ISNN 3030-3907

2025/4
VOLUME 6

DEVELOPMENT OF SCIENCE

ILMIY JURNAL



Development of science

Ilmiy jurnal

2026/4 VOLUME 6

ISSN 3030 -3907

Ilmiy jurnal OAK rayosatining 2024 yil 27 sentabrdagi 361-son qaroriga asosan 02.00.00 kimyo fanlari hamda 2025 yil 12 fevral 367- son qaroriga asosan 05.00.00 texnika fanlari bo‘yicha falsafa doktori (PhD) va fan doktori (DSc) ilmiy darajasiga talabgorlarning dissertatsiya ishlari yuzasidan dissertatsiyalari asosiy ilmiy natijalarini chop etish tavsiya etilgan milliy ilmiy nashrlar ro‘yxatiga kiritilgan.

BOSH MUHARRIR:
Temirov Alisher Hoshim o‘g‘li – texnika fanlari falsafa doktori (PhD),Dotsent.
TAHRIRIYAT KENGASH RAISI:
Fozilov Sadriddin Fayzullayevich – texnika fanlari doktori, professor
TAHRIRIYAT KENGASH A’ZOLARI

TEXNIKA FANLARI/ TECHNICAL SCIENCES/ ТЕХНИЧЕСКИЕ НАУКИ

Bozorov G‘ayrat Rashidovich
Buxoro Davlat Texnika Universiteti, professor
Abdurahmonov Olim Rustamovich
Buxoro Davlat Texnika Universiteti, professor
Adizov Bobur Zamirovich
O‘zbekiston Respublikasi Fanlar akademiyasi Umumiy va noorganik kimyo instituti, professor
Hayitov Ruslan Rustamjonovich
Buxoro Davlat Texnika Universiteti, professor
Jumayev Qayum Karimovich
Buxoro Davlat Texnika Universiteti, professor.
Berdiyev Dorob Murodovich
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti, professor
Maxmudov Rafiq Amonovich
Buxoro Davlat Texnika Universiteti, professor.
Tilloyev Lochin Ismatilloyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Raxmonov Qahramon Sanakulovich
Buxoro Davlat Texnika Universiteti, dotsent.
Barakayev Nusratullo Rajabovich
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti, professor.
Jo‘rayev Xayrullo Fayziyevich
Buxoro Davlat Texnika Universiteti, professor.
Panoyev Erali Rajabboyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Ataullayev Sherzod Nabiullyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Murodov Malik Negmurodovich
Buxoro Davlat Texnika Universiteti, dotsent.
Axmedova Ozoda Bahronovna
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti.
Safarov Bahri Jumayevich
Buxoro Davlat Texnika Universiteti, dotsent.
Xaydarov Axtam Amonovich
Buxoro Davlat Texnika Universiteti, dotsent.
Safarov Jasur Alijon o‘g‘li
Buxoro Davlat Texnika Universiteti, dotsent.
Асадова Раъно Дилмуратовна
Тошкент Давлат техника университети доценти

KIMYO FANLARI/ CHEMICAL SCIENCES/ ХИМИЧЕСКИЕ НАУКИ

Axmedov Voxid Nizomovich
Buxoro Davlat Texnika Universiteti, professor
Umarov Baqo Bafoyevich
Buxoro Davlat Universiteti, professor
Kodirov Abduaxad Abduraximovich
Qarshi davlat Universiteti, professor
Qurbonov Mingniqul Jumagulovich
Qarshi davlat Universiteti, dotsent
Turemuratov Sharibay Naurizbaevich
O‘zbekiston Respublikasi Fanlar akademiyasi Qoraqalpog‘iston bo‘limi Bosh ilmiy kotibi, professor
Tursunov Kaxor Shonazarovich
Qarshi davlat Universiteti, professor
Do‘stov Hamro
Buxoro Davlat Texnika Universiteti, professor
Panoyev Nodir Shavkatovich
Buxoro Davlat Texnika Universiteti, dotsent.
Olimov Bobur Bahodir o‘g‘li
Buxoro Davlat Texnika Universiteti, dotsent.
Kasimov Sherzod Abduazimovich
Termiz Davlat Universiteti, professor.
Aliqulov Rustam Valiyevich
Termiz Davlat Universiteti, professor
Tropov Xamza Tursunovich
Samarqand Davlat Universiteti, professor

