UO‘K 37.018.43:004.9

MASOFAVIY TA’LIMDA TALABA FAOLIYATINI MONITORING QILISH VA BASHORATLASH DASTURIY TA’MINOTI

Zikrilloyeva Feruza Abdurashid qizi,
Buxoro davlat universiteti Amaliy matematika va dasturlash texnologiyalari mutaxassisligi 1-bosqich magistranti
zikrillayevaferuza28@gamil.com
Shirinov Ziyomat Zoyirovich,
Buxoro davlat universiteti Amaliy matematika va dasturlash texnologiyalari kafedrasi katta o‘qituvchisi
ziyomatshirinov@umail.uz

Maqolada masofaviy ta’limda talabalar faoliyatini monitoring qilish va akademik xavfni erta bashoratlash dasturiy ta’minotining arxitekturasi hamda algoritmik asoslari taklif etiladi. Tizim o‘quv boshqaruv tizimi (LMS) jurnallari, test natijalari va topshiriqlar bajarilishi ma’lumotlarini birlashtirib, har bir talabaning o‘zlashtirmaslik ehtimolini baholaydi hamda tyutor va o‘qituvchilarga ogohlantirish yuboradi. Logistik regressiya, Random Forest va Gradient Boosting usullari 2000 talabadan iborat sintetik ma’lumotlar to‘plamida solishtirildi; eng yaxshi model ROC-AUC bo‘yicha 0,90 natija ko‘rsatdi. Ilmiy yangilik monitoring, bashoratlash va tushuntiriladigan ogohlantirishni yagona oqimda birlashtiruvchi modulli yechimdan iborat. Natijalar xavf omillarini ochiq ko‘rsatib, o‘qituvchiga o‘z vaqtida pedagogik aralashuv imkonini beradi.

Kalit so‘zlar: *masofaviy ta’lim, monitoring, bashoratlash, ta’lim tahlili, mashinali o‘qitish, LMS, akademik xavf, erta ogohlantirish, logistik regressiya, dasturiy ta’minot.*

SOFTWARE FOR MONITORING AND PREDICTING STUDENT PERFORMANCE IN DISTANCE EDUCATION

The article proposes the architecture and algorithmic basis of software for monitoring student activity and predicting academic risk in distance education. The system integrates learning management system (LMS) logs, quiz results and assignment submission data, estimates the probability of academic failure for each student, and sends alerts to tutors and instructors. Logistic regression, Random Forest and Gradient Boosting were compared on a synthetic dataset of 2,000 students; the best model achieved a ROC-AUC of 0.90. The scientific novelty lies in a modular solution that combines monitoring, prediction and interpretable risk-based alerting in a single workflow adapted to higher education.

Keywords: *distance education, monitoring, prediction, learning analytics, machine learning, LMS, academic risk, early warning, logistic regression, software.*

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ МОНИТОРИНГА И ПРОГНОЗИРОВАНИЯ ДЕЯТЕЛЬНОСТИ СТУДЕНТОВ В ДИСТАНЦИОННОМ ОБУЧЕНИИ

В статье предлагаются архитектура и алгоритмическая основа программного обеспечения для мониторинга деятельности студентов и прогнозирования академического риска в дистанционном обучении. Система объединяет журналы системы управления обучением (LMS), результаты тестов и данные о выполнении заданий, оценивает вероятность неуспеваемости каждого студента и направляет предупреждения тьюторам и преподавателям. Логистическая регрессия, Random Forest и Gradient Boosting сравнены на синтетическом наборе данных из 2000 студентов; лучшая модель показала ROC-AUC 0,90. Научная новизна состоит в модульном решении, объединяющем мониторинг, прогнозирование и интерпретируемое оповещение по уровням риска в едином процессе. Результаты позволяют прозрачно выявлять факторы риска и организовать своевременное педагогическое вмешательство.

