

LEXICON-BASED VA MACHINE LEARNING-BASED MODELLARI ASOSIDA SENTIMENT TAHLILNI AMALGA OSHIRISH

Elov Botir Boltayevich,

Texnika fanlari bo'yicha falsafa doktori PhD, dotsent

elov@navoiy-uni.uz

ToshDO'TAU

Malika Suyunova Odil qizi,

Tayanch doktorant

malikasuyunova0@gamil.com

ToshDO'TAU

Annotatsiya. Ushbu maqolada sentiment tahlili bo'yicha mavjud yondashuvlar, xususan, leksikonga asoslangan va mashinani o'rganishga asoslangan metodlar tahlil qilinadi hamda ularning samaradorligi o'rganiladi. Sentiment tahlili tabiiy tilni qayta ishlashning muhim yo'nalishlaridan biri bo'lib, u matnlarning hissiy kayfiyatini (ijobiy, salbiy yoki neytral) aniqlash uchun qo'llaniladi. Maqolada leksikonga asoslangan yondashuv so'zlarning oldindan belgilangan lug'atlar va sentiment resurslari (masalan, WordNet) asosida baholanishini tushuntiradi. Ushbu metodda so'zlarning hissiy qiymati oldindan belgilanib, ulardan foydalanib matnning umumiy sentimentini aniqlanadi. Shuningdek, tadqiqotda mashinani o'rganishga asoslangan yondashuv keng yoritilib, unda model katta hajmdagi annotatsiyalangan matnlar asosida o'qitilishi va sentimentni avtomatik ravishda aniqlashi tushuntiriladi. Ushbu usulda NV va SVM kabi algoritmlar qo'llanilib, sentimentni bashorat qilish uchun matn xususiyatlari o'rganiladi. Maqolada har ikki yondashuvning natijalari taqqoslanib, ularning kuchli va zaif tomonlari muhokama qilinadi. Tadqiqot natijalariga ko'ra, agar matnni sentiment tahlil qilish uchun yetarli hajmda ma'lumot mavjud bo'lsa, mashina o'rganishga asoslangan yondashuv aniqroq natijalar bera oladi. Biroq, leksikonga asoslangan yondashuv ko'p resurs talab qilmaydi va turli tillarga oson moslashtirilishi mumkin. Shu sababli, maqolada ushbu ikkala yondashuvni birlashtirgan kombinatsiyalangan yondashuv taklif qilinib, uning sentiment tahlilidagi samaradorligi ko'rsatilgan.

Annotation. This article analyzes existing approaches to sentiment analysis, particularly lexicon-based and machine learning-based methods, and examines their effectiveness. Sentiment analysis is one of the key areas of natural language processing, used to determine the emotional tone of texts (positive, negative, or neutral). The article explains that the lexicon-based approach evaluates words based on predefined dictionaries and sentiment resources (such as WordNet). In this method, the emotional value of words is predetermined and used to determine the overall sentiment of the text. The study also extensively discusses a machine learning-based approach that involves training a model on large volumes of

annotated text to automatically detect sentiment. In this method, algorithms such as Naive Bayes (NB) and Support Vector Machines (SVM) are used to analyze text features for sentiment prediction. The article compares the results of both approaches and discusses their strengths and weaknesses. According to the research findings, if a sufficient amount of data is available for sentiment analysis, the machine learning-based approach can provide more accurate results. However, the lexicon-based approach requires fewer resources and can be easily adapted to different languages. Therefore, the article proposes a combined approach that integrates both methods and demonstrates its effectiveness in sentiment analysis.

Абстрактный. В статье анализируются существующие подходы к анализу настроений, в частности методы, основанные на лексике и машинном обучении, а также изучается их эффективность. Анализ тональности текста является одной из важных областей обработки естественного языка и используется для определения эмоционального настроения (положительного, отрицательного или нейтрального) текстов. В статье объясняется, что подход на основе лексики оценивает слова на основе предопределенных словарей и ресурсов тональности (например, WordNet). В этом методе эмоциональная ценность слов предопределена и используется для определения общего настроения текста. В исследовании также подробно рассматривается подход к машинному обучению, объясняющий, как модель может обучаться на больших объемах аннотированного текста и автоматически определять тональность. Этот метод использует такие алгоритмы, как NB и SVM, для изучения особенностей текста и прогнозирования настроений. В статье сравниваются результаты обоих подходов и обсуждаются их сильные и слабые стороны. Согласно исследованию, если имеется достаточный объем данных для анализа настроений текста, подход на основе машинного обучения может обеспечить более точные результаты. Однако подход, основанный на лексике, не требует больших ресурсов и может быть легко адаптирован к разным языкам. Поэтому в статье предлагается комбинированный подход, объединяющий эти два подхода и демонстрирующий его эффективность в анализе настроений.