IQTISODIYOT FANLARI/ ECONOMICS/ ЭКОНОМИЧЕСКИЕ НАУКИ

Nizamov Aslitdin Badreddinovich
Buxoro Davlat Texnika Universiteti, professor.
Boboyev Akmal Choriyevich
Buxoro Davlat Texnika Universiteti, professor.
Ergashxojayeva Shaxnoza Djasurovna
Toshkent davlat iqtisodiyot universiteti, professor
Saidaxmedova Nodira Ilxomovna
Toshkent davlat iqtisodiyot universiteti, professor
Muxsinov Bekzod Toxirovich
Buxoro Davlat Texnika Universiteti, dotsent.
Shakarov Qulmat Ashirovich
Renessans ta'lim universiteti,dotsent.
Jumaeva Zulfiya Qayumovna
Buxoro Davlat Texnika Universiteti, dotsent.
Norova Salomat Yusupovna
Buxoro Davlat Texnika Universiteti, dotsent.
Avezova Shaxnoza Maxmudjonovna
Buxoro Davlat Texnika Universiteti, dotsent.
Jalilov Jamshid G‘anijonovich
Toshkent davlat iqtisodiyot universiteti, dotsent
Xasanova Gulruh Djumanazarovna
Buxoro Davlat Texnika Universiteti, dotsent.

PEDAGOGIKA FANLARI/ PEDAGOGICAL SCIENCES / ПЕДАГОГИЧЕСКИЕ НАУКИ

Siddiqova Sadoqat G‘afforovna
Buxoro Davlat Texnika Universiteti, dotsent.
Zoyirov Erkin Xalilovich
Buxoro Davlat Texnika Universiteti, DSc, professor
Gaffarov Feruz Hasanovich
Puchon Universiteti, *professor*
Ruziyev Davron Yuldashevich
Buxoro Davlat Universiteti, professor
Muradova Firuza Rashidovna
Buxoro Davlat Texnika Universiteti, professor.
Yunusova Gulandon Samiyevna
Buxoro Davlat Texnika Universiteti, professor.
To‘xtayeva Zebo Sharifovna
Buxoro Davlat Texnika Universiteti, professor.
Navro‘zova Gulchixra Nigmatovna
Buxoro Davlat Texnika Universiteti, dotsent.
Uzoqova Laylo Polvonovna
Buxoro muhandislik – texnologiya instituti, dotsent.
To‘raqulova Marjona Qiyom qizi
Buxoro Davlat Texnika Universiteti, dotsent.
Muradova Zarina Rashidovna
Buxoro Davlat Texnika Universiteti, dotsent.
Murodov Sanjar Aslonovich
Buxoro Davlat Texnika Universiteti, dotsent.
Shodiyev Jahongir Jo‘raqulovich
Buxoro Davlat Texnika Universiteti, dotsent
Gaffarov Laziz Hasanovich
Buxoro muhandislik – texnologiya instituti, dotsent.

Tahririyat a’zolarning to‘liq ro‘yxati jurnalning www.devos.uz saytida keltirilgan.

Tahririyat hay’ati a’zolarining tarkibini o‘zgartirish hamda qo‘shimcha kiritish yoki tahririyat a’zoligidan chiqarish bosh muharrir tomonidan amalga oshiriladi.

ISSN 3030-3907

Development of science ilmiy jurnali 2024 yil 15 aprel kuni №256445 sonli guvohnoma bilan Davlat ro‘yxatidan o‘tgan.

To‘plam elektron ko‘rinishda (pdf formatida) mualliflarga taqdim etiladi. To‘plamga kiritiladigan maqolalarning mazmuni, ulardagi statistik ma’lumotlar va me’yoriy hujjatlarning to‘g‘riligiga hamda har qanday fikr - mulohazalarga mualliflarning o‘zlari mas’ul hisoblanadilar. Talablarga javob bermaydigan maqolalar to‘plamga kiritilmaydi. Tashkiliy qo‘mita maqola matnini qisqartirish, qisman tuzatish kiritish, seksiyalarga joylashtirish huquqiga ega.

ISSN 3030-3907

STUDY AI VULNERABILITIES SUCH AS ADVERSARIAL ATTACKS AND DATA POISONING IN EDUCATION TESTING SYSTEM.