Ключевые слова: *дистанционное обучение, мониторинг, прогнозирование, аналитика обучения, машинное обучение, LMS, академический риск, раннее оповещение, логистическая регрессия, программное обеспечение.*

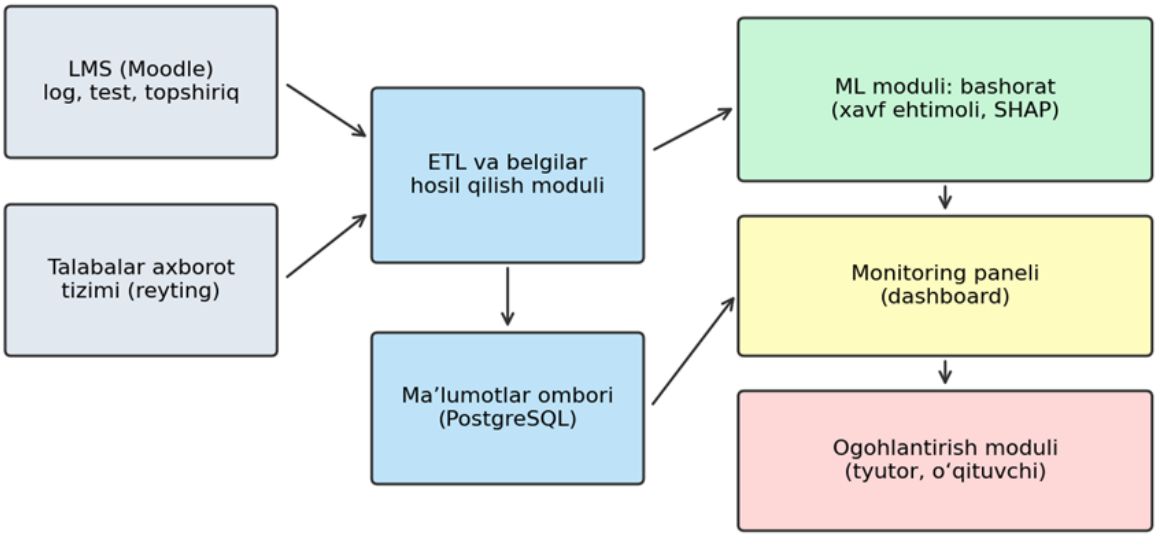
Kirish. Ta’lim jarayonining raqamlashtirilishi masofaviy ta’limni oliy ta’lim tizimining doimiy tarkibiy qismiga aylantirdi. O‘zbekistonda «Raqamli O‘zbekiston – 2030» strategiyasi ta’lim sohasida raqamli platformalar va elektron xizmatlarni rivojlantirishni ustuvor vazifa sifatida belgilagan [1]. Biroq ta’limning bu shaklida ta’lim jarayoni ishtirokchilari ustidan pedagogik kuzatuv zaiflashadi: o‘qituvchi talabaning qiziqishi pasayayotganini, topshiriqlarni kechiktirayotganini yoki o‘qishni tark etish arafasida ekanini ko‘pincha semestr oxirida, ya’ni choralar ko‘rish uchun kech bo‘lganda bilib oladi.

Mavzuga doir adabiyotlar tahlili. Ta’lim tahlili (learning analytics) yo‘nalishi bu muammoni o‘quv muhitida to‘planadigan ma’lumotlarni yig‘ish, o‘lchash va tahlil qilish orqali hal etishni taklif qiladi [2]. Ta’lim ma’lumotlarini intellektual tahlil qilish usullari talabaning o‘zlashtirishi va o‘qishni tark etish xavfini bashoratlashda keng qo‘llanilmoqda [3; 4; 7]. Purdue universitetidagi Course Signals tizimi bunday yondashuvning dastlabki amaliy namunalaridan bo‘lib, talabaga «yo‘l chiroq» ko‘rinishidagi signal berish orqali o‘zlashtirishni oshirishga qaratilgan [6].

