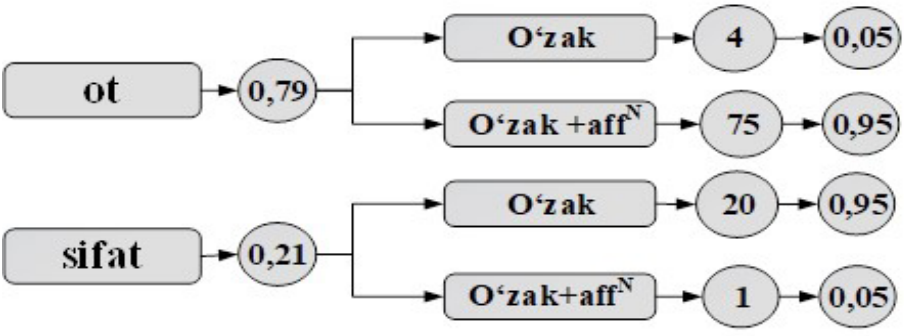
Ushbu gapda *alomat* so‘zi omonim bo‘lib, quyidagi ma’nolarni ang- latadi.



4-rasm: Alomat so’zining ma’nolaridan

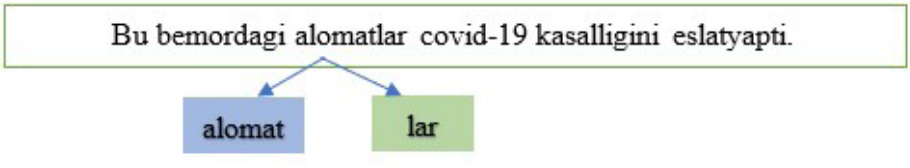
Alomat so‘zini yuqoridagi parametrlar bo‘yicha tasniflashda o‘zbek tili milliy korpusi ma’lumotlaridan foydalanamiz.

Uzschoolcorpora.uz sayti orqali *alomat* so‘zi qidirilganda jami 6823 ta ma’lumot topildi, shulardan 100 tasi tahlil qilinganda 79 tasi ot so‘z turkumi 21 tasi sifat so‘z turkumi tahlil qilingan 100 ta ma’lumot o‘zak va qo‘shimchalarga ajratildi va quyidagi natijalar aniqlandi.



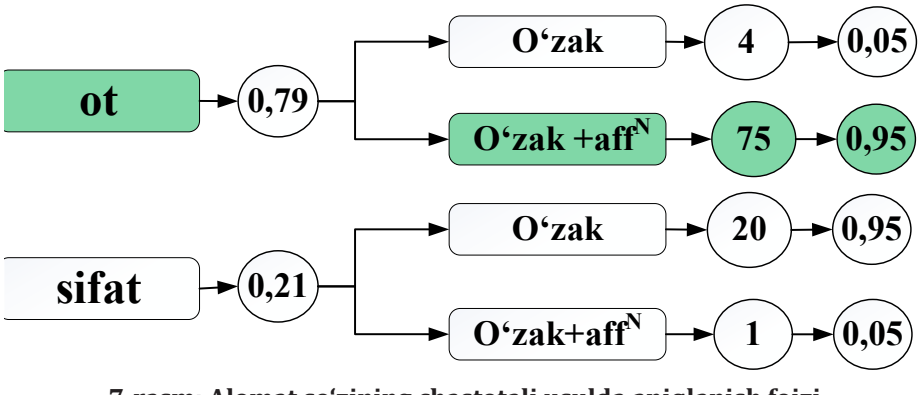
5-rasm: Alomat so‘zining tasniflash parametrlari bo‘yicha statistik ma’lumot

Olingan statistik ma’lumotlarga asosan kiritilgan gapdagi omonim so‘zni qaysi so‘z turkumiga oid ekanligini aniqlaymiz.



6-rasm: o’zak+aff parametriga mos jumla

Demak, bu gapdagi *alomat* so‘zi *o‘zak+aff*^{N} tasniflash parametriga mos ekanligini aniqlagan holda quyidagicha qaror qabul qilinadi.