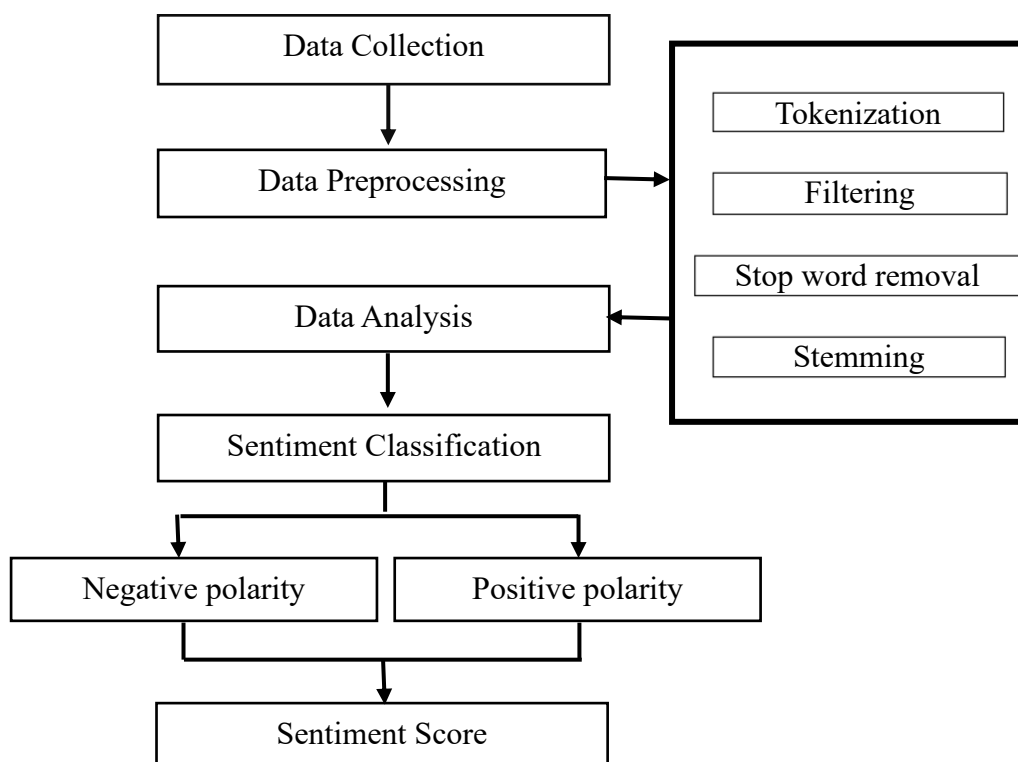
Kalit so‘zlar: *sentiment tahlili, leksikonga asoslangan yondashuv, mashina o‘rganishga asoslangan yondashuv, Naïve Bayes, Support Vector Machine (SVM), sentiment klassifikatsiyasi.*

Kirish

Sentiment tahlili – bu so‘nggi yillarda ko‘plab sohalarda qo‘llanilayotgan muhim texnologiyalardan biri bo‘lib, insonlarning fikrlarini, hissiyotlarini yoki baholarini avtomatik tarzda tahlil qilish imkonini beradi. Ijtimoiy tarmoqlar iste’molchilar va tashkilotlar o‘rtasidagi aloqa vositasi bo‘lib, hozirda keng qo‘llaniluvchi Twitter, Facebook va boshqa tarmoqlar foydalanuvchilarga turli mavzular, shaxslar,

mahsulotlar va xizmatlar haqidagi fikrlarini, xohish-istaklarini ifodalash imkonini bermoqda. Sentiment tahlil esa insonlarning ijtimoiy tarmoqlarda bildirgan ijobiy, salbiy yoki neytral munosabatlarini aniqlash uchun ishlatiladi. Sentiment tahlilni bugungi kunda barcha sohalarida, jumladan, ijtimoiy tarmoqlarda – foydalanuvchilarning fikrlarini tahlil qilishda, marketing sohasida – mahsulot yoki brendga nisbatan mijozlarning munosabatini o'rganishda, siyosatda – jamoatchilik fikrini o'rganishda, servis, xizmat, sog'liqni saqlash sohalarida esa mijozlarning xizmatlar haqidagi fikrlarini tahlil qilishda ishlatamiz. Sentiment tahlil turli ko'rinishda ifodalanishi mumkin. Masalan, ikkilik shaklida (ijobiy va salbiy), uchlik ko'rinishida (ijobiy, salbiy, neytral) yoki baho qiymatida (-5 dan 5 gacha). Ushbu ma'lumotlar turli maqsadlar uchun, jumladan, bozor tadqiqotlari, brendni boshqarish va qaror qabul qilish uchun foydali bo'lishi mumkin.

