

TIBBIY TASHXIS QILISH MASALALARINI YECHISHDA INTELLEKTUAL TAHLIL USULLARI ANSAMBLARIDAN FOYDALANISH

To'rayev Jo'rabek Nurbek o'g'li

Iqtisodiyot va pedagogika universiteti "Matematika" kafedrası stajyor-assistenti.

Email: 0511jurabek@gmail.com Tel: (+99891)-222-43-18

Shonazarov Uchqun Nurbek o'g'li

Qarshi davlat universiteti Amaliy matematika mutaxassisligi, 1-kurs magistranti.

Email: shonazarovuchqun@gmail.com

<https://doi.org/10.5281/zenodo.15492753>

Annotatsiya. Maqolada tibbiy tashxis qilish jarayonida qo'llaniladigan ko'p sinfli klassifikatsiya masalalarini yechish uchun ansamblli o'rganish yondashuvlaridan foydalanish masalasi qaralgan. Stacking, Random Forest va LightGBM ansambl usullaridan foydalanib o'quv tanlanmalaridagi klassifikatsiya aniqligini hisoblash kabi ishlar natija sifatida olingan.

Kalit so'zlar: Ensemble learning, bagging, random forest, Boosting, XGBoost, LightGBM, CatBoost.

ИСПОЛЬЗОВАНИЕ МЕТОДОВ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ПРИ РЕШЕНИИ ЗАДАЧ МЕДИЦИНСКОЙ ДИАГНОСТИКИ

Аннотация. В статье рассматривается применение ансамблевых подходов к обучению для решения задач многоклассовой классификации, используемых в процессе медицинской диагностики. Результаты включают расчет точности классификации на обучающих выборках с использованием методов Stacking, Random Forest и ансамбля LightGBM.

Ключевые слова. Ensemble learning, bagging, random forest, Boosting, XGBoost, LightGBM, CatBoost.

USE OF INTELLECTUAL ANALYSIS METHODS IN SOLVING MEDICAL DIAGNOSIS PROBLEMS

Abstract. The article deals with the use of ensemble learning approaches to solve multi-class classification problems used in medical diagnosis. The results include calculating the classification accuracy of training samples using the ensemble methods Stacking, Random Forest and LightGBM.

Keywords: Ensemble learning, bagging, random forest, Boosting, XGBoost, LightGBM, CatBoost.

Kirish. Tibbiy tashxis qilish jarayonida qo'llaniladigan ko'p sinfli klassifikatsiya masalalarini yechish uchun ansamblli o'rganish (ensemble learning) yondashuvlaridan foydalanish yuqori aniqlik va yuqori tezlikni ta'minlaydi. Ushbu usullardan foydalanib tibbiy tashxis qilishni amalga oshiruvchi dasturiy vositalar ishlab chiqish va bu dasturiy vositalar uchun xizmat qiluvchi o'quv tanlanmalarini shakllantirish masalalarini yechish mumkin. Hozirgi kunda keng foydalanilayotgan samarali ansambl usullari hamda ularning kombinatsiyalarini quyida keltiramiz:

- Bagging (Bootstrap Aggregating) va Random Forest algoritmlari ansambli – bu ansambl yuqori aniqlik va turg'unlikka ega. Bunda har bir daraxt mustaqil ravishda o'rgatiladi, bu esa

ortiqcha moslashish (overfitting) muammosini kamaytiradi. Shuningdek bu usul parallel ishlov berish imkoniyatiga egaligi sababli yuqori tezlikda ishlay oladi. Bu algoritmdan o'quv tanlanmasidagi ma'lumotlar katta hajmda bo'lganda va unda o'rtacha murakkablikka ega alomatlar mavjud bo'lganda foydalanish maqsadga muvofiq.

- Boosting va XGBoost / LightGBM / CatBoost – bu ansamblda XGBoost (Ekstremal gradient boosting) tezlik va aniqlik bo'yicha samarali hisoblanadi, LightGBM katta hajmli o'quv tanlanmasiga moslashgan bo'lib, u tezkor hamda kam resurs talab qiladi, CatBoost kategoriyali ma'lumotlar bilan samarali ishlaydi, unda kam moslashuvchanlik(overfitting) muammosi mavjud. Bu ansambldan aniqlik juda muhim bo'lganda va murakkab alomatlar mavjud bo'lsa foydalanish maqsadga muvofiq.

