



**ЎЗБЕКИСТОН РЕСПУБЛИКАСИ
ОЛИЙ ВА ЎРТА МАХСУС ТАЪЛИМ ВАЗИРЛИГИ**

**МИРЗО УЛУҒБЕК НОМИДАГИ ЎЗБЕКИСТОН МИЛЛИЙ
УНИВЕРСИТЕТИ ЖИЗЗАХ ФИЛИАЛИ**



**ИННОВАЦИОН ЁНДАШУВЛАР ИЛМ-ФАН
ТАРАҚҚИЁТИ КАЛИТИ СИФАТИДА:
ЕЧИМЛАР ВА ИСТИҚБОЛЛАР**

**Республика миқёсидаги илмий-техник
анжумани материаллари тўплами**

2020 йил 8-10 октябрь

ЖИЗЗАХ-2020

ОЛИЙ ВА ЎРТА МАХСУС ТАЪЛИМ ВАЗИРЛИГИ

**МИРЗО УЛУҒБЕК НОМИДАГИ ЎЗБЕКИСТОН
МИЛЛИЙ УНИВЕРСИТЕТИ
ЖИЗЗАХ ФИЛИАЛИ**

**ИННОВАЦИОН ЁНДАШУВЛАР ИЛМ-ФАН
ТАРАҚҚИЁТИ КАЛИТИ СИФАТИДА: ЕЧИМЛАР ВА
ИСТИҚБОЛЛАР**

*мавзусидаги Республика миқёсидаги илмий-техник
анжумани материаллари тўплами
(2020 йил 8-10 октябрь)*

ЖИЗЗАХ-2020

Таснифлашни баҳолаш ўлчовлари

О. Ж. Бабомурадов¹, Л. Б. Бобоев², О.Х.Тўрақулов³

1 Муҳаммад ал-Хоразмий номидаги ТАТУ

*2 Муҳаммад ал-Хоразмий номидаги ТАТУ ҳузуридаги АКТ илмий-инновацион
маркази*

3 Жиззах политехника институти

ozod_b_76@mail.ru

replytolochin@gmail.com

o_xolmirzayevich@mail.ru

Аннотация: ушбу мақолада матнли хужжатларни таснифлашни амалга оширишда айнан алгоритмларни баҳолашни амалга ошириш масаласи қараб ўтилган. Алгоритм ишини баҳолаш учун макро ва микро ўртача баҳолаш ёндашувлари таклиф этилган.

Таянч сўзлар: матн, таснифлаш, баҳолаш, recall, precision, accuracy, F-measure, micro average, macro average.

Кириш

Тадқиқот ҳамжамиятида алгоритмларни баҳолаш учун бўлишиладиган ва таққосланадиган ишлаш кўрсаткичларига эга бўлиш афзал саналади. Бироқ, аслида бундай ўлчовлар фақат бир нечта усуллар учун мавжуд бўлиши мумкин. Матнларни таснифлаш усуллари баҳолашда асосий муаммо - маълумотларни йиғишнинг стандарт протоколларининг йўқлиги ҳисобланади. Умумий йиғиш усули мавжуд бўлса ҳам (масалан, *Reuters* янгиликлар корпуси), шунчаки турли хил ўқитиш ва тест тўпламларини танлаш моделларнинг ишлашида номувофиқликларни келтириб чиқариши мумкин [1]. Усулни баҳолаш билан боғлиқ яна бир муаммо - бу алоҳида тажрибаларда ишлатиладиган турли хил ишлаш кўрсаткичларини таққослаш. Ишлаш кўрсаткичлари одатда таснифлаш вазифаларини бажаришнинг ўзига хос жиҳатларини баҳолайди ва шу билан ҳар доим ҳам бир хил маълумотларни тақдим этмайди. Ушбу бўлимда баҳолаш кўрсаткичлари ва ишлаш кўрсаткичлари муҳокама қилинган ва таснифлагичларнинг ишлаш кўрсаткичларини таққослаш усуллари келтирилган. Турли хил баҳолаш