Sentiment tahlilning jarayon oqimi quyidagicha[1]:



1-rasm. Sentiment tahlil ko'rinishi

Bu jarayon quyidagicha kechadi:

1. **Ma'lumotlarni yig'ish** – turli ijtimoiy tarmoqlardan to'plangan matnli ma'lumotlar to'planadi;
2. **Ma'lumotlarni oldindan qayta ishlash** – bu bosqichda matnni tayyorlash ishlari amalga oshiriladi:
 - Tokenizatsiya – matn so'zlarga ajratiladi;

- Filtrlash – keraksiz elementlar chiqarib tashlanadi;
 - Stop-soʻzlarni olib tashlash – muhim boʻlmagan soʻzlar chiqarib tashlanadi;
 - Stemming – soʻzlar asosiy asos qismga keltiriladi.
3. **Maʼlumotlarni tahlil qilish** – tozalangan matn maʼlumotlari sentiment tahlili uchun tayyorlanadi;
4. **Sentiment klassifikatsiyasi** – matn ijobiy yoki salbiy polaritetga ega ekanligi aniqlanadi:
- Salbiy polaritet;
 - Ijobiy polaritet.
5. **Sentiment bahosi** – bu matnning ijobiy yoki salbiy ekanligi haqidagi yakuniy baho.

Ijtimoiy tarmoqlardagi sentiment matnlari 3 ta asosiy elementdan iborat boʻladi: opinion holder – fikr bildiruvchi, object – obyekt yaʼni fikr bildirilayotgan narsa yoki shaxs, opinion – matn shaklida ifodalangan munosabat [2]. Ushbu elementlarning har biri matnning aniq tahlil qilinishini taʼminlaydi. Shunday qilib, sentiment tahlili nafaqat individual fikrlarni baholash, balki ijtimoiy jarayonlarning umumiy yoʻnalishini tushunish uchun ham kuchli vositadir.

Sentiment tahlilni 3 usul vositasida amalga oshirish mumkin:

1. **Lexicon-based** (leksikaga asoslangan sentiment tahlili);
2. **Machine learning-based** (mashina oʻrganishga asoslangan sentiment tahlili);
3. **Deep learning** (gibrid yondashuv). Ushbu tadqiqot leksikaga asoslangan va mashina oʻrganishga asoslangan yondashuvlarning kuchli va zaif tomonlarini taqqoslaydi.

Adabiyotlar sharhi

N. Imanina Zabha va boshqalar (2019)[3] Twitter maʼlumotlari asosida leksikonga asoslangan usullar yordamida koʻp tilli hissiy tahlilni oʻrgangan. Natijada esa klassifikator hissiyotlarni toʻgʻri aniqlashga muvaffaq boʻlgan. A. Mitra va boshqalar (2020)[4] esa filmlar boʻyicha maʼlumotlar toʻplamida leksikonga asoslangan yondashuvdan foydalangan holda qaysi filmlar ommaga koʻproq yoqishini aniqlashga harakat qilishgan. Bu tadqiqot kino prodyuserlari va teatr egalari uchun foydali hisoblanadi.

M.Sani va boshqalar (2022)[5] Hausa tilidagi tvitlarning hissiy holatini mashinani oʻqitish usullaridan foydalanib tahlil qilishgan. Ushbu tadqiqotda machine learning-based va lexicon-based Hausa matnlarini tasniflashni qanday yaxshilashi mumkinligi oʻrganilgan.

M.S.Hajrahimova, M.I.Ismaylova va boshqalar (2021)[6] COVID-19 pandemiyasi davrida Twitter foydalanuvchilarining munosabatlarini machine learning yordamida tahlil qilishgan. Bu tadqiqotda pandemiya bilan bog'liq tvitlarning hissiyotlari Supporty Vector Machine, Naive Bayes, Random Forest va Neural Network neyron tarmoqlari yordamida baholangan. Bunda Random Forest klassifikatori 79.37% aniqlik bilan eng yaxshi natijani ko'rsatgan.

N.M.Sham, A.Mohamed va boshqalar (2022)[7] iqlim o'zgarishlari bilan bog'liq tvitlarning hissiy tahlilini o'rganishgan. Ular yettita leksikonga asoslangan usul, uchta mashina o'qitish usulini va gibril algoritmlarni solishtirganlar. Eng samarali usul sifatida 75.3 % natija bilan gibril yondashuv F1-Score bo'yicha aniqlangan. Lemmatizatsiya mashina o'qitish va gibril usullarni 1.6% ga oshirgan. Shuningdek, T.Islam, Md.Alif Sheakh va boshqalar (2024)[8] turli tadqiqotlarda ishlatilgan yondashuvlarni taqqoslab, ulardan eng maqbuli sifatida LSTM modelining 98% aniqlik darajasiga ega ekanligini isbotladi.