- Stacking (Stacked Generalization) va Meta Learning – bu ansambl turli model ansambllarini birlashtirish orqali samarali bashorat qilish imkoniyatini beradi. Misol uchun bu yerda Random Forest + XGBoost + LightGBM kombinatsiyasini qo'llash mumkin. Bu algoritmlardan model kombinatsiyasidan optimal ansambl yaratish kerak bo'lgan hollarda va alomatlar turli toifaga tegishli bo'lganda foydalanish tavsiya etiladi.

- Voting Ensemble (Hard Voting yoki Soft Voting) – bu ansambl bir nechta model natijasini birlashtirib, eng yaxshi natijaviy bashoratni tanlaydi. Hard Votingda har bir modelning bergan ovozigga ko'ra eng ko'p ovoz olgan sinf tanlanadi. Soft Voting har bir modelning ehtimollikni hisoblaydi. Bunda eng katta ehtimollik tanlanadi. Bu ansambldan turli toifadagi modellarning birlashuvi natijani yaxshilashi mumkin bo'lganda va model yakka o'zi yaxshi ishlamaganda foydalaniladi.

Quyida ansambl usullarining xususiyatlari va imkoniyatlarini taqqoslashni amalga oshiramiz(1-jadval):

Ansambl turi	Tezlik	Aniqlik	Ma'lumot turi
Random Forest	Tez	Yuqori	O'rtacha hajmli, strukturalangan
XGBoost	O'rtacha	Juda yuqori	Katta hajmli, murakkab ma'lumot
LightGBM	Juda tez	Yuqori	Juda katta hajmli ma'lumot
CatBoost	O'rtacha	Yuqori	Kategorik ma'lumotlar
Stacking	Sekin	Eng yuqori	Murakkab model ansambli
Voting Ensemble	O'rtacha	Yuqori	Har xil modellar birlashganda

Jadvalda keltirilgan ma'lumotlardan kelib chiqib, quyidagi xulosalarga kelimiz:

- LightGBM + Voting ansambli eng tez va aniq ishlovchi hisoblanadi;
- Stacking (Random Forest + XGBoost + LightGBM + Meta Model) ansambli yuqori aniqlikni ta'minlaydi;
- Voting +(XGBoost yoki LightGBM) ansambli samaradorligi yuqori hisoblanadi.

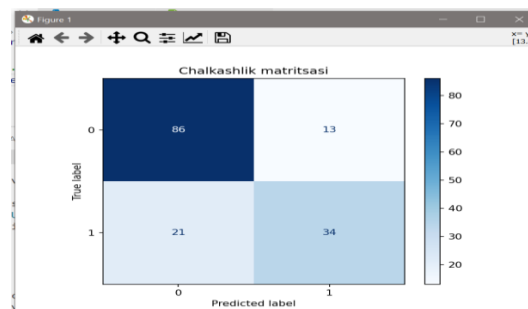
Yuqorida keltirilgan ansambllardan birini tanlash uchun quyidagilarni hisobga olish kerak bo'ladi. Yuqori aniqlik talab etilganda Stacking yoki XGBoostdan, yuqori samaradorlik va tezlikka ehtiyoj sezilganda LightGBM yoki Random Forest, kategoriyali ma'lumotlar bilan ishlashda CatBoost, murakkab strukturali ma'lumotlar bilan ishlaganda Stacking yoki Deep Learning ansambllaridan foydalanish mumkin.

Ansambl usullarini qo'llashga misol sifatida diabetes [https://www.kaggle.com/datasets/akshaydattatraykhare/diabetes-dataset] o'quv tanlanmasining klassifikatsiya aniqligini hisoblash hamda unga mos chalkashlik matritsasini ansambl metodlari yordamida hosil qilishga ansambl usullarini qo'llashni ko'ramiz.

Diabetesda 768 ta mijoz haqida ma'lumotlar joylashgan bo'lib, har bir mijoz uchun unda qand kasalligining bor yoki yo'qligini aniqlashga mo'ljallangan 8 ta tahlil natijalari hamda kasallik bor yoki yo'qligini bildiruvchi(0 yo'q, 1 bor) sinf mavjud, ya'ni tanlanma 768 ta satr, 9 ta ustundan iborat jadval ko'rinishida shakllantirilgan.