Chuqur o‘rganish (deep learning) modellari ommaviy ochiq onlayn kurslarda (MOOC) o‘qishni tark etishni bashoratlashda qo‘llanilmoqda va shaxsiylashtirilgan aralashuv uchun asos bo‘lib xizmat qiladi [8]. MOOC platformalarida talabalarning harakatlari (video ko‘rish, test topshirish, forumdagi faollik) vaqt bo‘yicha ketma-ketlik ko‘rinishida qayd etiladi. Takroriy neyron tarmoqlar (RNN, LSTM) kabi arxitekturalar bunday ketma-ketliklardagi yashirin qonuniyatlarni belgilarni qo‘lda tanlamasdan aniqlay oladi. Shu tufayli ular katta hajmdagi ma’lumotlarda klassik usullarga (logistik regressiya, tasodifiy o‘rmon) nisbatan yuqoriroq aniqlik berishi mumkin. Biroq bunday modellar ko‘p ma’lumot va hisoblash resurslarini talab qiladi.

Mavjud tadqiqotlar tahlili shuni ko‘rsatadiki, ko‘pchilik ishlar bashorat modelining aniqligiga e’tibor qaratadi, monitoring, bashoratlash va pedagogik aralashuvni yagona dasturiy oqimda birlashtirish esa yetarlicha yoritilmagan [5; 7]. Bundan tashqari, izohlanmaydigan («qora quti») modellar o‘qituvchining natijaga ishonchini pasaytiradi, shu sababli qaror omillarini tushuntirish zarur [13].

Tadqiqot metodologiyasi. Taklif etilayotgan dasturiy ta’minot 1-rasmda ko‘rsatilgan modullardan tarkib topadi: ma’lumot manbalari; ETL va belgilar hosil qilish moduli; ma’lumotlar ombori; ML moduli; monitoring paneli; ogohlantirish moduli.



1-rasm. Monitoring va bashoratlash dasturiy ta’minotining arxitekturasi (LMS – o‘quv boshqaruv tizimi; ETL – ma’lumotlarni ajratib olish, o‘zgartirish va yuklash; ML – mashinali o‘qitish)

Talaba faoliyati yettita belgi bilan tavsiflandi: haftalik kirishlar soni, haftalik onlayn vaqt, topshiriqlarning o‘z vaqtida topshirilish ulushi, o‘rtacha test bali, forum faolligi, ko‘rilgan video ulushi va oxirgi kirishdan o‘tgan kunlar. Belgilar tarkibi Open University Learning Analytics dataset kabi ochiq ma’lumotlar to‘plamlarida qo‘llanilgan faoliyat ko‘rsatkichlariga tayangan [10].

Bashorat vazifasi ikkilik tasnif sifatida qo‘yildi: y = 1 – talaba o‘zlashtirmaslik xavfi ostida, y = 0 – xavf yo‘q. Asosiy model sifatida logistik regressiya ishlatildi:

$$P(y=1|x)=\frac{1}{1+e^{-z}},\quad z=w_0+\sum_{i=1}^{m}w_i x_i \qquad (1)$$

bu yerda $P(y=1|x)$– xavf ehtimoli; x_i – i-belgi qiymati; w_i – model koeffitsiyenti; m – belgilar soni (m = 7). Solishtirish uchun Random Forest va Gradient Boosting ansambl usullari tanlandi [12]. Modellar

Python tilida scikit-learn kutubxonasi yordamida qurildi [11], qaror omillarini izohlash uchun SHAP yondashuvi nazarda tutildi [13].

Tajribada 2000 talabadan iborat sintetik ma’lumotlar to‘plami ishlatildi. Sintetik to‘plam yashirin «qobiliyat» va «motivatsiya» omillari orqali belgilar o‘rtasida haqiqiy o‘quv muhitiga xos o‘zaro bog‘liqlikni beradigan statistik generator bilan hosil qilindi; xavf ostidagi talabalar ulushi 42,3 % ni tashkil etdi. Sintetik ma’lumotdan foydalanish shaxsiy ma’lumotlar maxfiligini saqlash va usulni dastlabki sinash uchun tanlandi. Ma’lumotlar 70 % o‘qitish va 30 % sinov qismlariga tabaqalashtirilgan holda ajratildi. Sifat accuracy, precision (P), recall (R), F1 va ROC-AUC ko‘rsatkichlari bilan baholandi:

$$F1 = \frac{2 \cdot P \cdot R}{P+R}, \quad P = \frac{TP}{TP+FP}, \quad R = \frac{TP}{TP+FN} \qquad (2)$$

bu yerda TP, FP, FN – mos ravishda to‘g‘ri musbat, noto‘g‘ri musbat va noto‘g‘ri manfiy natijalar soni. Xavf sohasida recall ko‘rsatkichi alohida ahamiyatga ega, chunki xavf ostidagi talabani o‘tkazib yuborish keraksiz ogohlantirishdan qimmatroq bahoga tushadi.

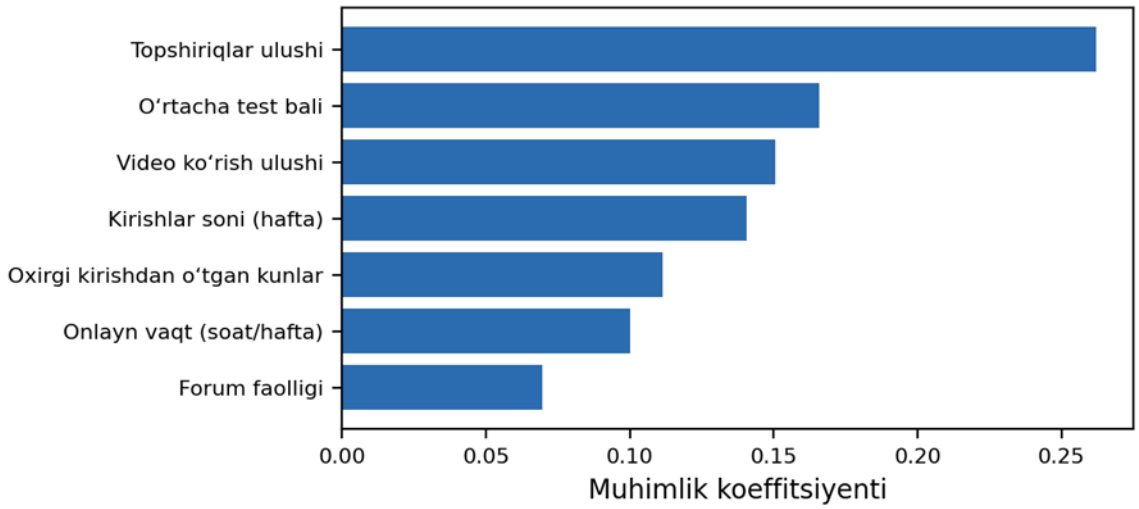
Olingan natijalar va ularning tahlili.

1-jadval. Bashorat modellarining sinov to‘plamidagi ko‘rsatkichlari

Model	Accuracy	Precision	Recall	F1	ROC-AUC
Logistik regressiya	0,832	0,805	0,795	0,800	0,901
Random Forest	0,803	0,774	0,756	0,765	0,884
Gradient Boosting	0,803	0,768	0,768	0,768	0,884

Eng yuqori natijani logistik regressiya ko‘rsatdi (ROC-AUC = 0,901, F1 = 0,800).

Random Forest bo‘yicha belgilarning muhimlik koeffitsiyentlari 2-rasmda berilgan. Eng muhim omil topshiriqlarning o‘z vaqtida topshirilgan ulushi (0,262) bo‘ldi, undan keyin o‘rtacha test bali (0,166) va ko‘rilgan video ulushi (0,150) turadi. Forum faolligi eng past ta’sir ko‘rsatdi (0,070). Bu natija talaba faolligi (engagement) ko‘rsatkichlari o‘zlashtirish bilan bog‘liqligini ko‘rsatgan tadqiqotlarga mos keladi [9].