2-jadval. Qisqa matnlar bilan bog'liq mavjud ishlarga taqqoslama jadvali

Manba	Hissa	Ma'lumotlar to'plami	Algoritmalar	Eng yaxshi aniqlik
Ushbu ish	Twitter sentiment analizida leksikon va chuqur o'rganish yondashuvlarini solishtirish	1,6 million tvitdan iborat Sentiment140 to'plami	Leksikon asosidagi yondashuv va Long Short-Term Memory (LSTM)	98% (LSTM)
[11]	Hausa tilidagi Twitter sentiment analizini mashinani o'rganish va leksikon usullari yordamida o'rganish	BBC Hausa Twitter API	Logistic Regression, Multinomial Naive Bayes (MNB) + Count Vectorizer va TF-IDF	86% (LR)
[12]	Haqiqiy vaqt rejimidagi sentiment analizini SVM, K-SVM va MNB yordamida amalga oshirish	Twitter ma'lumotlari va bahsli hududlar yangilik saytlari	Support Vector Machine (SVM), Kernel SVM (K-SVM), Multinomial Naive Bayes	80% (K-SVM)
[13]	Sentiment analiz uchun leksikon va mashinani o'rganish usullarini solishtirish	TF-IDF	Logistic Regression, Support Vector Machine (SVM), AFINN, Vader lexicon	96.3% (SVM)
[15]	Twitter foydalanuvchilarining COVID-19 bo'yicha munosabatlarini mashinani o'rganish yordamida tahlil qilish	Kaggle'dagi COVID-19 tvitlari	Naive Bayes, Random Forest, SVM, Neural Network	79.37% (Random Forest)
[17]	Hindistondagi COVID-19 va vaksina bo'yicha tvitlarni chuqur o'rganish va leksikon yordamida tahlil qilish	COVID-19 va vaksina tvitlari (Hindiston)	Recurrent Neural Network (RNN) + Bi-LSTM va GRU texnikalari	93.03% (GRU)

Chuqur o'rganish asosidagi usullar, ayniqsa, LSTM-ga asoslangan modellar, matn ketma-ketliklaridagi kontekstual ma'lumotlarni va bog'liqliklarni olishda sezilarli salohiyatni ko'rsatdi[9].

Lexicon-based yondashuvi

Leksikon asosidagi sentiment tahlili usuli, matndagi so'zlarni va ifodalangan his-tuyg'ularni tahlil qilish uchun oldindan tuzilgan lug'atlar yoki leksikonlardan foydalanadi. Bu leksikonlarda har bir so'zga tegishli bo'lgan ijobiy yoki salbiy his-tuyg'u (polaritet) ko'rsatilgan. Bu usulda, matndagi so'zlar yoki iboralar tekshiriladi va ular leksikonda mavjud bo'lgan his-tuyg'u polariteti bilan solishtiriladi[10]. Leksikonga asoslangan yondashuv hissiy tahlil uchun oddiy va tushunarli mexanizmni taqdim etadi. Avvaldan belgilangan leksikondagi so'zlarni hissiy polaritet bilan bog'lash orqali, u matnning hissiyotini ijobiy yoki salbiy iboralar mavjudligiga asoslanib tezda tasniflaydi [11]. Biroq, leksikonlar kontekstga bog'liq kayfiyat, ironiyani yoki jargonni ushlay olmaydi, bu esa xato natijalarga olib kelishi mumkin.

Leksikonga asoslangan yondashuvda hissiyotni bildiruvchi so'z yoki iboralar oldindan lug'at holida shakllantirilgan bo'ladi. Ushbu lug'at orqali sentiment tahlilning ijobiy yoki salbiy ekanligi ishonchli tahlil qilinadi. Matnning hissiy yo'nalishini aniqlash uchun leksikon hujjat bilan taqqoslanadi va undagi ijobiy va salbiy so'zlarning soni aniqlanadi. Keyin esa quyidagi formula yordamida matnning hissiy yo'nalishi belgilanadi [12]:

$$x_1(s) = 1 \text{ if } \text{sum}(1,s) > 0 \quad (1)$$

$$x_1(s) = 0 \text{ if } \text{sum}(1,s) = 0 \quad (2)$$

$$x_1(s) = -1 \text{ if } \text{sum}(1,s) < 0 \quad (3)$$

Bu yerda $\text{sum}(1,s)$ – hujjatda topilgan ijobiy va salbiy hissiyotga ega so'zlarning umumiy farqini ifodalaydi. Agar ijobiy so'zlar soni salbiy so'zlardan ko'p bo'lsa natija $x_1(s) = 1$ (matn ijobiy), agar ijobiy va salbiy so'zlar teng bo'lsa $x_1(s) = 0$ (matn neytral), agar salbiy so'zlar soni ijobiylardan ko'p bo'lsa $x_1(s) = -1$ (matn salbiy) deb qaraladi. Leksikonlarni kengaytirish uchun sinonimlar, WordNet, SenticNet, emojilardan foydalanish mumkin. Sentiment tahlilni amalga oshiruvchi leksikon lug'atlar 3 usul orqali yaratiladi[13,14]:

- a) *qo'lda yondashuv*: bunda hissiyotni ifodalovchi so'zlar shaxsiy bilim va va til tushunchasi orqali qo'lda tuziladi. Bu juda ko'p vaqt talab qiladi va insonning til bo'yicha bilimiga bog'liq bo'ladi. Odatda, ushbu yondashuv

avtomatlashtirilgan usullar bilan birlashtiriladi, chunki bular ba'zi xatolarga yo'l qo'yishi mumkin.

- b) *lug'at asosidagi yondashuv*: hissiyotni bildiruvchi so'zlar turli lug'aviy manbalardan yig'iladi, onlayn lug'at va leksikografik resurslardan foydalanadi, sinonim, antonimlar orqali yangi hissiy so'zlar aniqlanadi[15].
- c) *Korpus asosidagi yondashuv*: bunda so'zlarning gap ichidagi o'rni so'zlar o'rtasidagi munosabati va birga ishlatilish chastotasini tahlil qilish orqali hissiy so'zlarni aniqlaydi[16].

Lexicon-based usulning afzalliklari

- Oddiy: leksikonga asoslangan algoritmlar murakkab machine learning-based usuliga qaraganda kam resurs talab qiladi.
- Tezkor: hissiy yo'nalishni ifodalovchi so'z yoki iboralar oldindan mavjudligi sababli natijalar juda tez chiqadi.
- Yaxshi natija beradi: agar lug'at yetarlicha boy bo'lsa, sentimentni aniqlash natijalari ancha aniq bo'ladi.
- Turli usullar bilan kengaytirish mumkin: lug'at WordNet, SenticNet, emoji va sinonimlar bilan kengaytirilishi mumkin.

Lexicon-based usulining kamchiliklari

- Ironiya va sarkazmni tushunmaydi: masalan, *“Xizmat ko'rsatishga gap bo'lishi mumkinmas, 4 soat kutdim!”* kabi iboralarni noto'g'ri tushunishi mumkin.
- Til bilan bog'liq muammo: har bir til uchun alohida sentiment lug'ati bo'lishi shart.
- Kontekstni tushunmaydi: masalan, *“Havo juda sovuq, bu yomon”*, degan gapning hissiyotini aniq tushunmaydi.

Leksikon asosidagi yondashuv juda oddiy, samarali va tushunarli hisoblanadi. Bunda mashina o'rganish talab qilinmaydi. Lekin bu usulda so'zning kontekstda anglatayotgan ma'nosi turlicha bo'lishi hisobga olinmaydi. Masalan, *yaxshi* va *yaxshi emas* so'zini ijobiy deb tahlil qilishi mumkin. Shuningdek, biror mahsulotga baho berishda qo'shtirnoq ichida teskari ma'no anglatuvchi birikma yoki gaplar ham noto'g'ri tahlil qilinishi mumkin: *“Gap bo'lishi mumkin emas”* – bu, umuman, yaxshi yoki juda yomon ma'nolarida ham ishlatiladi.

Machine learning-based yondashuvi

Bu yondashuv mashina o'rganishga asoslangan usul bo'lib, sentiment tahlilda katta hajmdagi ijtimoiy tarmoq postlari, mahsulot sharhlari va xabarlar

ustida ishlashda samarali hisoblanadi. Bu usul kompyuter tizimlariga aniq dasturlashtirishsiz o'z tajribasi asosida o'rganish va qaror qabul qilish imkonini beradi. Ushbu yondashuv ikki asosiy turga bo'linadi[17]:

1. **Nazoratli o'rganish** – oldindan belgilangan toifalarga ega bo'lgan matn ma'lumotlari asosida modelni o'rgatish usulidir. Uning afzalligi yuqori aniqlik bilan yaxshi ishlash imkoniyatiga ega. Taqdim etilgan sohaga mos ravishda optimallashtirish mumkin. Kamchiligi esa ko'p miqdordagi belgilangan o'quv ma'lumotlari talab qiladi, yangi yoki kontekstga bog'liq so'zlarni tushunmasligi mumkin.
2. **Nazoratsiz o'rganish** – avvaldan belgilangan ma'lumotlarga tayanmagan holda sentimentni aniqlaydi. Afzalligi annotatsiyalangan ma'lumotlarni talab qilmaydi. Yangi yoki noma'lum mustaqil so'zlarning hissiy belgisini mustaqil o'rganishi mumkin. Kamchiligi esa natijalarni oldindan bashorat qilganligi sababli aniqlik kamroq bo'lishi mumkin.