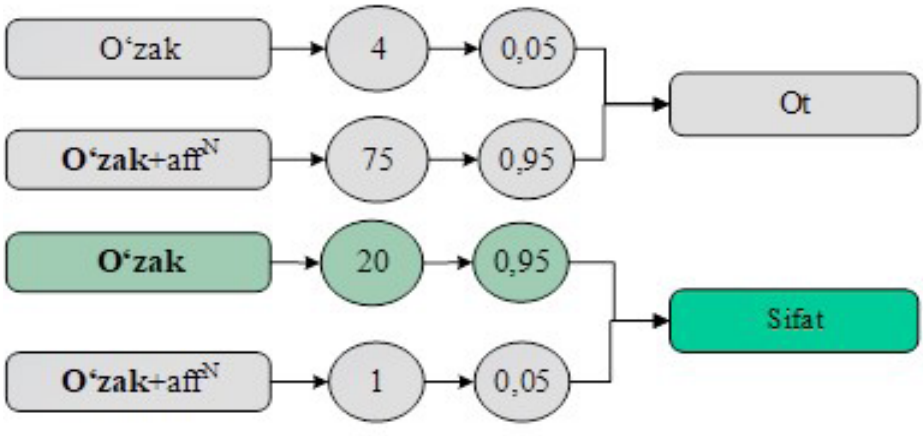
7-rasm: Alomat so‘zining chastotali usulda aniqlanish foizi

Bu chizmadagi ma’lumotlar shuni ko‘rsatadiki, berilgan gapdagi alomat so‘zi 95% ehtimollik bilan ot so‘z turkumidagi omonim so‘z ekanligi va u “belgi, ko‘rsatkich” ma’nosini anglatadi.

Xuddi shunday ikkinchi gapdagi omonim so‘zni ham aniqlaymiz:

8-rasm: o‘zak parametriga mos jumla

Bu gapdagi alomat so‘zi tarkibida qo‘shimcha mavjud emas, demak, o‘zak shaklida uchrash holatining ehtimolligini qaraymiz.



9-rasm: Berilgan jumladagi omonim so‘zning sifat so‘z turkumiga oid ekanlik ehtimolli

Ikkinchi gapdagi alomat so‘zi esa 95% ehtimollik bilan sifat so‘z turkumiga oid va u *odatdagidan o‘zgacha, tushunib bo‘lmaydigan* degan ma’noni bildiradi. Berilgan ikkala gapda uchragan *alomat* so‘zining semantik tahlili statistik usullar yordamida 95% aniqlikka ega. Qolgan 5% semantik aniqlik qoidalarga asoslangan va Mashinali o‘qitishga asoslangan usullar yordamida aniqlanadi. Shu tarzda *chastotali usul* orqali omonimiyani aniqlash mumkin. Ko‘rinib turibdiki, chastotali usuldan foydalanish uchun turli so‘z turkumlari doirasidagi omonim so‘zlarning har biri uchun tasniflash parametrlari aniqlanishi va ular asosida statistik hisob-kitoblar amalga oshirilishi talab qilinadi.

Chastotali usul yordamida omonimiyani aniqlash uchun quyidagicha dataset zarur bo‘ladi.

№	**Tasniflash parametrlari**
1	O‘zak
2	O‘zak+aff

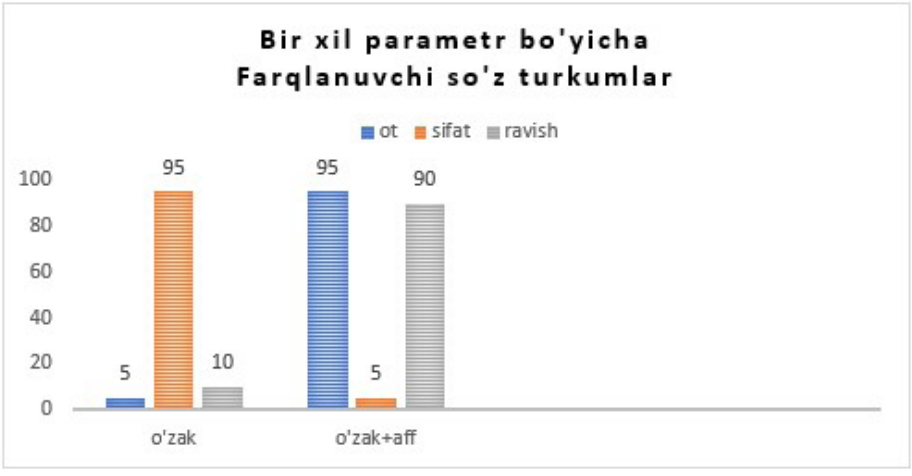
Chastotali usul yordamida farqlanishi mumkin bo‘lgan so‘z turkumlari ro‘yxati.