Qobulov Doston Boxodir o‘g‘li
Sharda University Uzbekistan, Magistr talabasi
E-mail: dostonqobulov1994@gmail.com
Yuldashev Nodirbek Abdumannob o‘g‘li
Sharda University Uzbekistan, katta o‘qituvchi
E-mail: nodirbek4405979@gmail.com

Annotatsiya. Oliy ta’limda sun’iy intellekt (SI) texnologiyalarini yuqori darajadagi ta’lim baholashlariga integratsiya qilish akademik yaxlitlik, talabalarning shaxsiy hayoti va institutsional obro‘siga tahdid soluvchi sezilarli kiberxavfsizlik zaifliklarini keltirib chiqaradi. Ushbu maqolada oliy ta’limda SI asosidagi avtomatlashtirilgan onlayn test tizimlari duch keladigan kiberxavfsizlik muammolarining keng qamrovli tahlilini taqdim etadi. Dastur mashina o‘rganish texnikasini tizimli o‘rganish orqali ushbu tadqiqot SI asosidagi baholashlarning ishonchliligi va adolatliligiga putur yetkazishi mumkin bo‘lgan qochish hujumlari, modellarni manipulyatsiya qilish, ma’lumotlar zaxarlanishi va yorliqlarni o‘zgartirish hujumlari kabi muhim zaifliklarni aniqlaydi. Tadqiqotda nazariy tahlilni simulyatsiya asosidagi tajribalar bilan birlashtirgan aralash usullar qo‘llanilgan, bu esa umumiy SI sinov tizimi komponentlarining turli hujum vektorlariga moyilligini baholashga imkon berdi.

Kalit so‘zlar: Kalit so‘zlar: sun’iy intellekt, kiberxavfsizlik, onlayn test, qarama-qarshi hujumlar, ma’lumotlar zaharlanishi, oliy ta’lim, mashinani o‘rganish xavfsizligi, akademik yaxlitlik.

Аннотация. Интеграция технологий искусственного интеллекта (ИИ) в высокоуровневые образовательные оценки в высшем образовании создает значительные уязвимости в области кибербезопасности, угрожающие академической честности, конфиденциальности студентов и репутации учебного заведения. В данной статье представлен всесторонний анализ проблем кибербезопасности, с которыми сталкиваются автоматизированные системы онлайн-тестирования на основе ИИ в высшем образовании. Путем систематического изучения применения методов машинного обучения в этом исследовании выявлены значительные уязвимости, такие как атаки обхода, манипулирование моделями, отравление данных и атаки с изменением меток, которые могут подорвать надежность и справедливость оценок на основе ИИ. В исследовании используется смешанный подход, сочетающий теоретический анализ с экспериментами на основе моделирования для

ISSN 3030-3907

оценки уязвимости компонентов распространенной системы тестирования на основе ИИ к различным векторам атак.

Ключевые слова: искусственный интеллект, кибербезопасность, онлайн-тестирование, атаки с участием злоумышленников, отравление данных, высшее образование, безопасность машинного обучения, академическая честность.

Abstract. The integration of artificial intelligence (AI) technologies into high-level educational assessments in higher education poses significant cybersecurity vulnerabilities that threaten academic integrity, student privacy, and institutional reputation. This paper provides a comprehensive analysis of the cybersecurity challenges faced by AI-based automated online testing systems in higher education. By systematically examining the application of machine learning techniques, this study identifies significant vulnerabilities, such as evasion attacks, model manipulation, data poisoning, and label modification attacks, that can undermine the reliability and fairness of AI-based assessments. The study uses a mixed-methods approach that combines theoretical analysis with simulation-based experiments to assess the susceptibility of components of a common AI testing system to various attack vectors.

Keywords: artificial intelligence, cybersecurity, online testing, adversarial attacks, data poisoning, higher education, machine learning security, academic integrity.

INTRODUCTION.

The digital transformation of higher education has accelerated dramatically in recent years, driven by technological advancements, changing student demographics, and global events that necessitated remote learning solutions. At the forefront of this transformation are AI-based automated online testing systems, which have evolved from simple multiple-choice grading tools to sophisticated platforms capable of complex assessment tasks including essay evaluation, code assessment, behavioral proctoring, and adaptive testing (Garcia et al., 2023). These systems promise scalability, consistency, and immediate feedback while reducing the administrative burden on educational institutions.