Shuningdek, ba'zi adabiyotlarda *mustahkamlovchi o'rganish* turi haqida ham so'z boradi. Mustahkamlovchi o'rganishda mukofot va jazolar asosida aniq bir qaror qabul qilishni o'rganadi. U tahlil natijasida mukofot yoki jazo olib, o'z strategiyasini yaxshilaydi.

ML yondashuvining afzalliklari:

- Yuqori aniqlikka ega (katta hajmdagi ma'lumotlarga moslashib mustaqil qaror qabul qiladi);
- So'zlarning kontekstda anglatayotgan ma'nosini tushuna olish qobiliyatiga ega;
- Mashina o'rgangan sari chiqargan natijalarini o'zgartirib, sifatini oshirib boradi.

ML yondashuvining kamchiliklari:

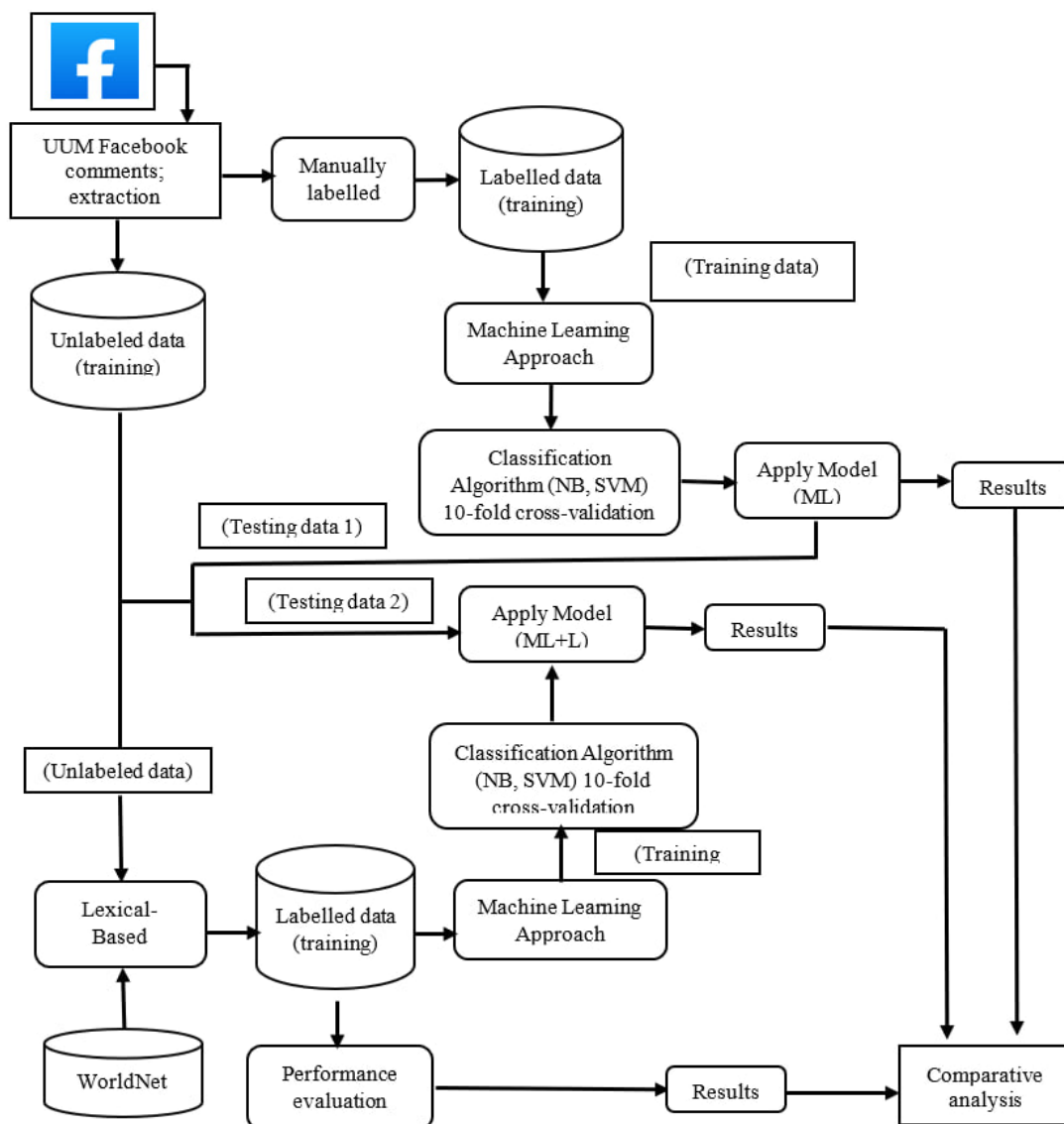
- Katta hajmdagi yorliqlangan ma'lumotlar talab qiladi;
- Tahlil qilish jarayoni ko'p vaqt talab qilishi mumkin;
- Ma'lumotlar tili yoki hissiy bahosi o'zgarganda model moslashishi qiyin bo'lishi mumkin.