№	***So‘z turkumlari***
1	Ot
2	Sifat
3	Ravish
4	Olmosh
5	Ko‘makchi
…	…

So‘z turkumlarini farqlovchi tasniflash parametrlaridan iborat jadval

1-so‘z turkumi	2-so‘z turkumi	Tasniflash parametri	Farqlash algoritmi
Ot	Sifat	O‘zak	5%, 95%
Sifat	Ot	O‘zak+aff	95%, 5%
Ot	Ko‘makchi	O‘zak	75%, 25%
Ko‘makchi	Ot	O‘zak+aff	25%,75%
…	…	…	…

Xuddi shunday boshqa so‘z turkumlari orasidagi omonim so‘zlarni aniqlashda ham chastotali usuldan foydalanish mumkin. Olib borilgan izlanishlar natijasida ot,sifat va ravish so‘z turkumlari orasidagi omonim so‘zlar farqlandi. Olingan natijalarni diagramma ko‘rinishida quyidagicha tasvirlash mumkin.



Xulosa

Chastotali usul statistik usullarning unigram modeli bo‘lib, unda faqat shu so‘zning o‘ziga xos bo‘lgan parametrlar aniqlanadi. Masalan joriy omonim so‘z ot so‘z turkumiga oid bo‘lganida qanday xususiyatlarga ega-yu, sifat so‘z turkumiga oid bo‘lganida qanday xususiyatlani o‘zida saqlaydi. Tasniflovchi parametrlar shunday tanlanishi kerakki, ularni hosil qilgan foizi boshqasiga qaraganda sezilarli darajada katta bo‘lsin. Chastotali usul yordamida faqat turli so‘z turkumlari doirasidagi omonimiyani aniqlashda foydalanish mumkin. Bir so‘z turkumiga oid omonim so‘zlarni farqlashda esa Mashinali o‘qitishning boshqa algoritmlaridan foydalanish maqsadga muvofiq. Chunki bir so‘z turkumi doirasidagi omonim so‘zlar bir xil xususiyatga ega. Ularni bu usul bilan farqlash ancha mushkul.

Foydalanilgan adabiyotlar

Рысаков С.В. Клышинский Э.С. Статистические методы снятия омонимии// Новые информационные технологии в автоматизированных системах. – 2015

Chirag Goyal-June 23,2021// https://www.analyticsvidhya.com/blog/2021/06/part-9-step-by-step-guide-to-master-nlp-semantic-analysis/

Elov B.B, Axmedova X.I. Uchta so‘z turkumi doirasidagi omonimiyani farqlovchi biznes jarayonni modellashtirish//O‘zbekiston respublikasi innovatsion rivojlanish vazirligining, ilm-fan va innovasion rivojlanish ilmiy jurnal 2022 / 1, 150-161-b.

Ш.К.Гулямова, Х.И. Ахмедова. Омоним сўзларнинг лингвистик асоси, математик модели ва алгоритмлари (от ва

феъл ҳамда сифат ва феъл туркуми доирасида)// Qo‘qonDPI. Ilmiy xabarlar,2021/1

Axmedova X. Ravish/fe’l va ravish/yuklama so‘z turkumlari orasidagi omonimlikni farqlovchi matematik modellar // “Zamonaviy leksikografiya, til korpuslari va turkiy tillar platformalarini yaratish muammolari” nomdagi sho‘ba materiallari. Elektron nashr / ebook. – Toshkent: ToshDO‘TAU, 18.10.2021. – 452 b.

Axmedova X., Xusainova Z. Fe’l/ko‘makchi va sifat/son so‘z turkumlari orasidagi omonimlikni farqlovchi matematik modellar //“Zamonaviy leksikografiya, til korpuslari va turkiy tillar platformalarini yaratish muammolari” nomdagi sho‘ba materiallari. Elektron nashr / ebook. – Toshkent: ToshDO‘TAU, 18.10.2021. – 452 b.

E. B. Boltayevich and A. X. Ilxomovna, "Business Process Modeling That Distinguishes Homonymy Within Three Parts of Speechs in The Uzbek Language," 2022 7th International Conference on Computer Science and Engineering (UBMK), Diyarbakir, Turkey, 2022, pp. 278-283, doi: 10.1109/UBMK55850.2022.9919453.