2-rasm. Random Forest modeli bo‘yicha belgilarning muhimlik koeffitsiyentlari

Bashorat natijalaridan amaliy foydalanish uchun xavf ehtimoli chegaraviy qiymatlar bo‘yicha uch darajaga ajratildi va har biriga pedagogik chora biriktirildi (2-jadval). Chegaraviy qiymatlar boshlang‘ich sifatida olingan bo‘lib, muassasa resurslariga qarab sozlanishi kerak.

2-jadval. Xavf darajalari va pedagogik aralashuv choralari

Daraja	Ehtimol (p)	Tizim harakati
Past	$p < 0,30$	Monitoring panelida yashil belgi; qo‘shimcha chora talab qilinmaydi
O‘rta	$0,30 \le p < 0,60$	Talabaga avtomatik eslatma; topshiriq muddatlari haqida xabar; tyutorga ro‘yxat
Yuqori	$p \ge 0,60$	O‘qituvchi va tyutorga shoshilinch ogohlantirish; xavf omillari izohi bilan individual maslahat rejalashtiriladi

Muhokama. Bu natijani sintetik ma’lumotlar generatori belgilar va xavf o‘rtasida asosan chiziqli bog‘lanish hosil qilgani bilan izohlash mumkin, shuning uchun ansambl usullarining ustunligi yo‘qligini umumlashtirib bo‘lmaydi. Haqiqiy ma’lumotlarda chiziqli bo‘lmagan bog‘lanishlar ko‘proq bo‘lgani uchun modellarni qayta baholash zarur. Shunga qaramay, logistik regressiya sodda, tez o‘qitiladi va koeffitsiyentlari bevosita izohlanadi; bu esa o‘qituvchi uchun ochiqlikni ta’minlaydi.

Taklif etilgan yechimning ustunligi shundaki, u faqat «kim xavfda» degan savolga emas, «nima uchun» va «nima qilish kerak» degan savollarga ham javob beradi. Cheklovlar: (1) natijalar sintetik ma’lumotlarga asoslangan va real ma’lumotlarda tasdiqlanishi shart; (2) belgilar faqat raqamli izlarni qamrab oladi; (3) bashorat sababiy bog‘liqlikni isbotlamaydi, uni o‘qituvchi pedagogik nuqtai nazardan baholashi kerak. Kelgusi ishlarda tizimni real semestr ma’lumotlarida sinash, aralashuvning talaba o‘zlashtirishiga ta’sirini nazorat guruhi bilan baholash va vaqt bo‘yicha (haftama-hafta) bashoratlashni rejalashtirish mumkin.

Xulosa. Masofaviy ta’limning ommalashishi talaba faoliyatini kuzatish va o‘qishdagi qiyinchiliklarni o‘z vaqtida aniqlash muammosini birinchi darajali vazifaga aylantirdi.Shu sababli tadqiqotda monitoring, bashoratlash va pedagogik aralashuvni yagona jarayonda birlashtiruvchi dasturiy ta’minot taklif etildi.

Ishning asosiy natijalari quyidagilardan iborat. Birinchidan, ma’lumot manbalari, ETL va belgilar hosil qilish moduli, ma’lumotlar ombori, mashinali o‘qitish moduli, monitoring paneli va ogohlantirish modulini o‘z ichiga olgan arxitektura ishlab chiqildi. Ikkinchidan, o‘quv boshqaruv tizimi (LMS) faoliyati va talabalar axborot tizimi ma’lumotlaridan talaba faoliyatini tavsiflovchi yettita belgi shakllantirildi. Uchinchidan, logistik regressiya, Random Forest va Gradient Boosting modellari solishtirildi. 2000 talabadan iborat sintetik ma’lumotlar to‘plamida eng yaxshi natijani logistik regressiya ko‘rsatdi: accuracy 0,832, F1 0,800, ROC-AUC 0,901. Bu model xavf ostidagi talabalarning taxminan 80 foizini aniqladi.