Mashina o'rganish usuli bugungi kunda sun'iy intellekt, tibbiyot, moliya, transport, marketing kabi ko'plab sohalarda qo'llanilmoqda.

Chuqur o'rganishga asoslangan yondashuv (gibrid)

Islam T., Sheakh Md.A., Sadik Md.R., Tahosin Mst.S., Foysal Md.M.R., Ferdush J. va M. Begumlar leksikonga asoslangan va chuqur o'rganishga asoslangan

yondashuvlarni taqqoslashgan (2024)[8]. Ushbu tadqiqotda ikki mashhur yondashuv – leksikon asosidagi va chuqur o'rganish asosidagi yondashuvlarni 1,5 milliondan ortiq ma'lumotga ega **sentiment140** deb nomlangan Twitter datasetidan foydalanib taqqoslashgan. Asosiy e'tiborni chuqur o'rganishning ketma-ketlik modeli – *Long Short-Term Memory (LSTM)* ga qaratadilar. Ushbu tadqiqot natijasida leksikon asosidagi yondashuv 95% aniqlikka, chuqur o'rganish algoritmlari model 98% aniqlikka ega ekanligini aniqlashgan. Shuningdek, Alaa Thamer Mahmood, Siti Sakira Kamaruddin va boshqalar (2020)[18] leksikonga asoslangan yondashuv va mashina o'rganishga asoslangan yondashuvni birlashtirib Facebook izohlarida sentiment tahlilni amalga oshirdilar.



3-rasm. Gibrid yondashuv asosidagi sentiment tahlil

Bunda UUM Facebook sahifasidan izohlar olinadi.

1. Mashina o'rganish yondashuvi:

- izohlarga ijobiy yoki salbiy yorliqlari qo'shiladi;
- yorliqlangan ma'lumotlar mashina o'rganish modeli uchun o'quv ma'lumoti sifatida ishlatiladi;
- NB va SVM algoritmlari yordamida 10 marta o'qitiladi;
- 1-natija tayyor bo'ladi.

2. Leksikonga asoslangan yondashuv:

- Facebook izohlari WordNet lug'atidan foydalanib baholanadi;
- lug'at asosida har bir so'z uchun sentiment aniqlanadi;
- bu ma'lumotlar yorliqlangan shaklda saqlanadi va mashina o'rganish model uchun ma'lumot sifatida ishlatiladi;
- mashina o'rganish yondashuvi qayta ishlanib 10 marta qayta tekshirilib o'qitiladi;
- ushbu yondashuv bo'yicha 2-natija hosil bo'ladi.

3. Gibrid yondashuv:

- leksikonga asoslangan yondashuvdan olingan natijalar mashina o'rganish modeli bilan birlashtiriladi;
- birlashtirilgach 3-natija hosil bo'ladi.

4. Natijalarni taqqoslash:

- 1-natija, 2-natija va 3-natija o'zaro taqqoslanib, qaysi yondashuv eng aniq sentiment tahlilini amalga oshirgani baholanadi [19].

Xulosa

Bugungi kunda ijtimoiy tarmoqlarda insonlar tomonidan bildirilayotgan fikrlar va sharhlar katta hajmli bo'lib, bu ma'lumotlarni inson omili vositasida hissiy tahlil qilish juda ko'p vaqt va resurs talab qiladi. Ushbu ma'lumotlarning sentimentini aniqlash uchun sun'iy intellekt va avtomatlashtirilgan tahlil usullari tobora keng qo'llanilmoqda. Ayniqsa, o'zbek tilida ham sentiment tahlili bo'yicha katta ishlar amalga oshirilmoqda. S.Allanazarova, D.Elova (2024)[20] sentiment tahlil uchun matndagi valentlik, hissiyot va boshqa ta'sirchan holatlarni aniqlaganlar. Ilyos Rabbimov (2021)[21] boshchiligidagi bir guruh izlanuvchilar o'zbek filmlariga qoldirilgan sharhlarni emojilar asosida o'rganishgan. Ushbu tadqiqotda esa sentiment tahlilni amalga oshirish uchun eng mashhur ikki yondashuv tahlil qilindi. Har ikkala usul afzallik va kamchiliklarga ega. Faqat bitta usulga tayanish sentiment tahlil uchun yetarli emas, chunki har bir yondashuvning cheklovlari mavjud. Leksikonga asoslangan yondashuvni mashina o'rganish modeli bilan birlashtirish sentimentlarni yanada aniqroq tahlil qilish imkonini beradi.