Anggraeni, M., Syafrullah, M., & Damanik, H. A. (2019). Literation Hearing Impairment (I-Chat Bot): Natural Language Processing (NLP) and Naïve Bayes Method. In Journal of Physics: Conference Series (Vol. 1201). Institute of Physics Publishing. https://doi.org/10.1088/1742-6596/1201/1/012057

Putong, M. W., & Suharjito. (2020). Classification model of contact center customers emails using machine learning. Advances in Science, Technology and Engineering Systems, 5(1), 174–182. https://doi.org/10.25046/aj050123

Bako, A. T., Taylor, H. L., Wiley, K., Zheng, J., Walter-McCabe, H., Kasthurirathne, S. N., & Vest, J. R. (2021). Using natural language processing to classify social work interventions. American Journal of Managed Care, 27(1), E24–E31. https://doi.org/10.37765/AJMC.2021.88580

Bogery, R., Babtain, N. A., Aslam, N., Alkabour, N., Hashim, Y. A., & Khan, I. U. (2019). Automatic semantic categorization of news headlines using ensemble machine learning: A comparative study. International Journal of Advanced Computer Science and Applications, 10(11), 689–696.

https://doi.org/10.14569/IJACSA.2019.0101190

Nahar, K. M. O., Jaradat, A., Atoum, M. S., & Ibrahim, F. (2020). Sentiment analysis and classification of arab jordanian facebook comments for jordanian telecom companies using lexicon-based approach and machine learning. Jordanian Journal of Computers and Information Technology, 6(3), 247–262. https://doi.org/10.5455/jjcit.71-1586289399

Taheri, S., & Mammadov, M. (2013). Learning the naive bayes classifier with optimization models. International Journal of Applied Mathematics and Computer Science, 23(4), 787–795. https://doi.org/10.2478/amcs-2013-0059

Foster, J., & Wagner, J. (2021). Naive Bayes versus BERT: Jupyter notebook assignments for an introductory NLP course. In Teaching NLP 2021 - Proceedings of the 5th Workshop on Teaching Natural Language Processing (pp. 112–114). Association for Computational Linguistics (ACL). https://doi.org/10.18653/v1/2021.teachingnlp-1.20

Sabita, H., Fitria, F., & Herwanto, R. (2021). ANALISA DAN PREDIKSI IKLAN LOWONGAN KERJA PALSU DENGAN METODE NATURAL LANGUAGE PROGRAMING DAN MACHINE LEARNING. Jurnal Informatika, 21(1), 14–22. https://doi.org/10.30873/ji.v21i1.2865

Sourav Kunal, Arijit Saha, Aman Varma, Vivek Tiwari. Textual Dissection Of Live Twitter Reviews Using Naive Bayes. International Conference on Computational Intelligence and Data Science (ICCIDS 2018). Procedia Computer Science 132 (2018) 307–313.

Granik, M., & Mesyura, V. (2017). Fake news detection using naive Bayes classifier. In 2017 IEEE 1st Ukraine Conference on Electrical and Computer Engineering, UKRCON 2017 - Proceedings (pp. 900–903). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/UKRCON.2017.8100379.

Bahri, S., Saputra, R. A., & Wajhillah, R. (2017). Analisa sentimen berbasis Natural Languange Processing (NLP) dengan Naïve-Bayes clasifier. Konferensi Nasional Ilmu Social & Technology, 1(1), 176–180. Retrieved from https://www.researchgate.net/

Rusli, N. L. I., Amir, A., Zahri, N. A. H., & Ahmad, R. B. (2019). Snake species identification by using natural language processing. Indonesian Journal of Electrical Engineering

and Computer Science, 13(3), 999–1006. https://doi.org/10.11591/ijeecs.v13.i3.pp999-1006

Kaur, C. (2020). Sentiment Analysis of Tweets on Social Issues using Machine Learning Approach. International Journal of Advanced Trends in Computer Science and Engineering, 9(4), 6303–6311. https://doi.org/10.30534/ijatcse/2020/310942020

Siddiqui, S., Rehman, M. A., Daudpota, S. M., & Waqas, A. (2019). Opinion mining: An approach to feature engineering. International Journal of Advanced Computer Science and Applications, 10(3), 159–165. https://doi.org/10.14569/IJACSA.2019.0100320

