



O'ZMU XABARLARI

ВЕСТНИК НУУЗ

АКТА NUUZ

MIRZO ULUG'BEK NOMIDAGI O'ZBEKISTON MILLIY
UNIVERSITETI ILMIY JURNALI

JURNAL
1997 YILDAN
CHIQA
BOSHLAGAN

2024
1/4/1

Ijtimoiy-
gumanitar
fanlar turkumi

Bosh muharrir:

MADJIDOV I.U. – t.f.d., professor.

Bosh muharrir o'rinbosari:

ERGASHOV Y.S. – f-m.f.d., professor.

SHIRINOVA R.X. – fil.f.d., professor.

Tahrir hay'ati:

Sagdullayev A.S. – t.f.d., akademik.

Ashirov A.A. – t.f.d., prof.

Balliyeva R. – t.f.d., prof.

Malikov A.M. – t.f.d., prof.

Yusupova D.Y. – t.f.d., prof.

Yunusova X.E. – t.f.d., prof.

Murtazayeva R.H. – t.f.d., prof.

Mo'minov A.G. – s.f.d., prof.

Abdulayeva N.B. – f.f.d., prof.

Madayeva Sh.O. – f.f.d., prof.

Tuychiyev B.T. – f.f.d., prof.

Utamuradov A. – f.f.n., prof.

Muxammedova D.G. – psix.f.d., prof.

Boltaboyev H. – fil.f.d., prof.

Rahmonov N.A. – fil.f.d., prof.

Siddiqova I.A. – fil.f.d., prof.

Sa'dullayeva N.A. – fil.f.d., dots.

Arustamyan Y.Y. – fil.f.d., dots.

Pardayev Z.A. – fil.f.f.d., PhD.

Mas'ul kotib: **PARDAYEV Z.A.**

TOSHKENT – 2024

MUNDARIJA

Tarix

Abbasov B. XX asr o'rtalarida Farg'ona vodiysida ko'chirilgan aholining yangi sharoitga moslashishi.....	4
Акимниязова Г. Отражение торговых практик и купечества в Каракалпакском фольклоре	7
Ahmedova N. Muhammad Yusuf Bayoniyning Xiva adabiy muhitida tutgan o'rni	10
Djumanov S. Mustaqillik yillarida Imom Moturidiy merosining o'rganilishi va uning diniy qadriyatlarimizni tiklashdagi o'rni. 13	13
Jalilov D. Sopolli madaniyatiga dasht aholisi madaniy ta'siri: tarixshunoslik masalalari (1970-1999-yillar ilmiy ishlari misolida)	16
Qodirov N. Siyosiy partiyalar "Yoshlar qanotlari"ning yoshlar ijtimoiy faolligiga ta'siri	20
Qurbonov B. O'zbekistonda haj ziyoratining qayta tiklanishining tarixiy bosqichlari	24
Ochilova O. O'rta Osiyo xonliklaridagi mashhur tabiblar faoliyati	27
Сабитова К. Проблема неудовлетворительной подготовки педагогических техникумов в Узбекистане в период 1932-1936 гг. и пути ее решения	30
Tilavova M. Rossiya imperiyasi va Buxoro amirligi o'rtasida ziddiyatlarning kuchayishi	34
Toshturov Sh. Mingyillik rivojlanish maqsadlari doirasida O'zbekistonda ekologik barqarorlikni ta'minlash masalalari.....	37
Usarov U. XIX asrning ikkinchi yarmi – XX asr boshlarida Toshkentning sug'orish ishlariga oid ba'zi mulohazalar.....	40
Ergashev U. Temir-qapig' – Sug'd va Tohariston orasidagi tarixiy tog' guzari	43

Falsafa. Pedagogika. Psixologiya. Metodika. Sotsiologiya. Siyosiy fanlar. Islomshunoslik

