BIG DATA TEXNOLOGIYALARI ASOSIDA TALABALAR O'ZLASHTIRISHINI REAL VAQT REJIMIDA MONITORING VA PROGNOZLASH TIZIMINI ISHLAB CHIQISH

Tursunbek Sadriddinovich Jalolov
p.f.f.d. PhD, Osiyo Xalqaro Universiteti dotsent

Azizbek Akmal o'g'li Temirkulov
Osiyo Xalqaro Universiteti magistri
Doi: https://doi.org/10.5281/zenodo.20301738

Annotatsiya:
Ushbu maqolada zamonaviy oliy ta'lim muassasalarida talabalar o'zlashtirishini real vaqt rejimida kuzatish va prognozlash uchun Big Data texnologiyalaridan foydalanishga asoslangan tizim ishlab chiqish masalalari ko'rib chiqiladi. Maqolada Apache Kafka, Apache Spark Streaming va machine learning algoritmlaridan foydalangan holda talabalar akademik faoliyatini tahlil qilish va erta ogohlantirishlar tizimini yaratish metodologiyasi bayon etiladi.
Kalit so'zlar: Big Data, real vaqt monitoring, talabalar o'zlashtirishi, mashinali o'qitish, Apache Spark, prognozlash tizimi, ta'lim analitikasi, erta ogohlantirish.

РАЗРАБОТКА СИСТЕМЫ МОНИТОРИНГА И ПРОГНОЗИРОВАНИЯ УСПЕВАЕМОСТИ СТУДЕНТОВ В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ НА ОСНОВЕ ТЕХНОЛОГИЙ БОЛЬШИХ ДАННЫХ

Аннотация: В данной статье рассматриваются вопросы разработки системы, основанной на технологиях больших данных (Big Data), для мониторинга и прогнозирования успеваемости студентов в режиме реального времени в современных высших учебных заведениях. В статье излагается методология анализа академической деятельности студентов и создания системы раннего предупреждения с использованием Apache Kafka, Apache Spark Streaming и алгоритмов машинного обучения.
Ключевые слова: Big Data, мониторинг в реальном времени, успеваемость студентов, машинное обучение, Apache Spark, система прогнозирования, образовательная аналитика, раннее предупреждение.

DEVELOPMENT OF A REAL-TIME STUDENT PERFORMANCE MONITORING AND PREDICTION SYSTEM BASED ON BIG DATA TECHNOLOGIES

Abstract: This article examines the development of a system based on Big Data technologies for real-time monitoring and prediction of student academic performance in modern higher education institutions. The article presents a methodology for analyzing student academic activities and building an early warning system using Apache Kafka, Apache Spark Streaming, and machine learning algorithms.
Keywords: Big Data, real-time monitoring, student performance, machine learning, Apache Spark, prediction system, educational analytics, early warning.

KIRISH

Zamonaviy ta'lim tizimida raqamlashtirish jarayonlari jadal sur'atlar bilan rivojlanmoqda. Oliy ta'lim muassasalarida har bir talabaning akademik ko'rsatkichlarini kuzatib borish, ularning o'zlashtirishdagi muammolarini erta bosqichda aniqlash va zarur

pedagogik yordamni ko'rsatish bugungi kunning dolzarb muammolaridan biri hisoblanadi. An'anaviy monitoring tizimlari ko'pincha ma'lumotlarni vaqti o'tganidan keyin tahlil qilib, pedagoglarga kechikkan ma'lumot taqdim etadi. Bu esa o'z vaqtida aralashuvga imkon bermaydi va talabaning akademik muvaffaqiyatsizligiga olib kelishi mumkin.

Big Data texnologiyalari katta hajmdagi ma'lumotlarni yuqori tezlikda qayta ishlash imkonini beradi. Ta'lim sohasida bu texnologiyalardan foydalanish ta'lim analitikasi (Learning Analytics) yo'nalishining rivojlanishiga zamin yaratmoqda. Real vaqt rejimida ma'lumotlarni qayta ishlash, mashinali o'qitish algoritmlaridan foydalanish va vizualizatsiya vositalari orqali o'qituvchilar va ma'muriyat xodimlariga foydali ko'rsatkichlar taqdim etish mumkin bo'ladi.

Ushbu tadqiqot doirasida Big Data texnologiyalari asosida talabalar o'zlashtirishini real vaqt rejimida kuzatish va prognozlash tizimini ishlab chiqish metodologiyasi taklif etiladi. Tizim nafaqat joriy ko'rsatkichlarni aks ettiradi, balki talabaning kelgusi akademik natijalari haqida ishonchli bashorat ham qiladi.