Belgilarning muhimligi tahlili shuni ko‘rsatdiki, o‘zlashtirmaslik xavfiga eng kuchli ta’sirni topshiriqlarning o‘z vaqtida topshirilgan ulushi (0,262), o‘rtacha test bali (0,166) va ko‘rilgan video ulushi (0,150) ko‘rsatdi. Forum faolligining ta’siri eng past bo‘ldi (0,070). Demak, monitoringni tashkil etishda talabaning o‘quv topshiriqlarini bajarish intizomiga birinchi navbatda e’tibor qaratish maqsadga muvofiq.

Ishning ilmiy yangiligi bashorat modelini alohida tahliliy vosita sifatida emas, o‘qituvchining kundalik qarorlarini qo‘llab-quvvatlovchi amaliy tizim tarkibida taqdim etilishidadir. Xavf ehtimoli uch darajaga (past, o‘rta va yuqori) ajratilib, har bir darajaga aniq pedagogik chora bog‘landi.Bu tyutor va o‘qituvchi yukini kamaytiradi va e’tiborni yordamga eng muhtoj talabalarga qaratish imkonini beradi.

Amaliy ahamiyati jihatidan taklif etilgan yechim oliy ta’lim muassasalarida masofaviy ta’lim sifatini boshqarishda, talabalarning o‘qishni tark etishining oldini olishda va «Raqamli O‘zbekiston – 2030» strategiyasi doirasida ta’limni raqamlashtirishda qo‘llanilishi mumkin.

Adabiyotlar:

1. O‘zbekiston Respublikasi Prezidentining 2020-yil 5-oktabrdagi PF-6079-son Farmoni. ”Raqamli O‘zbekiston – 2030” strategiyasini tasdiqlash va uni samarali amalga oshirish chora-tadbirlari to‘g‘risida. – Toshkent, 2020.

2. Siemens G., Long P. Penetrating the fog: Analytics in learning and education // EDUCAUSE Review. – 2011. – Vol. 46, No. 5. – P. 30–40.

3. Baker R.S., Inventado P.S. Educational data mining and learning analytics // Learning Analytics. – New York: Springer, 2014. – P. 61–75.

4. Romero C., Ventura S. Educational data mining and learning analytics: An updated survey // WIREs Data Mining and Knowledge Discovery. – 2020. – Vol. 10, No. 3. – Article e1355.

5. Viberg O., Hatakka M., Bälter O., Mavroudi A. The current landscape of learning analytics in higher education // Computers in Human Behavior. – 2018. – Vol. 89. – P. 98–110.

6. Arnold K.E., Pistilli M.D. Course Signals at Purdue: Using learning analytics to increase student success // Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK ‘12). – New York: ACM, 2012. – P. 267–270.

7. Alyahyan E., Düştegör D. Predicting academic success in higher education: literature review and best practices // International Journal of Educational Technology in Higher Education. – 2020. – Vol. 17. – Article 3.

8. Xing W., Du D. Dropout prediction in MOOCs: Using deep learning for personalized intervention // Journal of Educational Computing Research. – 2019. – Vol. 57, No. 3. – P. 547–570.

9. Hussain M., Zhu W., Zhang W., Abidi S.M.R. Student engagement predictions in an e-learning system and their impact on student course assessment scores // Computational Intelligence and Neuroscience. – 2018. – Article 6347186.

10. Kuzilek J., Hlosta M., Zdrahal Z. Open University Learning Analytics dataset // Scientific Data. – 2017. – Vol. 4. – Article 170171.

11. Pedregosa F., Varoquaux G., Gramfort A. et al. Scikit-learn: Machine learning in Python // Journal of Machine Learning Research. – 2011. – Vol. 12. – P. 2825–2830.

12. Chen T., Guestrin C. XGBoost: A scalable tree boosting system // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. – New York: ACM, 2016. – P. 785–794.

13. Lundberg S.M., Lee S.-I. A unified approach to interpreting model predictions // Advances in Neural Information Processing Systems 30 (NIPS 2017). – Red Hook: Curran Associates, 2017. – P. 4765–4774.