Abdullaeva S., Yuldasheva M. Dual ta'lim – afzalliklar va muammolar.....	46
Abdullayeva F., Umarova Z., Rashidov A. Zamonaviy moda tendensiyalarining soha doirasidagi ta'lim jarayonlari modernizatsiyasidagi aktualligi	49
Abduraimov I. Integrativ yondashuv asosida tarix fani o'qituvchilarini kasbiy rivojlanish metodikasini takomillashtirish	52
Akramov K. Kotte tipidagi va jamoat binolarini loyihalash estetikasi.....	55
Amanulla Z. Teaching methodics of Russian in China and in other countries teaching written translation competence of chinese and russian languages (for example of the China)	58
Anorboyev Q. Tibbiyot oliy ta'lim muassasalarida biofarmatsiya fanini o'qitishda talabalarninig integrativ kasbiy kompetentligini rivojlantirish	61
Arapbayeva D. Shaxs milliy g'urning folklor san'atida namoyon bo'lishi	64
Arapbayeva D. Ajrashish ijtimoiy-psixologik hodisa sifatida	67
Arushanova M. Yangi O'zbekistonda turizmni rivojlanishida Islom Karimov muzeyining o'rni	70
Atajonova A. Bo'lajak boshlang'ich sinf o'qituvchilarining mediasavodxonligini rivojlantirish	73
Axmatova G. Gender yondashuv asosida ayol va erkakning jamiyatdagi o'rni.....	76
Axmedova D. Yangi O'zbekiston yoshlari-kelajagimiz poydevori	79
Axmedova M. Noto'liq oilaning ijtimoiy-psixologik xususiyatlari.....	82
Ahmadaliyev S. Ta'lim tizimida tarbiyaviy ishlarni tashkil etishning pedagogik zarurati.....	85
Bahronov B. Jadid ma'rifatchiligining ijtimoiy-falsafiy rivojida Abdulla Avloniyning xizmatlari	88
Bekkiyeva N. O'smirlarda diniy tushunchalarning shakllanishida huquqiy madaniyatni psixologik xususiyatlari.....	91
Berdikulova S. Jamiyat milliy taraqqiyotida ijtimoiy normalarning ijtimoiy-falsafiy jihatlarini	94
Bobanov X. Ontogenezda psixologik madaniyat rivojlanishining ijtimoiy-psixologik omillari.....	97
Burkhonov Sh. Martin Heidegger's profound inquiry into the moral and practical dimensions of pure existence.....	100
Yoqubjonov Sh. Ommaviy axborot vositasi – siyosiy kommunikatsiyaning asosiy faktori sifatida.....	103
Ikramov I. Jamiyatda milliy ijtimoiy xotiraning shakllanish mexanizmlari	106
Irisova S., Suyarov Z., Mamatov B. Maktabgacha yoshdagi bolalarga steam xalqaro dasturini qo'llash metodikasi	109
Ismanova A. Ota-onalarning pedagogik madaniyatini oshirish-yoshlar tarbiyasining zamonaviy muhim usuli sifatida	112
Kadirova F. Development of reading abilities.....	115
Qarshiyev N. Markaziy Osiyoda davlatlarida siyosiy partiyalarning genezisi va ularning mafkuraviy yo'nalishdagi faoliyatining shakllanishi	118
Latibjonov R. Pedagoglarning kasbiy faoliyatida kommunikativ ko'nikmalarni rivojlantirishning ahamiyati	122
Mamatkarimova B. Texnik yo'nalishdagi talabalarning professional leksikasini takomillashtirishda integrativ yondashuvdan foydalanish asoslari.....	125
Mardankulov J. Matematika ta'limi natijalarini baholash va diagnostika qilish uchun test usuli va vositalardan foydalanish.....	128
Maxmudova A. Zamonaviy jamiyatda bioetikaning asosiy muammolari	131
Muxammadiyeva X. Bo'lajak o'qituvchilarda konfliktologik kompetentlikni rivojlantirishda pedagogik-psixologik fanlarning imkoniyatlari.....	134
Nimatova M. Kredit-modul tizimida maktabgacha ta'lim yo'nalishlarida mustaqil topshiriqlar tuzish usullari.....	136
Nishonova N. Boshqaruvda xotin-qizlarning qobiliyati: mulohaza va takliflar.....	139
Otaqulov Sh. Gender tenglikni ta'minlashning asosiy tamoyillari.....	142
Payzibayeva M. Boshlang'ich sinflarda tarbiya fani asosida o'quvchilarni ma'naviy dunyoqarashini shakllantirish usullari	145
Pardayeva K. Raqamli ta'lim va innovatsion texnologiyalar sharoitida maktabgacha yoshdagi bolalar nutqini o'stirish metodikasi	148
Rahimova G. Oilaviy hayot dinamikasida shaxs inqirozlarining oila mustahkamligiga psixologik ta'siri	151
Reymbaeva R., Ernazarova R., Daryabaeva A. Talaba-yoshlarning ijtimoiy faoliyati hamda vatan oldidagi mas'uliyati	154
Saidova M. Xarizmatik shaxslarning sotsial qiyofasi.....	157
Saparbboyev K. O'zbekistonda xavfsizlikni ta'minlashning konstitutsiyaviy asoslari.....	160
Saparov B. Yoshlarda vatanparvarlik tuyg'usini shakllantirish omillari.....	163
Sobirov D. Maqsadli ta'lim asosida talabalarning mustaqil ishlarini takomillashtirish metodikasi.....	166
Sodirjonov M. Zamonaviy o'zbek jamiyatida ijtimoiy transformatsiyaning etnosotsiologik omillari	170
Soyipov S. "Ravzat ush-shuhado" dostonida she'riy san'atlarning qo'llanilishi	173
Soliyev U. Ingliz tilida sohaviy matnlarni o'rgatishda yuzaga keladigan qiyinchiliklarni bartaraf etish usullari	176
Тахирова М. Проектирование целей обучения	179
Tilabjonova S. Pedagogik amaliyotni o'tkazishning zamonaviy usullari.....	182
To'ymanova I. Talabalarda til o'rgatish jarayonini "Pen pal" asosida loyihalash metodikasi	185

