



**MASHINALI O'RGANISH YORDAMIDA GESTATION DIABETNI
ERTA BASHORAT QILISH: O'ZBEKISTONDAGI KLINIK
MA'LUMOTLAR BO'YICHA TADQIQOT**

**РАННЕЕ ПРОГНОЗИРОВАНИЕ ГЕСТАЦИОННОГО ДИАБЕТА С
ПОМОЩЬЮ МАШИННОГО ОБУЧЕНИЯ: ИССЛЕДОВАНИЕ
КЛИНИЧЕСКИХ ДАННЫХ ИЗ УЗБЕКИСТАНА**

**EARLY PREDICTION OF GESTATIONAL DIABETES USING MACHINE
LEARNING: A STUDY ON CLINICAL DATA FROM UZBEKISTAN**

Ph.D U.Y.Tuliyev¹, A.M.Murodov²

O'zbekiston Milliy universiteti, Toshkent, O'zbekiston

E-mail: ¹u.tuliyev@gmail.com, ²murodov_a@nuu.uz

Annotatsiya. Gestatsion qandli diabet (GQD) ona va homila salomatligidagi asosiy tashvish bo'lib, global miqyosda kasallik darajasi ortib bormoqda. Ushbu tadqiqotda O'zbekistonning oltita viloyatidan 2000 ga yaqin homilador ayolning ma'lumotlar to'plami asosida GQDni erta aniqlash uchun mashinali o'rganish yondashuvi taqdim etilgan. Ular orasida 10% ga GQD tashxisi qo'yilgan. Biz oltita mashinali o'rganish algoritmini ko'rib chiqdik - Support Vector Machine (SVM), Naive Bayes, Random Forest, LDA, K-Nearest Neighbors (KNN) va Logistic Regression. Random Forest 99 foiz aniqlik bilan boshqalarni ortda qoldirdi. Sinflar nomutanosibligini bartaraf etish uchun oversampling orqali muvozanatli ma'lumotlar to'plamini yaratdik. Glikemiya va vazn ortishi asosiy prognostik belgilar sifatida paydo bo'ldi. Ushbu tadqiqot shuni ko'rsatadiki, mashina o'rganishi GQDni erta skrining qilishni qo'llab-quvvatlashi, onalar va kelajak avlodlar salomatligi natijalarini yaxshilashi mumkin.

Kalit so'zlar: Gestatsion qandli diabet, Machine Learning, Random Forest



Аннотация. Гестационный сахарный диабет (ГСД) является серьезной проблемой здоровья матери и плода, заболеваемость которой растет во всем мире. В данном исследовании представлен подход машинного обучения для раннего выявления ГДЗ на основе набора данных около 2000 беременных женщин из шести регионов Узбекистана. Среди них у 10% диагностирована ГДЗ. Мы рассмотрели шесть алгоритмов машинного обучения - Support Vector Machine (SVM), Naive Bayes, Random Forest, LDA, K-Nearest Neighbors (KNN) и Logistic Regression. Random Forest превзошел остальных с точностью 99%. Для устранения классового дисбаланса мы создали сбалансированный набор данных путем oversampling. Главными прогностическими признаками стали гликемия и прибавка веса. Это исследование демонстрирует, что машинное обучение может способствовать раннему скринингу ГДЗ, улучшая здоровье матерей и будущих поколений.

Ключевые слова: гестационный сахарный диабет, машинное обучение, случайный лес

Abstract. Gestational diabetes mellitus (GDM) is a major concern in maternal and fetal health, with rising incidence globally. This study presents a machine learning approach for early detection of GDM based on a dataset of about 2,000 pregnant women from six regions of Uzbekistan. Among them, 10% were diagnosed with GDM. We examined six machine learning algorithms—Support Vector Machine (SVM), Naive Bayes, Random Forest, LDA, K-Nearest Neighbors (KNN), and Logistic Regression. Random Forest outperformed others with 99% accuracy. To address class imbalance, we created a balanced dataset via oversampling. Glycemia and weight gain emerged as key predictive features. This research demonstrates that machine learning can support early screening of GDM, improving health outcomes for mothers and future generations.

Key words: Gestational diabetes mellitus, Machine Learning, Random Forest



1. Muammoning o'rganilganligi

Gestatsion qandli diabet deb birinchi marta homiladorlik paytida aniqlangan, ammo aniq qandli diabet parametrlariga mos kelmaydigan diabetga aytiladi [9]. GQD, odatda, yumshoq giperglikemiya nazarida tutadi, ammo shunga qaramay, homiladorlik, tugʻruqning noxush oqibatlari va kelajakda diabet [5], yurak-qon tomir kasalliklari, metabolik sindrom, semizlik rivojlanishining yuqori xavfi bilan bogʻliq va mos ravishda tuzatilishi kerak [2, 8]. Shu sababli, GQD bilan ogʻrigan homilador ayollarni olib borish jihatlari yosh ayollarda diabet va boshqa kasalliklarning oldini olish nuqtai nazaridan yanada muhim ahamiyat kasb etadi [11].