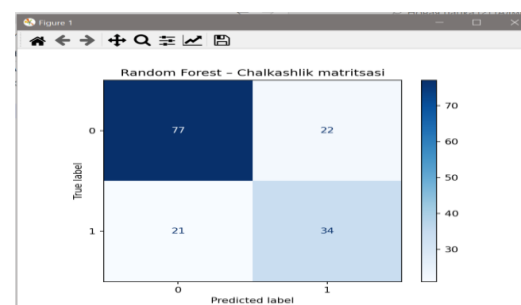
Stacking. Klassifikatsiya hisoboti:

	precision	recall	f1-score	support
0	0.80	0.86	0.83	99
1	0.71	0.62	0.66	55
accuracy			0.77	154
macro avg	0.76	0.74	0.74	154
weighted avg	0.77	0.77	0.77	154



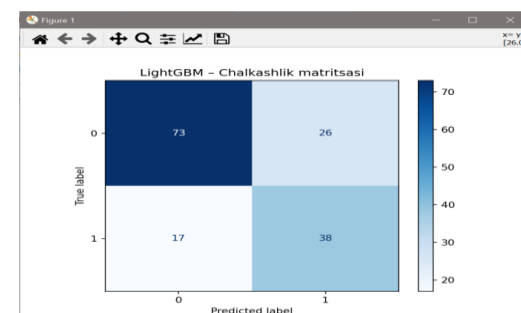
Random Forest. Klassifikatsiya hisoboti:

	precision	recall	f1-score	support
0	0.79	0.78	0.78	99
1	0.61	0.62	0.61	55
accuracy			0.72	154
macro avg	0.70	0.70	0.70	154
weighted avg	0.72	0.72	0.72	154



LightGBM. Klassifikatsiya hisoboti:

	precision	recall	f1-score	support
0	0.81	0.74	0.77	99
1	0.59	0.69	0.64	55
accuracy			0.72	154
macro avg	0.70	0.71	0.71	154
weighted avg	0.73	0.72	0.72	154



olingan natijalar shuni ko'rsatadiki, stackingda aniqlik yuqori, random forest va lightgbmda aniqlik biroz past bo'lsada algoritmlarning ishlash tezligi va samaradorligini ko'rish mumkin.

Xulosa. Yuqorida ko'rib chiqilgan usullardan foydalanib, o'quv tanlanmalaridan informativ alomatlar fazosini shakllantirish, tanlanmaning klassifikatsiya aniqligini hisoblash, ko'p sinfli klassifikatsiya masalalarini yechish kabi bir qator keng qo'llanmaga ega bo'lgan masalalarni hal etish mumkin.

Olingan natijalardan tibbiy tashxis qilish masalalarida foydalaniladigan o'quv tanlanmalaridagi informativ alomatlarini aniqlash, tanlanmalarining klassifikatsiya aniqligini hisoblash uchun foydalanilganda yuqori aniqlikda tashxis qiluvchi dasturiy vositalar ishlab chiqish imkoniyati tug'iladi.

REFERENCES

1. Ignatov N.A., Madrahimov Sh.F. "Berilganlarning intellektual tahlili"
2. Григорьева Г.М. Реализация методов интеллектуального анализа данных в медицинской практике / Г.М.Григорьева, В.Ю. Ходченков // Системы компьютерной математики и их приложения. – 2018. – № 19. – С.124-129.
3. Shonazarov U.N., To'rayev J.N. "Intellektual tahlil usullarini tibbiy tashxislashda qo'llash"
4. Asqar, M., & Jorabek, T. (2024). *KOSHI MASALASI YECHIMINING TURG'UNLIGI. JOURNAL OF THEORY, MATHEMATICS AND PHYSICS*, 3 (10), 3–5.
5. To'rayev, J. R., & Temirova, Z. (2025). *FUNKSIYANING EGILISH NUQTALARI. Modern Science and Research*, 4(4), 1359-1363.
6. Bozarov, D. (2023). Bo 'lajak iqtisodchi talabalarning iqtisodiy kompetensiyasini rivojlantirishning matematik tahlili. *Академические исследования в современной науке*, 2(27), 84-90.
7. Bozarov, D. (2022). CHIZIQLI VA KVADRATIK MODELLASHTIRISH MAVZUSINI MUSTAQIL O'RGANISHGA DOIR MISOLLAR. *Евразийский журнал математической теории и компьютерных наук*, 2(6), 24-28.
8. Bozarov, D., & Rahmonov, B. (2024). Kombinatorikaning paydo bo 'lishi va hayotiy masalalarga tadbiqu. *Modern Science and Research*, 3(6).
9. Uralovich, B. D. (2022). CHIZIQLI ALGEBRAIK TENGLAMALAR SISTEMALARIGA OID MASALALAR. *Science and innovation*, 1(A2), 163-171.
10. Bozarov, D. (2025). EKONOMETRIYADA CHIZIQLI VA CHIZIQSIZ REGRESSIYA MODELLARINING MATEMATIK TAHLILI. *Modern Science and Research*, 4(4), 1020-1025.