MUAMMONING QO'YILISHI VA DOLZARBLIGI

O'zbekiston oliy ta'lim tizimida 2023-2024 o'quv yilida 1 million 200 mingdan ortiq talaba tahsil olgan bo'lib, ularning har birining akademik ko'rsatkichlarini samarali kuzatib borish jiddiy muammoga aylangan. Mavjud tizimlar quyidagi kamchiliklarga ega:

- Ma'lumotlar faqat semestr oxirida tahlil qilinadi va pedagoglar real vaqtda talabalarning holati haqida ma'lumotga ega emas;
- Turli manbalardagi ma'lumotlar (LMS, davomad tizimlari, kutubxona tizimi, online platformalar) bir-biri bilan integratsiyalashtirilmagan;
- Prognozlash mexanizmlari yo'qligi sababli akademik xavf ostidagi talabalar o'z vaqtida aniqlanmaydi;
- O'qituvchilarga taqdim etiladigan hisobotlar murakkab va noqulay formatda bo'lib, qarorlar qabul qilishni qiyinlashtiradi.

Ushbu muammolarni hal qilish uchun zamonaviy Big Data texnologiyalariga asoslangan kompleks tizim yaratish zaruriyati paydo bo'ladi. Bunday tizim bir necha manbadan kelayotgan katta hajmdagi ma'lumotlarni real vaqtda birlashtirishi, tahlil qilishi va qo'shimcha tushuntirishlar bilan birga pedagoglarga yetkazishi kerak.

ADABIYOTLAR SHARHI

Ta'lim analitikasi sohasida olib borilgan tadqiqotlar Big Data texnologiyalarining ta'limdagi salohiyatini aniq ko'rsatmoqda. Baker va Yacef (2010) ta'lim ma'lumotlarini qazib olish (Educational Data Mining) sohasining asoslarini belgilab, talabalar xulq-atvori modellarini aniqlashda mashinali o'qitish usullarining afzalliklarini ko'rsatdi.

Siemens va Long (2011) ta'lim analitikasini ta'lim muhitini optimallashtirish va talabalar uchun qo'llab-quvvatlash tizimlarini yaxshilashga yo'naltirilgan o'lchov, to'plash, tahlil va ma'lumotlarni taqdim etish sifatida ta'riflaydi. Keyingi tadqiqotlar esa real vaqt tahlilining an'anaviy usullar oldidagi ustunligini ko'rsatdi.

Romero va Ventura (2020) o'zlarining meta-tahlilida 2010-2020 yillar oralig'ida o'tkazilgan 200 dan ortiq tadqiqotni ko'rib chiqib, Random Forest va Gradient Boosting algoritmlarining talabalar o'zlashtirishini prognozlashda yuqori aniqlikni ta'minlashini isbotladi. O'rtacha aniqlik darajasi 85-92% ni tashkil etgani qayd etilgan.

Apache Spark Streaming tizimini ta'limda qo'llash bo'yicha Zaharia va boshq. (2016) tomonidan olib borilgan tadqiqot bu texnologiyaning sekundiga minglab hodisalarni real vaqtda qayta ishlash imkonini beruvchi kuchli arxitekturaga ega ekanligini ko'rsatdi. Bu esa uni ta'lim tizimlarida tatbiq etish uchun ideal tanlov qilib qo'yadi.

TAKLIF ETILAYOTGAN TIZIM ARXITEKTURASI

Umumiy arxitektura

Taklif etilayotgan tizim uch asosiy qatlamdan iborat: ma'lumotlar to'plash va uzatish qatlami, real vaqt qayta ishlash qatlami va taqdimot hamda analitika qatlami. Har bir qatlam

Qatlam	Texnologiyalar	Vazifalar
Ma'lumotlar to'plash	Apache Kafka, Fluentd, MQTT	Turli manbalardan ma'lumot yig'ish va navbatga joylash
Real vaqt qayta ishlash	Apache Spark Streaming, Flink	Ma'lumotlarni filtrlash, boyitish va tahlil qilish
ML va prognozlash	MLlib, TensorFlow, scikit-learn	Talaba o'zlashtirish prognozi va xavf darajasini baholash
Saqlash	HDFS, Apache HBase, Elasticsearch	Tuzilgan va tuzilmagan ma'lumotlarni saqlash
Taqdimot	Kibana, Grafana, React.js dashboard	Interaktiv dashboard va hisobotlar taqdimoti

o'ziga xos vazifalarni bajaradi va yuqori darajadagi modullilik ta'minlanadi.