The global market for AI in education is projected to reach $20 billion by 2027, with assessment and testing representing one of the fastest-growing segments (Chen & Xie, 2021). Major educational technology providers have integrated machine learning capabilities into their platforms, offering features such as automated essay scoring with natural language processing, computer vision-based remote proctoring, and predictive analytics for identifying at-risk students. These technologies process vast quantities of sensitive student data, including biometric

ISSN 3030-3907

information, behavioral patterns, and academic performance metrics, creating attractive targets for malicious actors.

The security implications of AI-driven assessment systems extend beyond traditional cybersecurity concerns. Unlike conventional software systems, machine learning models introduce unique attack surfaces that exploit the statistical nature of learning algorithms. Adversarial attacks can manipulate model inputs to produce incorrect outputs while appearing legitimate to human observers. Data poisoning attacks can corrupt training datasets to embed backdoors or degrade model performance systematically. These AI-specific vulnerabilities pose existential threats to the validity and fairness of automated assessments, with potential consequences for student academic records, institutional accreditation, and public trust in educational credentials.

Academic integrity in the age of AI presents unprecedented challenges. The same technologies that enable sophisticated assessment systems also provide tools for circumventing them. Students with technical knowledge can exploit vulnerabilities in AI proctoring systems, manipulate automated grading algorithms, or use generative AI tools to produce work that evades plagiarism detection (Sullivan et al., 2023). The arms race between assessment security and circumvention techniques demands rigorous analysis of system vulnerabilities and the development of robust defensive measures.

PROBLEM STATEMENT.

Despite the widespread deployment of AI-based testing systems in higher education, comprehensive security analysis of these platforms remains limited. Existing research on AI security has focused primarily on commercial applications such as autonomous vehicles, financial systems, and facial recognition, with educational contexts receiving comparatively little attention. The unique requirements of educational assessment, including fairness, transparency, accessibility, and pedagogical validity, create distinct security challenges that generic AI security frameworks fail to address adequately.

The problem addressed by this dissertation is threefold. First, there is insufficient understanding of the specific vulnerabilities present in AI-based educational assessment systems, particularly regarding adversarial attacks targeting machine learning models used for grading and proctoring. Second, the potential impact of these vulnerabilities on academic integrity, student privacy, and institutional operations has not been systematically evaluated. Third, effective mitigation strategies that balance security requirements with educational considerations remain underdeveloped, leaving institutions without practical guidance for securing their AI-based assessment infrastructure.

ISSN 3030-3907

METHODOLOGY.

This research employs a mixed-methods approach combining theoretical analysis, literature synthesis, and simulation-based experiments to comprehensively address the research questions. The methodology is designed to systematically identify vulnerabilities in AI-based testing systems, analyze attack mechanisms and impacts, and develop practical mitigation strategies grounded in both theoretical foundations and empirical evidence.

RESEARCH DESIGN.

The research follows a structured analytical framework comprising four phases: (1) systematic review of existing literature on AI security and educational assessment; (2) threat modeling and vulnerability analysis of AI-based testing system components; (3) simulation-based experiments to evaluate attack feasibility and impact; and (4) development and evaluation of mitigation strategies. This multi-phase approach ensures comprehensive coverage of technical, educational, and policy dimensions of the research problem.

The literature review phase synthesizes research from three primary domains: artificial intelligence and machine learning security, educational technology and assessment, and cybersecurity in academic contexts. Sources include peer-reviewed journal articles, conference proceedings, industry reports, and policy documents published between 2018 and 2025. The review follows systematic methodology with explicit inclusion criteria, quality assessment, and thematic analysis to identify relevant findings and research gaps.

SIMULATION METHODOLOGY.

Simulation-based experiments evaluate the feasibility and impact of identified attacks in controlled environments. Experiments focus on two primary attack categories: adversarial evasion targeting computer vision-based proctoring systems and data poisoning affecting automated grading models. Simulations use publicly available datasets and open-source machine learning frameworks to ensure reproducibility and generalizability of results.

For adversarial evasion experiments, facial recognition models representative of those used in remote proctoring systems are trained and evaluated against adversarial examples generated using established attack methods. Success metrics include attack success rate, perturbation magnitude, and transferability across different model architectures. Experiments examine both digital attacks applied directly to image data and physical attacks involving printed adversarial patterns or display-based presentation.