Xasanova S. Maktabgacha yoshdagi bolalarda xavotirlanishni psixodiagnostik tahlili.....	188
Xoldorova M. Elektron ta'lim vositalari asosida bo'lajak muhandislarning kasbiy kompetentligini rivojlantirish	191
Xoliqova R. Ona tili ta'limida o'quv topshiriqlarini lingvodidaktik kompetensiyalar asosida rivojlantirish	195
Khujaakhmatova Q., Muradova S., Habibulloyeva G. Using innovative technologies in teaching FL to esp learners	198
Xuramova F. Mobil texnologiyalardan foydalanib ta'lim jarayonini rivojlantirishning didaktik imkoniyatlari	200
Xayitmuradova S. Talaba-yoshlarni tafakurini shakllantirishda pedagogik texnologiyalarning ahamiyati	203
Xakimova I. Sog'lom shaxs hayotida psixologik xizmatning o'rni	206
Halimov Sh. Insoniyat madaniy faoliyati integratsiyasining falsafiy tahlili	209
Hoshimxonov M. Mashrabning taxalluslari va hasbi holi	211
Shamsutdinova N. Ibn sinoning "Risola ut-tayr" asarida tasavvuf mohiyati	214
Shermuxeimedov A., Shermuxeimedova N. Sport takomillashuvi guruh futbolchilarining o'quv-mashg'ulot yuklama parametrlarining musobaqa yuklama parametrlariga nisbatan mutanosibligini tadqiq etish	217
Ergasheva S. Madaniyat san'at sohasi menejmenti fanini o'qitish metodikasi	220
Eshbekov D. Davlat boshqaruvida parlament institutining o'rni va ahamiyati	222
Yuldashev F. Philosophical aspects of studying youth needs and interests in social life	225
Yusubov J. O'zbekistonda davlat fuqarolik xizmati sohasida kadrlar tayyorlash tizimining takomillashtirilishi	228
Filologiya	
Abduqayumova S. Semantic features of english idioms with flower component.....	231
Abdurakhmonov M., Vokhidov A. Ergonyms of educational and cultural spheres in Shakhrisabz.....	234
Авлакулов Я. Проблемы исследования ономастического пространства узбекского языка.....	236
Akbayeva X. Chet tili darslarida innovatsion texnologiyalardan foydalanish	239
Alihonova M. Diachronic approach towards food technology terminology	242
Asatjanova I. Ahmad Tabibiy g'azallarining vazn xususiyatlari	245
Ahmedov P. Художественное освещение проблемы технофобии в сборнике рассказов «Я, робот» Айзека Азимова	248
Achilova M. Shaxslararo munosabatlarda empatiya hissi ahamiyatining ilmiy adabiyotlarda o'rganilganlik holati.....	251
Baxadirova M. The gender factor in women's and men's linguistic worldviews	255
Berdiyeva N. Ingliz va o'zbek adabiyotida funktsional chegaralangan so'zlarning qo'llanishi va tahlili	258
Berdimurodov A. Nemis va o'zbek tillarida inkor gaplarning yasallishi va inkorni ifodalovchi so'zlarning semantik xususiyatlari	261
Biksalieva R. Productive word-forming affixes in the russian and uzbek languages as a means of reflecting socio-cultural realities.....	264
Bobokalonov O. Ziddiyat kontentli leksemalar va antonimlar paradigmasidagi farqli jihatlar	267
Davletov K. Fransuz va o'zbek tillarida qo'rquv konseptiga kiruvchi sifatlar tasnifi.....	270
Davronova Z., Azamjonova S. Peculiarities of national and cultural realities in uzbek works translation into french language (on the example of Chulpan's novels "Day and night" and Aibek's "Navoi").....	274
Djafarova D. Tilshunoslikda "Tarixiy xotira"ga oid frazeologik birliklarning fransuz va o'zbek tillarida tahlili.....	277
Jo'raev J. "Gulshan ul -asror" qo'lyozma nusxalarida Amir Temur va unga chizilgan miniatyuralar talqini.....	280
Jurayeva M. Basic principles of studying language phenomena in cognitive linguistics	284
Zoxitova N. Internet nashrlarning migratsiya mavzusini yoritishda til xususiyatlari	287
Ibragimov A. Orzu Urganjiyning Fuzuliyga taxmislari	291
Kamolova S. O'zbek antonimlarining badiiy matnlarda lug'aviy ma'nolarining ifodalanishi	294
Kayumova M., Ismatullayeva N. Haldun Taner ijodida dramatik janrning o'rni	297
Qudratova S. Hikoya janrining asosiy belgilari va tabiati	300
Matkarimova A., Yuldasheva D. Muloqotda shaxsiy munosabat kategoriyasi	303
Mirsanov G., Qarshiyeva Sh. Ingliz tili ergativ fe'llarining o'rganilish masalalari	306
Misirov S. Ingliz tilida pedagogika va psixologiya yo'nalishi talabalarining kasbiy kompetensiyasini rivojlantirishda sohaviy matnlarni o'qishni o'rgatish metodikasi.....	309
Muxtorova U. Chingiz Aytmatov asarlarida inson va qismat tushunchasi.....	312
Nabiyeva N. O'zbek ayollar she'riyatida hijron motivi.....	315
Navruzova N. So'zning semantik tuzilishida konnotativ ma'noning ahamiyati.....	318
Namazova M. O'tkir Hoshimov asarlarida inson ruhiyati tasviri.....	321
Namozova F. Lingvosensorika – zamonaviy o'zbek tilshunosligining istiqbolli yo'nalishi sifatida.....	324
Norova R. O'zbek va ingliz adabiy ertaklarida she'rli vositalarni qo'llashda milliy folklorning o'rni.....	327
Ortikova M. Today, uzbek and english are mandatory units of language expressing relationships.....	330
Primov A. Badiiy adabiyot va folklorizm (Abdulla Qodiriy nasri misolida).....	333
Pulatova S. A window into Nabokov's mind through his novel "Ada"	336
Saydalieva N. Formation of written translation competence during teaching Chinese in Uzbekistan	339
Safoyeva S. Well so'zining diskursiv-pragmatik ma'nosi. Kino diskursida tarkibiy zanjir sifatida	342
Siddikova I. Stomatologik terminlarning shakllanishida intralingvistik va ekstralingvistik omillar	345
Tojiyeva K. Tilshunoslikda so'z va terminning o'zaro munosabati	349
Tuychiyev S. Gapning sintaktik tahliliga yangicha yondashuvlar haqida ayrim mulohazalar.....	353
O'rmonova N. Qur'oni Karim tarjimai muammolari	356
Usmonova M. Ingliz tilida "Boylarning paydo bo'lishi va qashshoqlik tasviri" tushunchalarining leksik va semantik xususiyatlari	359
Xotamova P. Ingliz va o'zbek tillarida sof substansial possessiv sintaksemaning ifodalanishi	362
Shermamatova S. Adabiyotshunoslikda motiv tushunchasi va uning genezisi	365
Shomurotova Ch. Isajon Sultonning "Alisher Navoiy" asari syujetida plastik tasvir xususiyatlari.....	368
Shukurova S. Examining suicidal motifs in western and eastern literature: a cross-cultural comparative analysis.....	372
Elov B., Alayev N., Yodgorov U. Tematik modellashirishning zamonaviy usullari	375
Yuldasheva D. E.Sapir's contribution to the development of linguistic typology.....	379
Юлдашева Ф., Тепирова Ш. Эволюция языка: исторические и современные тенденции	382
Yuldosheva N. Gap qo'llari voqelanishida etnomadaniy omillar ta'siri xususida	385