GQD uchun nafaqat ona organizmiga, balki homilaga ham ta'siri, tugʻayotgan ayollar va avlod uchun uzoq muddatli oqibatlar mavjudligi va shuning uchun uning oldini olish va davolash masalasi alohida ahamiyatga ega. Homiladorlik va tugʻruq davrida onalar va chaqaloqlarda asoratlarni davolash va oldini olish masalalari [7, 13] keng o'rganilmoqda. Tibbiy jihatlardan tashqari, iqtisodiy jihatlar ham mavjud: GQD bilan ogʻrigan bemorni davolash (Angliyada) kasalliksiz homilador ayolga xizmat koʻrsatishga qaraganda 34% qimmatroq ekanligi hisob-kitob qilingan [7]. Gestatsion qandli diabet (GQD) - bu birinchi marta homiladorlik davrida, odatda ikkinchi yoki uchinchi trimestrda aniqlangan glyukozani koʻtara olmaslik bilan tavsiflangan holat. Preeklampsiya, makrosomiya va neonatal gipoglikemiya kabi asoratlarning xavfini oshiradi, shuningdek ona va bolada 2-toifa diabet uchun xavf omili hisoblanadi. GQDni oʻz vaqtida aniqlash parhez, jismoniy mashqlar va zarur boʻlganda insulin terapiyasi orqali erta davolanishga imkon beradi.

Oʻzbekistonda, koʻplab past va oʻrta daromadli mamlakatlarda boʻlgani kabi, glyukozani doimiy nazorat qilish va tashxislash vositalaridan foydalanish imkoniyatining cheklanganligi odatiy klinik maʼlumotlardan foydalanadigan bashorat qilish vositalariga ehtiyojni keltirib chiqaradi. Ushbu tadqiqotda biz Oʻzbekistonning oltita viloyati boʻyicha toʻplangan haqiqiy klinik va demografik



ma'lumotlardan foydalangan holda GQDni bashorat qilish uchun mashinali o'rganish metodlarini qo'llaymiz.

Bir qator tadqiqotlar tibbiy diagnostikada, xususan, homiladorlik bilan bog'liq asoratlarni bashorat qilishda mashinali o'rganishning imkoniyatlaridan foydalanishni tahlil qilgan. [3]da birinchi trimestrda elektron sog'liqni saqlash yozuvlaridan GQDni bashorat qilish uchun mashinali o'rganish modelini ishlab chiqilgan va AUC 0,85 ga erishilgan. Xitoylik mutaxassislar logistik regressiya va qarorlar daraxti modellarini qo'llagan holda [12] och qoringa o'lchangan glyukoza miqdori, onalik yoshi va TVIni GQDni aniqlashda muhim rol o'ynovchi alomatlar ekanligini aniqlashdi. [6]da ko'rsatilishicha, GDMni erta bosqichda aniqlash uchun ansambl o'rganish modellaridan foydalangan va Random Forest boshqa algoritmlardan ustunligini aniqlagan. Hindiston kasalxonalarida ma'lumotlari bo'yicha SVM, KNN va Naive Bayesdan foydalangan [10] va SMOTE (Sintetik ozchilikni ortiqcha namunalash texnikasi) yordamida sinf nomutanosibligini bartaraf etilgan. Ushbu tadqiqotlar GQDni bashorat qilishda mahalliyashtirilgan ma'lumotlar, muvozanatli namuna olish va talqin qilinadigan modellarning muhimligini ta'kidlaydi.

2. Berilganlar bazasi (Dataset) va uni dastlabki qayta ishlash

Ushbu tadqiqot doirasida O'zbekistonda birinchi marta homilador ayollar orasida skrining o'tkazish orqali o'zbek populyatsiyasidagi homilador ayollar orasida GQD uchrash chastotasini o'rganish va ustunlik qiluvchi xavf omillarini aniqlash amalga oshirildi.

O'zbekistonning 6 mintaqasida 1842 nafar homilador ayol tekshirildi, ulardan 192 nafarida GQD aniqlandi, bu 10,42% ni tashkil qiladi (1-jadval).