метрикаларининг асосий механикаси турлича бўлиши мумкинлиги сабабли, ушбу кўрсаткичларнинг ҳар бири аниқ нимани англатишини ва улар қандай маълумотларни етказмоқчи эканликларини тушуниш таққослаш учун жуда муҳим ҳисобланади. Ушбу кўрсаткичларга мисол қилиб *recall*, *precision*, *accuracy*, *F-measure*, *micro average* va *macro average*-ларни олиш мумкин . Ушбу кўрсаткичлар тўғри ижобий (TP), тўғри салбий (TN), нотўғри ижобий (FP) ва нотўғри салбий (FN) каби элементларни ўз ичига олган "чалкашлик матрицаси" га асосланган [2]. Ушбу тўрт элементнинг аҳамияти таснифлаш дастурига қараб фарқ қилиши мумкин. Тўғри башоратларнинг барча башоратларга нисбатан қисми *accuracy* (аниқлик) деб аталади (тенглама (1)). Тўғри прогноз қилинган маълум ижобий қисмга *sensitivity* (сезгирлик), яъни тўғри ижобий даражаси ёки *recall* дейилади (тенглама (2)). Тўғри прогноз қилинган салбийларнинг нисбати *specificity* (ўзига хослик) (тенглама (3)) деб номланади. Тўғри башорат қилинган ижобийларнинг барча ижобийларга нисбати *precision*, яъни ижобий башорат қиймат деб аталади (тенглама (4)).

$$accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)} \quad (1)$$

$$sensitivity = \frac{TP}{(TP + FN)} \quad (2)$$

$$specificity = \frac{TN}{(TN + FP)} \quad (3)$$

$$precision = \frac{\sum_{l=1}^L TP_l}{\sum_{l=1}^L TP_l + FP_l} \quad (4)$$

$$recall = \frac{\sum_{l=1}^L TP_l}{\sum_{l=1}^L TP_l + FN_l} \quad (5)$$

$$F1 - Score = \frac{\sum_{l=1}^L 2TP_l}{\sum_{l=1}^L 2TP_l + FP_l + FN_l} \quad (6)$$

Асосий қисм

Тўпламни қайта ишлаш учун бир нечта бинар синф таснифлагичларидан фойдаланилганда битта умумий ўлчов талаб қилинади. Макро ўртача синфлар бўйича ўртача кўрсаткични беради, микро ўртача эса ҳар бир ҳужжат бўйича қарорларни синфлар бўйича бирлаштиради ва кейин кутилмаган ҳолатлар жадвалида самарали ўлчовни чиқаради [3]. Макро ўртача натижаларини қуйидагича ҳисоблаш мумкин:

$$B_{macro} = \frac{1}{q} \sum_{\lambda=1}^q B(TP_{\lambda} + FP_{\lambda} + TN_{\lambda} + FN_{\lambda}) \quad (7)$$

бу ерда B - тўғри ижобий (TP), нотўғри ижобий (FP), нотўғри салбий (FN) ва тўғри салбий (TN) асосида ҳисобланган иккилик баҳолаш ўлчови ва $L = \{\lambda_j : j = 1 \dots q\}$ эса барча ёрликлар тўплами.

Микро ўртача натижалар қуйидагича ҳисобланиши мумкин [4,5]:

$$B_{macro} = B\left(\sum_{\lambda=1}^q TP_{\lambda}, \sum_{\lambda=1}^q FP_{\lambda}, \sum_{\lambda=1}^q TN_{\lambda}, \sum_{\lambda=1}^q FN_{\lambda}\right) \quad (8)$$

Микро ўртача балл натижада ҳар бир ҳужжат учун тенг оғирликларни белгилайди ва у ҳар бир ҳужжат учун ўртача ҳисобланади. Бошқа томондан, ўртача балл ўртача частотани ҳисобга олмаган ҳолда ҳар бир тоифага тенг оғирликларни белгилайди ва шунинг учун бу ўртача ҳар бир тоифага тўғри келади.