Botir ELOV,

Toshkent davlat o'zbek tili va adabiyoti universiteti dotsenti

E-mail: elov@navoiy-uni.uz

ORCID: 0000-0001-5032-6648

Narzullo ALAYEV,

Toshkent davlat o'zbek tili va adabiyoti universiteti tayanch doktoranti

ORCID: 0009-0009-4625-5539

Umidjon YODGOROV,

Toshkent davlat o'zbek tili va adabiyoti universiteti o'qituvchisi

E-mail: yodgorov@navoiy-uni.uz

ORCID: 0000-0002-7666-8395

O'zMU professori N.Abduraxmonova taqrizi asosida

ТЕМАТИК МОДЕЛЛАСHTIRISHNING ZAMONAVIY USULLARI

Аннотация

Тематик моделлаштириш – tabiiy tilni qayta ishlash (Natural Language Processing, NLP) vazifalarining muhim komponenti bo'lib, turli sohalarida ishlatilishi mumkin. Tematik modellarida ikkita asosiy muammo mavjud: natijani bashorat qilish uchun ishlatilishi mumkin bo'lgan mavzularni aniqlash va allaqachon topilgan mavzularni tushunishni osonlashtirish. Ushbu maqolada, tematik modellashtirishning zamonaviy usullari hisoblangan manfiy bo'lmagan matritsa faktorizatsiyasi (Non-negative Matrix Factorization), yashirin semantik taqsimot (Latent Semantic Allocation, LSA), ehtimoliy yashirin semantik tahlil (Probabilistic Latent Semantic Analysis, PLSA), lda2vec chuqur o'rganish modeli va tBERT haqida fikr-mulohaza yuritiladi. Tematik modellashtirish odatda katta hajmdagi til korpusiga qo'llaniladi. Tematik modellashtirish uchta turdagi so'zlarning klasterlarini hosil qiladi – birgalikda keladigan so'zlar; so'zlarning taqsimlanishi va mavzu bo'yicha so'zlarning gistogrammasi. Ushbu maqolada o'zbek tili korpusidagi namunaviy matnlarga LDA usulini qo'llash orqali olingan natijalar keltiriladi.

Kalit so'zlar: Tematik modellashtirish, LDA, LSA, PLSA, lda2vec, tBERT, nazoratsiz mashinali o'qitish usullari, til korpusi.

MODERN METHODS OF THEMATIC MODELING

Annotation

Thematic modeling is an important component of Natural Language Processing (NLP) tasks and can be used in various fields. Topic models have two main problems: identifying topics that can be used to predict outcomes, and facilitating understanding of topics that have already been found. In this article, the modern methods of thematic modeling are calculated Non-negative Matrix Factorization (NMF), Latent Semantic Allocation (LSA), Probabilistic Latent Semantic Analysis (PLSA), lda2vec deep learning, the learning model and tBERT are discussed. Topic modeling is typically applied to large language corpora. Topic modeling is typically applied to large language corpora. Thematic modeling creates clusters of three types of words - co-occurring words; distribution of words and a histogram of words by topic. This article presents the results obtained by applying the LDA method to sample texts in the Uzbek language corpus.

Key words: Thematic modeling, LDA, LSA, PLSA, lda2vec, tBERT, unsupervised methods, language corpus.