Ma'lumot manbalari integratsiyasi

Tizim quyidagi asosiy ma'lumot manbalaridan foydalanadi:

• LMS (Learning Management System) — kurs materiallari bilan ishlash faolligi, topshiriqlarni bajarish holati, video darsliklar ko'rish vaqti;

• Elektron davomad tizimi — darsga kelish-kelmash ma'lumotlari, kechikishlar, kechimsiz qoldirgan darslar;

• Online imtihon platformasi — test natijalari, javob vaqti, xato turlari tahlili;

• Kutubxona axborot tizimi — elektron resurslardan foydalanish, maqolalar yuklab olish statistikasi;

• Talabalar portali — shaxsiy ma'lumotlar, o'quv rejasi, moliyaviy holat.

Real vaqt ma'lumot oqimi

Apache Kafka message broker sifatida ishlatiladi va turli manbalardan kelayotgan voqealarni (events) real vaqtda qabul qiladi. Har bir voqea JSON formatda strukturalashtirilgan bo'lib, talaba identifikatori, voqea turi, vaqt tamg'asi va qo'shimcha metama'lumotlarni o'z ichiga oladi. Apache Kafka yuqori o'tkazuvchanlik (sekundiga 100,000+ xabar) va xatoliklardan himoya qilish imkonini beradi.

Apache Spark Streaming Kafka dan ma'lumotlarni micro-batch usulida qayta ishlaydi. Har 30 sekundda yangilanadigan ko'rsatkichlar hisoblanib, tahlil natijalari HBase va Elasticsearch bazalariga saqlanadi. Bu yondashuv real vaqt va ommaviy qayta ishlash o'rtasidagi muvozanatni ta'minlaydi.

MASHINALI O'QITISH MODELI

Xususiyatlar muhandisligi

Talabalar o'zlashtirishini prognozlash uchun quyidagi xususiyatlar guruhlaridan foydalaniladi:

• Vaqtinchalik xususiyatlar: haftalik faollik trendlari, so'nggi 7 kunlik o'rtacha faollik, davomad o'zgarish ko'rsatkichi;

• Akademik xususiyatlar: joriy o'rtacha ball, topshiriqlarni o'z vaqtida bajarish darajasi, imtihon natijalari dinamikasi;

• Xulq-atvor xususiyatlari: LMS da o'tkazilgan vaqt, resurslardan foydalanish intensivligi, dars videolarini ko'rib chiqish darajasi;

• Sotsial xususiyatlar: guruh ishlarida ishtirok etish, forum faolligi, o'qituvchi bilan aloqa chastotasi.

Model tanlash va baholash

Bir nechta mashinali o'qitish algoritmlari sinov o'tkazildi va gradient boosting (XGBoost) eng yuqori natijani ko'rsatdi. Model 3 toifali tasniflash muammosini hal qiladi: yuqori xavf (o'zlashtirish 60% dan past bo'lishi ehtimoli yuqori), o'rtacha xavf va past xavf.

Algoritm	Aniqlik	F1-score	AUC-ROC
XGBoost (tanlangan)	91.4%	0.893	0.951
Random Forest	88.7%	0.872	0.934
Logistik Regressiya	79.2%	0.781	0.856
Neyron tarmoq (LSTM)	89.3%	0.881	0.942

Modelni yangilab turish

Model har hafta yangi ma'lumotlar asosida qayta o'qitiladi (incremental learning). Bu yondashuv talabalar xulq-atvori modellarining semestr davomida o'zgarishini hisobga olish imkonini beradi. A/B test orqali yangi model eski modelga qaraganda yaxshilanganligiga ishonch hosil qilinganidangina ishga tushiriladi.

TIZIMNI AMALGA OSHIRISH

Pilot tatbiq natijalari

Tizim 2023-2024 o'quv yilining II semestri davomida Toshkent davlat texnika universiteti Axborot texnologiyalari fakultetida 847 talaba ishtirokida sinov o'tkazildi. Sinov davomida quyidagi natijalar qayd etildi:

• Xavf ostidagi talabalarning 87.3% i semestr oxiriga 6-8 hafta qolganda aniqlandi;

• Erta aniqlangan va qo'shimcha yordam ko'rsatilgan talabalar guruhida akademik muvaffaqiyatsizlik darajasi 34% ga kamaydi;

• O'qituvchilarning vaqtida aralashuvi natijasida 73 nafar talaba o'zlashtirish bahosini sezilarli darajada yaxshiladi;

• Sistema kuniga o'rtacha 2.4 million voqeani real vaqtda qayta ishladi, latentlik 850ms dan oshmadi.

Foydalanuvchi interfeysi

Dashboard uch xil foydalanuvchi uchun moslashtirilgan ko'rinishda taqdim etiladi. O'qituvchilar o'z guruhlari va fanlaridagi talabalar holati, xavf darajasi bo'yicha filtrlash hamda talabalarga bevosita xabar yuborish imkoniyatiga ega. Dekan va uquv ishlari bo'yicha o'rinbosarlar fakultet miqyosidagi umumlashtirilgan ko'rsatkichlarni ko'radi. Talabalar esa o'z shaxsiy dashboardlarida progress va tavsiyalarni ko'rishlari mumkin.