F_{β} - классификаторни баҳолаш учун энг оммалашган баҳолаш кўрсаткичларидан бири [2]. Параметр b *recall* ва *precision*- ни мувозанатлаш учун ишлатилади ва қуйидагича аниқланади:

$$F_{\beta} = \frac{(1 + \beta^2)(precision \times recall)}{\beta^2 \times precision + recall} \quad (9)$$

Одатда ишлатиладиган $\beta = 1$ учун, яъни F_1 , *recall* ва *precision* га тенг оғирликлар берилган ҳолда тенглама (9) қуйидагича соддалаштирилиши мумкин:

$$F_1 = \frac{2TP}{2TP + FP + FN} \quad (10)$$

F_β , *recall* ва *precision* га асосланганлиги сабабли, у чалкашлик матричасини тўлиқ акс эттирмайди.

Хулоса ва тавсиялар

Матнларни туркумлаш бўйича статистик ёндашувларни баҳолади ва таснифлагич алгоритмларини таққослашда эътиборга олиниши зарур бўлган куйидаги муҳим омилларга эгаллигини кўрсатди:

- Усуллар ва тажрибалар бўйича таққослама баҳолаш, натижада ишлашнинг ўзгариши асосидаги омиллар тўғрисида тушунча беради ва келажакда баҳолаш методологиясини яхшилайти;
- Йиғишдаги ўзгарувчанликнинг таъсири, масалан, ёрлиқга эга бўлмаган ҳужжатларни ўқитиш ёки тестлаш тўпламига қўшиш ва уларни салбий ҳолатлар деб ҳисоблаш жиддий муаммо бўлиши мумкин;
- Категориялар рейтингини баҳолаш ва иккилик таснифлашни баҳолаш интерактив дастурларда таснифлагичнинг фойдалилигини кўрсатади ва уларнинг навбати билан оммавий режимда ишлатилишини таъкидлайди. Таснифлагичлар учун ишлаш ўлчовларининг ҳар иккала турига эга бўлиш ажратиш стратегиясининг таъсирини аниқлашда ёрдам беради;

Катта категориялар фазосида таснифлагичларнинг кўламини баҳолаш камдан-кам ўрганиладиган соҳа ҳисобланади.

Фойдаланилган адабиётлар

16. Yang, Y. An evaluation of statistical approaches to text categorization. *Inf. Retr.* 1999, 1, 69–90.
17. Lever, J.; Krzywinski, M.; Altman, N. Points of significance: Classification evaluation. *Nat. Methods* 2016, 13, 603–604 .
18. Manning, C.D.; Raghavan, P.; Schütze, H. Matrix decompositions and latent semantic indexing. In *Introduction to Information Retrieval*; Cambridge University Press: Cambridge, UK, 2008; pp. 403–417.

19. Sebastiani, F. Machine learning in automated text categorization. ACM Comput. Surv. (CSUR) 2002, 34, 1–47.

20. Tsoumakas, G.; Katakis, I.; Vlahavas, I. Mining multi-label data. In Data Mining and Knowledge Discovery Handbook; Springer: Berlin/Heidelberg, Germany, 2009; pp. 667–685.

БИОЛОГИК ПОПУЛЯЦИЯ МАСАЛАСИНИ ГЕТЕРОГЕН МУҲИТДА ЕЧИШ УСУЛЛАРИ

Д.К. Мухамедиева

ТАТУ ҳузуридаги “АКТ илмий –инновацион” маркази, Тошкент
matematichka@inbox.ru

Аннотация. Мақолада бир жинсли бўлмаган муҳитда ночизикли диффузияли популяция жараёнининг квазичизикли параболик тенгламалари билан ифодаланувчи ночизикли математик моделлари сонли ва аналитик тарзда таҳлил қилинди. Ночизикли диффузияли популяция моделларини қуришнинг асосий ёндашувлари, ночизикли диффузияли популяция моделларининг автомодел ечимлари тадқиқ қилинди. Параболик турдаги квазичизикли тенглама билан ифодаланувчи биологик популяциянинг математик модели учун ночизикли парчалаш алгоритмини асослаб берилди

Калит сўзлар. Биологик популяция, ночизикли диффузия, квазичизикли тенглама, автомодел ечим, парчалаш алгоритми.