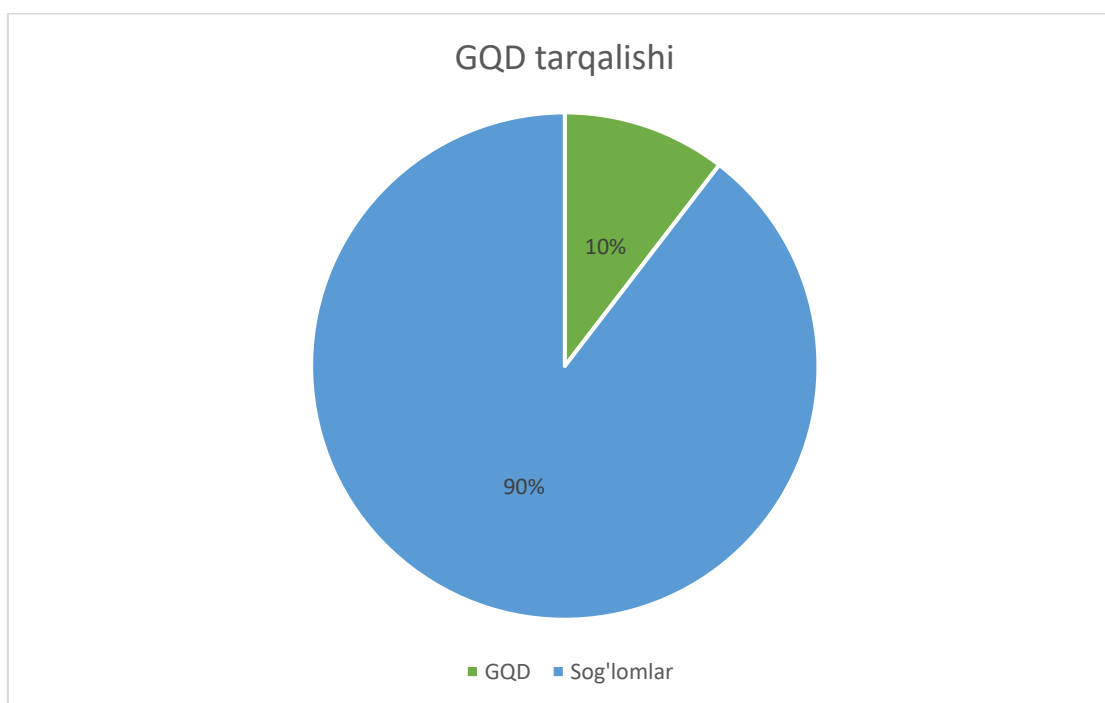
Hududlar bo'yicha GQD uchrash chastotasi: Samarqand 11,7%, Farg'ona - 9,6%, Xorazm - 8,9%, Surxondaryo - 13,3%, Namangan - 6,6%, Qashqadaryo - 13,57% ni tashkil etdi (1-jadval). GQD bilan kasallangan homilador ayollarning



eng ko'p soni boshqa hududlarga nisbatan tug'ilishda Surxondaryo - 13,1% viloyatida aniqlangan.

1-jadval. Bazaga kiritilgan berilganlarning viloyatlar kesimida soni.

Hudud	Jami	GQD	GQD, %
Xorazm	358	32	8,9
Samarqand	274	32	11,7
Farg'ona	298	27	9,06
Surxondaryo	292	39	13,3
Namangan	318	21	6,6
Qashqadaryo	302	41	13,57
Jami	1842	192	10,42



1-rasm. Bazadagi berilganlarda GQD va sog'lom homiladorlar ulushi.

-Ma'lumotlar manbasi: O'zbekistonning 6 ta viloyatidan viloyat tug'ruqxonalari bilan hamkorlikda yig'ilgan **1842** nafar atrohofilador ayolning tibbiy kartasi.



- Maqsadli o'zgaruvchi: GQD diagnostikasi (binar sinf: GQD yoki GQD bo'lmagan).

- Xususiyatlari: Yoshi, TVI, glikemiya, homiladorlik paytida vazn ortishi, oilaviy anamnezda diabet, qon bosimi va boshqa muntazam o'lchovlar.

- Ma'lumotlar nomutanosibligi: yozuvlarning 10 foizi ijobiy (GQD) sinfiga tegishli. Ushbu nomutanosiblikni yumshatish uchun biz Tasodifiy ortiqcha namuna olishni qo'lladik.

3. Masalani yechish metodologiyasi

Biz sinflarni kuchli ajratishni ta'minlaydigan kombinatsiyalarni aniqlash uchun juft belgilar tahlilini o'tkazdik. Glikemiya va vazn ortishi sinflar o'rtasidagi eng aniq chegarani ko'rsatdi.



Mashinali o'rganish algoritmlari:

- Logistik regressiya - sigmoid funksiya yordamida ikkilik natijalarni baholovchi chiziqli model.



- Tasodifiy oʻrmon: Tasodifiy tanlash va yigʻish bilan qaror daraxtlari ansambli.