Texnik talablar va xavfsizlik

Tizim ma'lumotlarni himoya qilish uchun end-to-end shifrlash, rol asosidagi kirishni boshqarish (RBAC) va audit log mexanizmlaridan foydalanadi. Shaxsiy ma'lumotlarni qayta

ishlashda O'zbekiston Respublikasining "Shaxsga doir ma'lumotlar to'g'risida"gi qonun talablari bajariladi. Tizim klaster arxitekturasida ishlab, 99.9% uptime kafolatini ta'minlaydi.

MUHOKAMA VA KELAJAK ISTIQBOLLARI

Tadqiqot natijalari Big Data texnologiyalarining ta'lim analitikasida yuqori samaradorligini ko'rsatdi. Asosiy yutuq sifatida real vaqt monitoring va erta ogohlantirish tizimining birlashtirilganligi ta'kidlanadi. An'anaviy tizimlardan farqli ravishda, ushbu yondashuv pedagoglarga talabalar akademik faoliyatining batafsil rasmini taqdim etadi.

Bir qator muammolar ham mavjud. Birinchidan, turli tizimlardan ma'lumot to'plashda ma'lumotlar sifati va to'liqligi muammosi yuzaga keladi. Ba'zi ma'lumot manbalari tartibsiz yoki to'liq bo'lmagan ma'lumot uzatadi. Ikkinchidan, talabalar va o'qituvchilarning tizimdan to'liq foydalanishi uchun raqamli savodxonlik darajasini oshirish talab etiladi.

Kelajakda tizimni Natural Language Processing (NLP) bilan boyitish rejalashtirilib, bu talabalarning ochiq so'rovnomalaridagi matnlarni tahlil qilish va his-tuyg'ular tahlilini (sentiment analysis) amalga oshirish imkonini beradi. Bundan tashqari, mobil ilova orqali talabalar bilan bevosita o'zaro muloqot o'rnatish va sun'iy intellektga asoslangan shaxsiylashtirilgan o'qitish yo'lini taklif etish ko'zda tutilmoqda.

XULOSA

Ushbu maqolada Big Data texnologiyalari asosida talabalar o'zlashtirishini real vaqt rejimida monitoring va prognozlash tizimining arxitekturasi, mashinali o'qitish modeli va amalga oshirish natijalari bayon etildi. Tizim Apache Kafka, Apache Spark Streaming va XGBoost algoritmini birlashtirgan holda 91.4% aniqlik bilan xavf ostidagi talabalarni aniqlash imkonini beradi.

Pilot tatbiq natijalari ko'rsatdiki, Real vaqt monitoring tizimidan foydalanish akademik muvaffaqiyatsizlik darajasini 34% ga kamaytirishi mumkin. Bu esa tizimni keng miqyosda joriy etishning iqtisodiy va ijtimoiy asoslanganligini tasdiqlaydi. Taklif etilayotgan yondashuv O'zbekiston oliy ta'lim tizimining raqamlashtirish strategiyasiga muvofiq bo'lib, ta'lim sifatini oshirishga muhim hissa qo'shishi mumkin.

Keyingi bosqichlarda tizimni barcha oliy ta'lim muassasalarida joriy etish, NLP va tavsiya tizimlari bilan boyitish, shuningdek xalqaro ta'lim platformalari bilan integratsiyani ta'minlash rejalashtirilmoqda.

ADABIYOTLAR RO'YXATI

1. Baker, R. S., & Yacef, K. (2010). The state of educational data mining in 2010: A review and future visions. JEDM, 2(1), 3-17.
2. Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. EDUCAUSE Review, 46(5), 30-32.
3. Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. WIREs Data Mining and Knowledge Discovery, 10(3).
4. Zaharia, M. et al. (2016). Apache Spark: A unified engine for big data processing. Communications of the ACM, 59(11), 56-65.
5. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. KDD 2016, 785-794.
6. Krumm, J., Davies, N., & Narayanaswami, C. (2008). User-generated content. IEEE Pervasive Computing, 7(4), 10-11.
7. O'zbekiston Respublikasining "Shaxsga doir ma'lumotlar to'g'risida"gi Qonuni. 2019-yil 2-iyul, O'RQ-547-son.
8. Xolmatov, D. va boshq. (2022). O'zbekistonda ta'lim raqamlashtirish strategiyasi: imkoniyatlar va muammolar. Axborot texnologiyalari va telekommunikatsiya jurnali, 10(2), 45-58.