Кириш

Ночизикли масалаларни ўрганиш жадал суръатлар билан ривожланиб боровчи синергетика фани билан узлуксиз боғлиқдир.

С.П.Капица, С.П.Курдюмов ва Г.Г.Малинецкийларнинг “Синергетика ва келажак башоратлари” [1] номли китобида синергетика фанининг пайдо бўлишига олиб келган ночизикли ҳодисаларга батафсил таъриф берилди.

Чет эл математикларининг бир қатор ишларида сингуляр ночизикликли

муносабатларини оптималлаштиш жараёнини автоматлаштириш....	99
17. А. Мирзамов, М.Дадахонов. Нутқ товушларини таниб олишда рақамли маълумотларни силлиқлаш.....	104
18. О.У.Бегулов, Б.У. Норбоев. JAVA FX муҳитида russion-uzbek аудио луғатини ишлаб чиқиш.....	111
19. Qodirov Farrux Ergash o'g'li. QR-kod asosida elektron kutubxona tizimini dasturiy taminotini yaratish.....	114
20. Рахматов Дилмурод Рустам ўғли. VR(virtual reality) ва AR (augmented reality) технологияларидан фойдаланиш имкониятлари	118
21. Saitov Sirojiddin Abduvaliyevich. Iqtisodiyotni rivojlantirishda kichik biznes va xususiy tadbirkorlik o'rni.....	124
22. O. Raximov. Fizika tarixining muhim bosqichlari.....	132
23. Н. Равшанов, Т.Р. Шафиев. Зарарли моддаларнинг атмосферада кўчиши ва диффузияси жараёнига таъсир этувчи асосий омилларни сонли тадқиқ қилиш.....	139
24. Suvanov Sh. M. Visual quality enhancement of images based on infrared data.....	145
25. О. Ж. Бабомурадов, Л. Б. Бобоев. Чуқур ўқитишга асосланган матнни таснифлаш моделлари.....	149
26. О. Ж. Бабомурадов, Л. Б. Бобоев, О.Х.Тўрақулов. Таснифлашни баҳолаш ўлчовлари.....	156
27. Д.К. Мухамедиева. Биологик популяция масаласини гетероген муҳитда ечиш усуллари.....	160
28. Д.К. Мухамедиева. Биологик популяция масалалари ечимнинг хоссалари.....	166
29. Д.Т.Мухамедиева. Норавадан маданий алгоритм асосида оптималлаштириш масаласини ечиш.....	171
30. Д.Т.Мухамедиева. Зарралар галаси усули асосида оптималлаштириш масаласини ечиш.....	177
31. Камилов М.М., Худайбердиев М.Х., Хамроев А.Ш. Тимсолларни аниқлаш масаласида декомпозицион модуллар комбинацияси.....	183
32. Raxmanov Q.S., Maxmanov B.Q. Davlat boshqaruvida davlat xizmatchilarga bo'lgan ehtiyojini baholashning o'rni.....	189
33. О. Ж. Бабомурадов Л. Б. Бобоев, Х.Т. Дусанов. Матннинг кетма-кетлик модели.....	195
34. С. С. Кобилов, И. М. Раббимов. Ўзбек тилида ёзилган матнларни автоматик синфлаш масаласи.....	200
35. А.Абдуллаев, О.Рахимов, И.Эгамов. Точечные дефекты в полупроводниках.....	206
36. М. А. Umarov. Yo'l belgilarining shaklga asoslangan klassifikatsiyasi..	212
37. Д.К. Шарипов. Разработка модели и веб-приложения для прогнозирования экологического состояния атмосферы.....	216
38. И.Е.Жуковская. Инновационные подходы совершенствования	