- Tayanch vektor mashinasi (SVM): Yuqori oʻlchamli fazoda optimal ajratish uchun gipertekisliklarni quradi.

- Sodda Bayes: Bayes teoremasidan foydalangan holda ehtimoliy tasniflagich.

- K-Eng yaqin qoʻshnilar (KNN) - eng yaqin qoʻshnilar boʻyicha

4. Tadqiqot natijalarining tahlili

Algoritm	Balanslangan dataset boʻyicha model aniqligi	Balanslangan dataset boʻyicha model aniqligi
Random Forest	99	99
LDA	95	92
SVM	49	94
KNN	89	93
Naive Bayes	99	98
Logistic Regression	95	93

Muvozanatlashtirilgandan soʻng, barcha modellar GQD holatlarini aniqlashda yaxshilandi. Random Forest barcha koʻrsatkichlar boʻyicha eng yaxshi natijalarni koʻrsatdi. Belgilar ahamiyati tahlili glikemiya va vazn ortishini kritik koʻrsatkichlar sifatida tasdiqladi.

5. Xulosa

Bizning natijalarimiz shuni koʻrsatadiki, Random Forest real dunyo maʼlumotlaridan foydalangan holda GQDni erta bashorat qilish uchun ishonchli



model hisoblanadi. Glikemiya va vazn ortishi kabi klinik belgilarning kiritilishi modelning ishonchliligini oshiradi. Mashinali o'rganish orqali erta tashxis qo'yish o'z vaqtida tibbiy aralashuvni osonlashtirishi va O'zbekistonda onalar va bolalar salomatligi uchun uzoq muddatli xavflarni kamaytirishi mumkin.

Tadqiqotning kelajakdagi istiqboli sifatida quyidagilarni ko'rsatish mumkin:

- Ma'lumotlar bazasini turmush tarzi va genetik omillarni o'z ichiga olgan holda kengaytirish;
- Modelni klinikalarda qaror qabul qilishni qo'llab-quvvatlash vositasi sifatida joylashtirish;
- Xalqaro ma'lumotlar to'plamlari bo'yicha samaradorlikni taqqoslash.

Foydalanilgan adabiyotlar

1. American Diabetes Association. (2014). Diagnosis and classification of diabetes mellitus. Diabetes care, 37(Supplement_1), S81-S90.
2. American Diabetes Association. 14. Management of diabetes in pregnancy: standards of medical care in diabetes – 2019. Diabetes Care 2019; 42, S165-S172.
3. Artzi, N. S., et al. (2020). Prediction of gestational diabetes based on nationwide electronic health records. Nature Medicine, 26(1), 71–76.
4. Breiman, L. (2001). Random forests. Machine learning, 45(1), 5–32.
7. Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20(3), 273–297.
5. Chen T., Liu D., Yao X. Progress of Clinical Trials for the Treatment of Gestational Diabetes Mellitus. Diabetes Metab Syndr Obes. 2021; 14:315-327. 115.



6. Cheung, Y., et al. (2019). Early detection of gestational diabetes using machine learning. *Journal of Biomedical Informatics*, 93, 103154.
7. Ignatev N. A. and Rahimova M. A. Formation and Analysis of Sets of Informative Features of Objects by Pairs of Classes // *Scientific and Technical Information Processing*, 2022, Vol. 49, No. 6, pp. 439–445.
8. Ignatev N. A. On Nonlinear Transformations of Features Based on the Functions of Objects Belonging to Classes // *Pattern Recognition and Image Analysis*. 2021. V. 31. № 2. P. 197–204.
9. National Institute for Health and Care Excellence Diabetes in Pregnancy: Management from Preconception to the Postnatal Period, NICE Guideline (NG3) 2015. 212.
10. Patil, A., et al. (2018). Prediction of gestational diabetes using machine learning techniques. *International Journal of Computer Applications*, 182(35), 6–10.
11. Wendland EM, Torloni MR, Falavigna M, et al. Gestational diabetes and pregnancy outcomes--a systematic review of the World Health Organization (WHO) and the International Association of Diabetes in Pregnancy Study Groups (IADPSG) diagnostic criteria. *BMC Pregnancy Childbirth* 2012;12:23.
12. Wu, Y., et al. (2021). Machine learning approaches for risk factor analysis and prediction of gestational diabetes mellitus. *Scientific Reports*, 11(1), 1–10.
13. Игнатъев Н.А., Акбаров Б.Х. Оценка близости структур отношений объектов обучающей выборки на многообразиях наборов латентных признаков // *Вестник Томского государственного университета. Управление, вычислительная техника и информатика*. 2023. № 65. С. 69–78. doi: 10.17223/19988605/65/7.