Foydalanilgan adabiyotlar:

1. Verma, B., & Thakur, R. S. (2018). Sentiment analysis using lexicon and machine learning-based approaches: A survey. In *Proceedings of International Conference on Recent Advancement on Computer and Communication: ICRAC 2017* (pp. 441-447). Springer Singapore.
2. Verma, B., & Thakur, R. S. (2018). Sentiment analysis using lexicon and machine learning-based approaches: A survey. In *Proceedings of International Conference on Recent Advancement on Computer and Communication: ICRAC 2017* (pp. 441-447). Springer Singapore.
3. Zabha, N. I., Ayop, Z., Anawar, S., Hamid, E., & Abidin, Z. Z. (2019). Developing cross-lingual sentiment analysis of Malay Twitter data using lexicon-based approach. *International Journal of Advanced Computer Science and Applications*, 10(1).
4. Mitra, A., & Mohanty, S. (2020). Sentiment analysis using machine learning approaches. *Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS*, 2, 63-68.
5. Sani, M., Ahmad, A., & Abdulazeez, H. S. (2022). Sentiment analysis of Hausa language tweet using machine learning approach. *Journal of Research in Applied Mathematics*, 8(9), 07-16.
6. Hajrahimova, M. S., & Ismaylova, M. I. (2021). Machine learning-based sentiment analysis of Twitter data.
7. Sham, N. M., & Mohamed, A. (2022). Climate change sentiment analysis using lexicon, machine learning and hybrid approaches. *Sustainability*, 14(8), 4723.
8. Islam, T., Sheakh, M. A., Sadik, M. R., Tahosin, M. S., Foysal, M. M. R., Ferdush, J., & Begum, M. (2024). Lexicon and deep learning-based approaches in sentiment analysis on short texts. *Journal of Computer and Communications*, 12(1), 11-34.
9. Bhowmik, N. R., Arifuzzaman, M., & Mondal, M. R. H. (2022). Sentiment analysis on Bangla text using extended lexicon dictionary and deep learning algorithms. *Array*, 13, 100123.
10. Catelli, R., Pelosi, S., & Esposito, M. (2022). Lexicon-based vs. Bert-based sentiment analysis: A comparative study in Italian. *Electronics*, 11(3), 374.
11. Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, 226, 107134.
12. Hota, H. S., Dinesh K. Sharma, and Nilesh Verma. "Lexicon-based sentiment analysis using Twitter data: a case of COVID-19 outbreak in India and abroad." *Data science for COVID-19*. Academic Press, 2021. 275-295.

13. Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams engineering journal*, 5(4), 1093-1113.
14. Khan, A., Baharudin, B., Lee, L. H., & Khan, K. (2010). A review of machine learning algorithms for text-documents classification. *Journal of advances in information technology*, 1(1), 4-20.
15. Hardeniya, T., & Borikar, D. A. (2016). Dictionary based approach to sentiment analysis-a review. *International Journal of Advanced Engineering, Management and Science*, 2(5), 239438.
16. Ye, Q., Zhang, Z., & Law, R. (2009). Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *Expert systems with applications*, 36(3), 6527-6535.
17. Zhang, H., Gan, W., & Jiang, B. (2014, September). Machine learning and lexicon based methods for sentiment classification: A survey. In *2014 11th web information system and application conference* (pp. 262-265). IEEE.
18. Mahmood, A. T., Kamaruddin, S. S., Naser, R. K., & Nadzir, M. M. (2020). A combination of lexicon and machine learning approaches for sentiment analysis on Facebook. *Journal of System and Management Sciences*.
19. Surroop, K., Canoo, K., & Pudaruth, S. (2016). A Novel Position-based Sentiment Classification Algorithm for Facebook Comments. *International Journal of Advanced Computer Science and Applications*, 7(10).
20. Allanazarova, S. (2024). SENTIMENT TAHLIL UCHUN MATNDAN VALENTLIK, HISSIYOT VA BOSHQA TA'SIRCHAN HOLATLARNI ANIQLASH. *Uzbekistan language and culture*, 2(02).
21. Allanazarova, S. (2021). Jamoatchilik fikrini organish yoki sentiment tahlil. *Computer linguistics: problems, solutions, prospects*, 1(1).
22. B.Elov, Sh.Hamroyeva, O.Abdullayeva, M.Uzoqova. O'zbek tilida pos tegging masalasi: muammo va takliflar. O'zbekiston: til va madaniyat (Amaliy filologiya), 2022, 5(4). 45-63 b.