СОВРЕМЕННЫЕ МЕТОДЫ ТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ

Аннотация

Тематическое моделирование является важным компонентом задач обработки естественного языка (НЛП) и может использоваться в различных областях. Тематические модели имеют две основные проблемы: определение тем, которые можно использовать для прогнозирования результатов, и облегчение понимания уже найденных тем. В этой статье рассчитываются современные методы тематического моделирования. Факторизация неотрицательных матриц (NMF), латентно-семантическое распределение (LSA), вероятностный латентно-семантический анализ (PLSA), глубокое обучение lda2vec, обсуждаются модель обучения и tBERT. Тематическое моделирование обычно применяется к большим языковым корпусам. Тематическое моделирование обычно применяется к большим языковым корпусам. Тематическое моделирование создает кластеры из трех типов слов – одновременно встречающихся слов; распределение слов и гистограмма слов по темам. В данной статье представлены результаты, полученные применением метода LDA к выборочным текстам корпуса узбекского языка.

Ключевые слова: Тематическое моделирование, LDA, LSA, PLSA, lda2vec, tBERT, неконтролируемые методы, языковой корпус.

Kirish. Tematik modellashtirish – bu hujjatlar to'plamidagi yashirin mavhum mavzularni avtomatik ravishda aniqlash uchun statistik va mashinali o'qitish modellaridan foydalanadigan matnni intellektual qayta ishlash usullari to'plami. Tematik modellashtirish, shuningdek, hujjatlardagi so'z va so'z birikmalari shablon (qolip)larini aniqlashga va hujjatlar to'plamini samarali tushunishga yordam beradigan

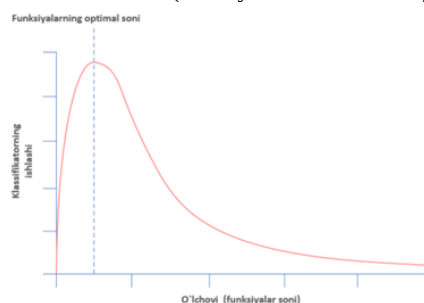
so'z guruhlari va shunga o'xshash birikmalarni avtomatik ravishda klasterlash qobiliyatiga ega bo'lgan nazoratsiz mashinali o'qitish usullari to'plamidir [1,2,3].

NLPda tematik modellashtirishning asosiy maqsadi o'zaro bog'liq bo'lgan so'zlarning birikmasi sifatida ifodalangan so'zlar klasteri bo'lgan mavzularni aniqlashdir. Har bir mavzu bir yoki bir nechta hujjatlarga tegishli

bo'lganligi sababli, o'z navbatida har bir hujjatni bir yoki bir nechta mavzularning kombinatsiyasi sifatida ifodalash mumkin [2,3,5].

O'lchov muammosi (Curse of Dimensionality, CoD). NLP modeliga qo'shimcha funktsiyalar qo'shilib, model samaradorligini oshirish uchun o'rgatish uchun zarur bo'lgan ma'lumotlar miqdori eksponent ravishda oshadi (1-

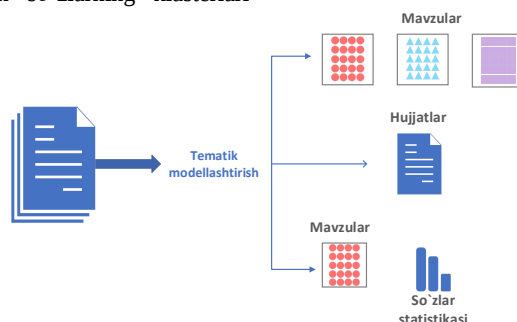
rasm). O'lchovlar NLP va ML modellari kontekstidagi xususiyatlar hisoblanadi [4,6,7]. O'lchovlar, ayniqsa, mashinali o'qitish algoritmlari uchun juda muhim bo'lib, bunda modelning aniqligi dastlab xususiyatlar soni ortishi bilan ortadi, lekin o'lchovlilik oshishi tufayli ma'lumotlarni doimiy ushlab turish bilan o'lchovlilik oshgani sayin yomonlasha boshlaydi.



1-rasm. NLPda mashinali o'qitishda o'lchov muammosi

NLPda tematik modellash Hujjatlar to'plamini tahlil chiqish jarayonida, har bir hujjat bir nechta fikrlar yoki g'oyalarning kombinatsiyasi bo'lishi mumkin. Bu holda, ushbu fikrga tegishli so'zlar hujjatlar bo'ylab turli nisbatlarda aniqlanadi. Shu sababli, har bir hujjatdagi so'zlarning statistikasini aniqlash va o'xshash so'zlarning klasterlari

bo'lgan mavzularni yaratish mumkin [8]. Tematik modellar bizga hujjatlarning butun matnida o'xshash ma'noga ega so'zlar birikmasi va har bir berilgan hujjatda mavzular birikmasi sifatida yashiringan potentsial mavzularni aniqlashga yordam beradi (2-rasm).