Chifu, A. G., & Ionescu, R. T. (2012). Word sense disambiguation to improve precision for ambiguous queries. Open Computer Science, 2(4), 398–411.https://doi.org/10.2478/s13537-012-0032-6

Pal, A. R., Saha, D., Naskar, S. K., & Dash, N. S. (2021). In search of a suitable method for disambiguation of word senses in Bengali. International Journal of Speech Technology, 24(2), 439–454. https://doi.org/10.1007/s10772-020-09787-8

Ku, C. H., & Leroy, G. (2014). A decision support system: Automated crime report analysis and classification for e-government. Government Information Quarterly, 31(4), 534–544. https://doi.org/10.1016/j.giq.2014.08.003

Determining homonymy between different word groups using statistical methods

Xolisa Axmedova[1]

Abstract:

The problem of automatic processing of natural language remains relevant for more than half a century. One of the important

[1] *Axmedova Xolisxon Ilxomovna*- PhD student of Tashkent State University of Uzbek Language and Literature named after Alisher Navoi.
E-mail: xolisa0629@gmail.com
ORCID: 0000-0002-9828-1650

problems in the field of NLP is the creation of a semantic analyzer, which in turn goes through a number of steps. Determining homonymy is important in the semantic analysis of sentences. The frequency method is used to determine homonymy between grammatically similar word groups. This method involves extracting homonym classification parameters.

Keywords: *NLP, semantic analyzer, homonomy, statistical methods, parameter classifiers, unigram.*

References

Rysakov S.V. Klyshinskiy E.S. Statisticheskiye metody snyatiya omonimii// Novye informatsionnye texnologii v avtomatizirovannyx sistemax. – 2015

Chirag Goyal-June 23,2021// https://www.analyticsvidhya.com/blog/2021/06/part-9-step-by-step-guide-to-master-nlp-semantic-analysis/

Elov B.B, Axmedova X.I. Uchta so‘z turkumi doirasidagi omonimiyani farqlovchi biznes jarayonni modellashtirish//O‘zbekiston respublikasi innovatsion rivojlanish vazirligining, ilm-fan va innovasion rivojlanish ilmiy jurnal 2022 / 1, 150-161-b.

Ш.К.Гулямова, Х.И. Ахмедова. Омоним сўзларнинг лингвистик асоси, математик модели ва алгоритмлари (от ва феъл ҳамда сифат ва феъл туркуми доирасида)// Qo‘qonDPI. Ilmiy xabarlar,2021/1

Axmedova X. Ravish/fe’l va ravish/yuklama so‘z turkumlari orasidagi omonimlikni farqlovchi matematik modellar // “Zamonaviy leksikografiya, til korpuslari va turkiy tillar platformalarini yaratish muammolari” nomdagi sho‘ba materiallari. Elektron nashr / ebook. – Toshkent: ToshDO‘TAU, 18.10.2021. – 452 b.

Axmedova X., Xusainova Z. Fe’l/ko‘makchi va sifat/son so‘z turkumlari orasidagi omonimlikni farqlovchi matematik modellar //“Zamonaviy leksikografiya, til korpuslari va turkiy tillar platformalarini yaratish muammolari” nomdagi sho‘ba materiallari. Elektron nashr / ebook. – Toshkent: ToshDO‘TAU, 18.10.2021. – 452 b.

E. B. Boltayevich and A. X. Ilxomovna, "Business Process Modeling That Distinguishes Homonymy Within Three Parts of Speechs in The Uzbek Language," 2022 7th International Conference on Computer Science and Engineering (UBMK), Diyarbakir, Turkey, 2022, pp. 278-283, doi: 10.1109/UBMK55850.2022.9919453.