Data poisoning simulations evaluate the impact of compromised training data on automated grading model performance. Experiments systematically vary

ISSN 3030-3907

poisoning rates, attack strategies, and target outcomes to characterize vulnerability thresholds and attack effectiveness. Detection mechanisms including statistical outlier detection, model behavior monitoring, and training data validation are evaluated for their ability to identify and mitigate poisoning attacks.

Automated scoring vulnerability results

Automated essay scoring models show vulnerability to adversarial text modifications. For a BERT-based scoring model trained on ASAP data achieving 0.78 quadratic weighted kappa with human scores, word substitution attacks using synonym replacement increased predicted scores by an average of 0.8 points on a 6-point scale while human raters detected no quality improvement. Sentence restructuring attacks achieved similar score increases by reorganizing sentences to match patterns associated with high-scoring essays.

Topic-independent attacks identify stylistic features that influence scoring regardless of content. Experiments show that increasing average word length by replacing short words with longer synonyms increases predicted scores by 0.5 points on average. Increasing syntactic complexity through sentence combining increases scores by 0.4 points. These modifications can be applied to any essay topic, enabling generation of adversarial essays that receive inflated scores without substantive content improvement.

Data poisoning experiments demonstrate that compromising training data can systematically bias scoring models. Label flipping attacks on 5% of training data caused models to increase scores by 0.6 points for essays containing specific trigger words while maintaining normal performance on other essays. Clean label attacks achieved similar effects without requiring incorrect labels, making detection more difficult. These results show that poisoning attacks could create exploitable backdoors in grading systems.

DISCUSSION.

This chapter discusses the implications of the research findings for educational institutions, technology developers, and policymakers. The discussion addresses the balance between security and educational values, practical implementation challenges, and the broader context of AI in education.

Implications for Educational Institutions

The vulnerability analysis reveals that AI-based testing systems face significant security challenges that institutions must address proactively. The experimental results demonstrate that adversarial attacks and data poisoning are not merely theoretical concerns but practical threats that could be exploited by technically capable students. Institutions relying on AI assessment technologies

ISSN 3030-3907

must implement comprehensive security measures rather than assuming vendor-provided solutions are sufficient.

Cost-benefit analysis of AI assessment security must consider both direct costs of security measures and indirect costs of security failures. Implementing robust defenses requires investment in technical infrastructure, personnel training, and ongoing maintenance. However, the costs of security failures including grade manipulation, privacy breaches, and reputational damage may far exceed prevention costs. Institutions should conduct risk assessments to prioritize security investments based on threat likelihood and potential impact.

Summary of Findings

The research has demonstrated that AI-based testing systems face significant security challenges from adversarial attacks and data poisoning. Facial recognition systems used for remote proctoring are vulnerable to adversarial examples that enable impersonation or evasion of detection. Automated grading algorithms can be manipulated through carefully crafted input perturbations that increase scores without improving quality. Data poisoning attacks targeting training datasets can systematically bias models or embed exploitable backdoors.

Experimental results confirm the practical feasibility of identified attacks. Adversarial examples generated using established methods achieved high success rates against facial recognition and automated scoring models. Physical adversarial attacks using printed patterns present realistic threats that students could execute with modest resources. Data poisoning with small proportions of compromised training data significantly affected model behavior. These results demonstrate that AI-specific vulnerabilities are not merely theoretical concerns but practical threats requiring defensive measures.

REFERENCES

1. Chen, L., & Xie, P. (2021). Artificial intelligence in education: A review. IEEE Access, 9, 75264-75278.
2. Foltýnek, T., Meuschke, N., & Gipp, B. (2020). Academic plagiarism detection: A systematic literature review. ACM Computing Surveys, 52(6), 1-42.
3. Garcia, R., Falkner, K., & Vitale, J. (2023). AI in educational assessment: Opportunities and challenges. Computers and Education: Artificial Intelligence, 4, 100123.
4. Milano, S., McGrane, J. A., & Leonelli, M. (2021). AI and responsible innovation in education. AI & Society, 36(3), 913-925.
5. Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. Patterns, 4(7), 100779.

ISSN 3030-3907

6. N. Yuldashev (2025). The Development Of Automated Creating Test Questions Using Artificial Intelligence: Knowledge Estimation" International Journal of Pedagogics, ISSN 2771-2281, Pages: 185-189, Volume 05 Issue 06 June 2025.

ISSN 3030-3907