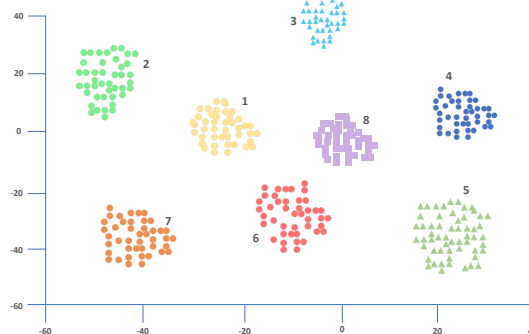


2-rasm. Tematik modellash jarayoni

Asosiy komponentlar tahlili (Principal Component Analysis, PCA). Asosiy komponentlar tahlili statistik usul bo'lib, asosiy g'oya ma'lumotlar to'plamida mavjud bo'lgan o'zgarishlarni imkon qadar ko'proq saqlab qolgan holda ko'p sonli o'zaro bog'liq o'zgaruvchilardan iborat ma'lumotlar to'plamining o'lchamini kamaytirishdir. PCA algoritmining ishlashi tamoyili quyidagicha [8,9,10]:

Ma'lumotlar to'plamidagi ustunlarni **asosiy komponentlar (principal components, PCs)** deb nomlangan yangi o'zgaruvchilar to'plamiga aylantirish lozim.

Asosiy komponentlar o'zaro bog'liq emas va ularning xarakterli qiymatlari asosida tartibga solinishi mumkin. Shuning uchun barcha asl o'zgaruvchilarda mavjud bo'lgan o'zgarishlarning ko'p qismini saqlab qolgan holda birinchi navbatda bir nechta asosiy komponentlarni tanlash mumkin.



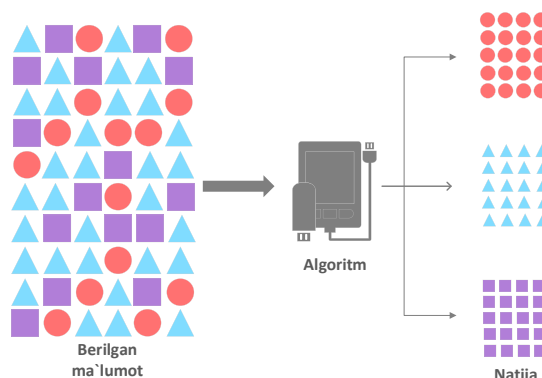
3-rasm. Asosiy komponentlar tahlili jarayoni

3-rasmdagi ranglar ma'lumotlar to'plamidagi maqsadli o'zgaruvchilar / guruhlar / klasterlar teglarini, koordinata o'qlari esa PCA usuli yordamida olingan dastlabki ikkita asosiy komponentni ifodalaydi.

Ma'lumotlarni klasterlashtirish. **Ma'lumotlarni klasterlash** – bu ma'lumotlar to'plamida kichik guruhlarini aniqlash yoki klasterlash uchun nazoratsiz mashinani o'rganish usuli (4-rasm) [9,10].

Klasterlashning asosiy g'oyasi quyidagicha: Ma'lumotlar to'plamidagi kuzatuvlarni turli xil guruhlariga

bo'lish orqali har bir guruh ichidagi kuzatuvlar bir-biriga juda o'xshash va turli guruhlardagi kuzatuvlar bir-biridan mutlaqo farq qilishi lozim. Kuzatishlar o'rtasidagi o'xshashlikni aniqlash uchun mos mezonlarni tanlash kerak. Bugungi kunda klaster tahlili uchun ko'plab usullar ishlab chiqilgan.



4-rasm. Asosiy komponentlar tahlili jarayoni

*Tematik modellash*tirish, *klasterlash* va *PCA*. Yuqoridagi uchta algoritim, *tematik modellash*tirish, *klasterlash* va *PCA* nazoratsiz mashinali o'qitish usuli bo'lib, ma'lumotlar to'plamini oz sonli xulosalar bilan soddalashtirish uchun ishlatiladi. Ushbu usullarning asosiy farqi ulardan qanday foydalanishda:

LDA (*tematik modellash*tirish) va *PCA* usullari o'lchamlarni kamaytirish uchun ishlatilishi mumkin, ammo *LDA* matn nuqtai nazaridan yaxshiroq aniqlik va tushuntirishni ta'minlaydi.

PCA usuli dispersiyaning bir qismini tushuntiradigan kuzatuvlarning past o'lchamli tasvirini topishga intiladi. Chiqish parametrlari bir-biriga bog'liq emasligi sababli, ularni klasterlash uchun kirish parametrlari sifatida foydalanish mumkin.

Klasterlash klasterlar orasidagi masofani maksimal darajada oshirish orqali kuzatishlar orasidagi kichik guruhlarini topishga asoslangan (klasterlar oldindan ma'lum emas)

Manfiy bo'lmagan matritsalarini faktorizatsiya qilish (*Non-Negative Matrix Factorization, NMF*) usuli

NMF – bu matritsani ikkita matritsaga ajratish usuli bo'lib, uchta matritsada ham manfiy elementlar mavjud emas [8,11]. *NMF* usuli asosan tavsiya tizimlari, signallarni qayta ishlash va bioinformatika sohalarida qo'llaniladi. *NMF* usulidagi asosiy g'oya dastlabki kirish matritsasi ikki (odatda) matritsalar (masalan, foydalanuvchilar va filmlar) o'rtasidagi o'zaro ta'sirlardan iborat bo'lgan yashirin xususiyatlar to'plamidan iborat bo'lib, assotsiatsiyalar kirish matritsasi o'g'irliklardir. Manfiy bo'lmagan matritsa faktorizatsiya usulining afzalligi shundaki, biz qanday o'zaro ta'sirlar sodir bo'lishi mumkinligini aniqlashimiz yoki ularni tartiblashimiz.