Anggraeni, M., Syafrullah, M., & Damanik, H. A. (2019). Literation Hearing Impairment (I-Chat Bot): Natural Language Processing (NLP) and Naïve Bayes Method. In Journal of Physics: Conference Series (Vol. 1201). Institute of Physics Publishing. https://doi.org/10.1088/1742-6596/1201/1/012057

Putong, M. W., & Suharjito. (2020). Classification model of contact center customers emails using machine learning. Advances in Science, Technology and Engineering Systems, 5(1), 174–182. https://doi.org/10.25046/aj050123

Bako, A. T., Taylor, H. L., Wiley, K., Zheng, J., Walter-McCabe, H., Kasthurirathne, S. N., & Vest, J. R. (2021). Using natural language processing to classify social work interventions. American Journal of Managed Care, 27(1), E24–E31.https://doi.org/10.37765/AJMC.2021.88580

Bogery, R., Babtain, N. A., Aslam, N., Alkabour, N., Hashim, Y. A., & Khan, I. U. (2019). Automatic semantic categorization of news headlines using ensemble machine learning: A comparative study. International Journal of Advanced Computer Science and Applications, 10(11), 689–696. https://doi.org/10.14569/IJACSA.2019.0101190

Nahar, K. M. O., Jaradat, A., Atoum, M. S., & Ibrahim, F. (2020). Sentiment analysis and classification of arab jordanian facebook comments for jordanian telecom companies using lexicon-based approach and machine learning. Jordanian Journal of Computers and Information Technology, 6(3), 247–262. https://doi.org/10.5455/jjcit.71-1586289399

Taheri, S., & Mammadov, M. (2013). Learning the naive bayes classifier with optimization models. International Journal of Applied Mathematics and Computer Science, 23(4), 787–795. https://doi.org/10.2478/amcs-2013-0059

Foster, J., & Wagner, J. (2021). Naive Bayes versus BERT: Jupyter notebook assignments for an introductory NLP course. In Teaching NLP 2021 - Proceedings of the 5th Workshop on Teaching Natural Language Processing (pp. 112–114). Association for Computational Linguistics (ACL). https://doi.org/10.18653/v1/2021.teachingnlp-1.20

Sabita, H., Fitria, F., & Herwanto, R. (2021). ANALISA DAN PREDIKSI IKLAN LOWONGAN KERJA PALSU DENGAN METODE NATURAL LANGUAGE PROGRAMING DAN MACHINE LEARNING. Jurnal Informatika, 21(1), 14–22. https://doi.org/10.30873/ji.v21i1.2865

Sourav Kunal, Arijit Saha, Aman Varma, Vivek Tiwari. Textual Dissection Of Live Twitter Reviews Using Naive Bayes. International Conference on Computational Intelligence and Data Science (ICCIDS 2018). Procedia Computer Science 132 (2018) 307–313.

Granik, M., & Mesyura, V. (2017). Fake news detection using naive Bayes classifier. In 2017 IEEE 1st Ukraine Conference on Electrical and Computer Engineering, UKRCON 2017 - Proceedings (pp. 900–903). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/UKRCON.2017.8100379.

Bahri, S., Saputra, R. A., & Wajhillah, R. (2017). Analisa sentimen berbasis Natural Languange Processing (NLP) dengan Naïve-Bayes clasifier. Konferensi Nasional Ilmu Social & Technology, 1(1), 176–180. Retrieved from https://www.researchgate.net/

Rusli, N. L. I., Amir, A., Zahri, N. A. H., & Ahmad, R. B. (2019). Snake species identification by using natural language processing. Indonesian Journal of Electrical Engineering and Computer Science, 13(3), 999–1006. https://doi.org/10.11591/ijeecs.v13.i3.pp999-1006

Kaur, C. (2020). Sentiment Analysis of Tweets on Social Issues using Machine Learning Approach. International Journal of Advanced Trends in Computer Science and Engineering, 9(4), 6303–6311. https://doi.org/10.30534/ijatcse/2020/310942020

Siddiqui, S., Rehman, M. A., Daudpota, S. M., & Waqas, A. (2019). Opinion mining: An approach to feature engineering. International Journal of Advanced Computer Science and Applications, 10(3), 159–165. https://doi.org/10.14569/IJACSA.2019.0100320

Chifu, A. G., & Ionescu, R. T. (2012). Word sense disambiguation to improve precision for ambiguous queries. Open Computer Science, 2(4), 398–411.https://doi.org/10.2478/s13537-012-0032-6

Pal, A. R., Saha, D., Naskar, S. K., & Dash, N. S. (2021). In search of a suitable method for disambiguation of word senses in Bengali. International Journal of Speech Technology, 24(2), 439–454. https://doi.org/10.1007/s10772-020-09787-8

Ku, C. H., & Leroy, G. (2014). A decision support system: Automated crime report analysis and classification for e-government. Government Information Quarterly, 31(4), 534–544. https://doi.org/10.1016/j.giq.2014.08.003