Yashirin Dirixle taqsimoti (*Latent Dirichlet Allocation, LDA*) usuli. *NLP*da *LDA* usuli yordamida tematik modellash tirish hujjatlardagi so'zlar asosida mumkin bo'lgan mavzularni aniqlash orqali matnli hujjatlar to'plamida **yashirin (latent)** mavzularni aniqlashga imkon beradi [12,13]. *Yashirin Dirixle taqsimoti* usulida har bir matn hujjat va korpusdagi har bir so'z, har bir mavzu va so'zlar o'rtasidagi munosabatlar yashirin o'zgaruvchilar yordamida modellash tiriladi.

Korpusdagi har bir hujjat yashirin o'zgaruvchilar (mavzular) bo'yicha **Dirixle taqsimoti** yordamida taqdim etiladi va har bir mavzu barcha hujjatlardagi barcha so'zlar bo'yicha boshqa *Dirixle taqsimoti* orqali hisoblanadi. *LDA* usulining umumiy modeli *Dirixle taqsimotlarida* o'qitilgan generativ ehtimollik modeliga ega **Bayes tizimiga** asoslangan. Bayes taxmini bilan so'zlarga shartli ehtimolliklarni qo'llash orqali kuzatilgan ma'lumotlar (so'zlar) asosida mavzularni modellash mumkin.

LDA usuli yordamida tematik modellash tirishni amalga oshirish. *NLP*da tematik modellash tirishni amalga oshirish uchun *LDA* usulida foydalanilganda, biz kirish matritsasi sifatida so'zlar to'plamini olamiz, chunki bu

ehtimollik modeli hisoblanadi. Keyingi qadamda *LDA* algoritmi matritsani ikkita kichik matritsaga ajratadi:

hujjatdan mavzu matritsasi;

so'zdan mavzu matritsasi.

Ushbu ikkita matritsa shunday optimallashtirilganki, ular bir-biriga ko'paytirilganda so'z matritsasini eng kam **xatolik (lowest error)** bilan hosil qilinadi.

Yashirin semantik taqsimlash (*Latent Semantic Allocation, LSA*). **Yashirin semantik taqsimlash** – *NLP*da tematik modellash tirish, shuningdek, taqsimlash semantikasi tushunchasidan foydalanadigan o'lchamlarni qisqartirish usuli bo'lib, ushbu usulda ma'nosi o'xshash so'zlar birgalikda uchraydi [14,15].

LSA usuli vositasida hujjat ma'lumotlaridagi yashirin tushunchalarni har bir hujjatdagi so'zlarni *termin hujjat matritsasi* (*term document matrix*, matnning vektor tasviri) sifatida ifodalanadi. Shuningdek, *LSA* usuli vositasida *IDF* matritsasini shakllantirish va keyingi qadamda uni *singular qiymat dekompozitsiyasi* (*Singular Value Decomposition*) dan foydalangan holda alohida *hujjat-mavzu matritsasi* (*document-topic matrix*) va *mavzu-termin matritsasi* (*topic-term matrix*)ga ajratiladi [14]. Shunday qilib, *LSA* usulida hujjatlar va terminlar o'zlarining past darajali vektor ko'rinishlariga ajratiladi. Shuning uchun *hujjat-hujjat* (*document-document*), *hujjat-termin* (*document-term*), *termin-termin* (*term-term*)ning o'xshashligi yoki semantik munosabatlari yashirin semantik o'lchovlar yordamida hisoblanishi mumkin.

Xususan, **$P(D,W)$** modelida, har qanday **d** hujjat va **w** so'zi uchun **$P(d,w)$** hujjat-termin matritsasi yozuvga mos keladi. Har bir hujjat bir nechta mavzularning kombinatsiyasi va har bir mavzu so'zlar to'plami bo'lgan mavzu modellari bilan solishtirganda, *LSA* ushbu taxmin(bashorat)larga ehtimollik elementini qo'shadi:

d hujjati berilgan bo'lsa, **z** mavzusi ushbu hujjatda **$P(z|d)$** ehtimoli bilan mavjud.

z mavzusi berilgan bo'lsa, **w** so'zi **z** dan **$P(w|z)$** ehtimolli bilan olingan.

lda2vec chuqur o'rganish modeli

lda2vec –*NLP*da tematik modellash tirish algoritmi bo'lib, *word2vec* modelining **skip-gram** arxitekturasini *Dirixle* optimallashtirilgan siyrak mavzu terminlari bilan aralashtirish orqali so'zlar va hujjatlarga mos tasvirlar yaratadi. *Lda2vec* modeli asosida kontekstli so'zlarni bashorat qilish uchun vektor so'zini to'g'ridan-to'g'ri ishlatish o'rni, biz bashorat qilish uchun **kontekst vektoridan** foydalanamiz. Kontekst vektorini so'z vektorlari va hujjat vektorlarining yig'indisi sifatida yaratiladi.

tBERT usuli. Bert – mavzularni BERT kabi oldindan tayyorlangan kontekstual tasvirlar bilan birlashtirgan *NLP*da tematik modellash tirish uchun yana bir mavzuni modellash tirish modelidir. *LDA* kabi tipik tematik modellash tirish usuli va Bert Base modellari birlashtirish tabiiy tilidagi ma'lumotlar to'plamlariga neyron tarmoqlari bo'yicha samarali ishlashni yaxshilaydi.

Tematik modellar, shuningdek, *tavsiya qiluvchi tizimlar, chatbotlar, qidiruv va virtual yordamchilar* kabi yangi NLP ilovalarida, *qidiruv tizimlarini so'rovlarni kengaytirish, mijozlarga xizmat ko'rsatish va fikr-mulohazalarni ko'rib chiqish* kabi NLP vazidalarida qo'llaniladi.

Xulosa. Til korpusi matnlarini tahlil qilishda o'ltchov, *xususiyatlar* soni juda katta bo'lgan modellarni o'qitishni qiyinlashtiradi va modellarning samaradorligini pasaytiradi. mashinali o'qitish Nazoratsiz vazifalarida qo'llaniladigan tematik modellashtirish teglash sifatida ko'rib chiqiladi va

birinchi navbatda, til korpusidan zarur ma'lumotlarni olish uchun ishlatiladi, hamda so'rovlarning bajarilish samaradorligining oshishiga yordam beradi. Tematik modellashtirish qidiruv tizimlarida mavzular bo'yicha foydalanuvchi qiziqishlarini xaritalashda keng qo'llaniladi. Bugungi kunda tematik modellashtirish usullari: *hujjatlarini tasniflash, toifalarga ajratish, umumlashtirish* kabi NLP vazifalarini hal qilishda qo'llanilmoqda. Shuningdek, tematik modellashtirish usullari ijtimoiy tarmoqlardagi foydalanuvchilarning his-tuyg'ularini tahlil qilish imkonini beradi.

ADABIYOTLAR

1. Gao, Q., Huang, X., Dong, K., Liang, Z., & Wu, J. (2022). Semantic-enhanced topic evolution analysis: a combination of the dynamic topic model and word2vec. *Scientometrics*, 127(3). <https://doi.org/10.1007/s11192-022-04275-z>
2. Zou, X., Zhu, Y., Feng, J., Lu, J., & Li, X. (2019). A novel hierarchical topic model for horizontal topic expansion with observed label information. *IEEE Access*, 7. <https://doi.org/10.1109/ACCESS.2019.2960468>
3. Korencic, D., Ristov, S., Repar, J., & Snajder, J. (2021). A Topic Coverage Approach to Evaluation of Topic Models. *IEEE Access*, 9. <https://doi.org/10.1109/ACCESS.2021.3109425>
4. Wallach, H. M., Murray, I., Salakhutdinov, R., & Mimno, D. (2009). Evaluation methods for topic models. *ACM International Conference Proceeding Series*, 382. <https://doi.org/10.1145/1553374.1553515>
5. Chang, J., Boyd-Graber, J., Gerrish, S., Wang, C., & Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. *Advances in Neural Information Processing Systems 22 - Proceedings of the 2009 Conference*.
6. Waal, A. de, & Barnard, E. (2008). Evaluating topic models with stability. *Annual Symposium of the Pattern Recognition Association of South Africa*.
7. Hou, C. K. J., & Behdinan, K. (2022). Dimensionality Reduction in Surrogate Modeling: A Review of Combined Methods. In *Data Science and Engineering* (Vol. 7, Issue 4). <https://doi.org/10.1007/s41019-022-00193-5>
8. Elov B., Aloyev N., Yuldashev A. SVD va NMF metodlari orqali tematik modellashtirish // *Труды XI Международной конференции «Компьютерная обработка тюркских языков» «TURKLANG 2023»*. Бухара, 20-22 октября 2023 г.
9. Ogunleye, B., Maswera, T., Hirsch, L., Gaudoin, J., & Brunson, T. (2023). Comparison of Topic Modelling Approaches in the Banking Context. *Applied Sciences (Switzerland)*, 13(2). <https://doi.org/10.3390/app13020797>
10. Qiu, J., Wang, H., Lu, J., Zhang, B., & Du, K.-L. (2012). Neural Network Implementations for PCA and Its Extensions. *ISRN Artificial Intelligence*, 2012. <https://doi.org/10.5402/2012/847305>
11. Wang, J., & Zhang, X. L. (2023). Deep NMF topic modeling. *Neurocomputing*, 515. <https://doi.org/10.1016/j.neucom.2022.10.002>
12. Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, 78(11). <https://doi.org/10.1007/s11042-018-6894-4>
13. Ramage, D., Hall, D., Nallapati, R., & Manning, C. D. (2009). Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora. *EMNLP 2009 - Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: A Meeting of SIGDAT, a Special Interest Group of ACL, Held in Conjunction with ACL-IJCNLP 2009*.
14. Zhuang, F., Karypis, G., Ning, X., He, Q., & Shi, Z. (2012). Multi-view learning via probabilistic latent semantic analysis. *Information Sciences*, 199. <https://doi.org/10.1016/j.ins.2012.02.058>
15. Hofmann, T. (2001). Unsupervised learning by probabilistic Latent Semantic Analysis. *Machine Learning*, 42(1-2).
16. <https://doi.org/10.1023/A:1007617005950>