



MASHINAVIY O'QITISH ALGORITMLARI ASOSIDA MAKTAB BITIRUVCHILARINING YO'NALISHLARINI ANIQLASH MODEL VA SUN'IY INTELLEKT

O.U. Mallayev^{1,3}, J.Q. Aliyev¹, J.A. G'azatov^{1,2},

¹ *Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti, Toshkent, O'zbekiston E-mail: jaloliddinaliyev0@gmail.com*

ORCID: <https://orcid.org/0009-0001-5093-1917>

² *Nizomiy nomidagi O'zbekiston milliy pedagogika universiteti, Toshkent, O'zbekiston*

³ *Alfraganus universiteti, Toshkent, O'zbekiston*

Annotatsiya: Ushbu maqolada maktab bitiruvchilarining akademik ko'rsatkichlari, qiziqishlari va psixometrik xususiyatlari asosida ularning kelajak kasbiy yo'nalishlarini aniqlash uchun mashinaviy o'qitish algoritmlariga asoslangan intellektual model taklif etiladi. Tadqiqot davomida bitiruvchilarning fanlardan olgan baholari, test natijalari, qiziqish so'rovlari va ijtimoiy-demografik omillari tahlil qilindi. Qaror daraxti (Decision Tree), Tasodifiy o'rmon (Random Forest), Logistik regressiya va K-eng yaqin qo'shnilar (KNN) algoritmlaridan foydalanildi. Eksperimental natijalar shuni ko'rsatdiki, Random Forest algoritmi eng yuqori aniqlikni ta'minladi. Taklif etilgan model ta'lim muassasalarida kasbiy yo'naltirish jarayonini avtomatlashtirish va samaradorligini oshirishga xizmat qiladi.

Kalit so'zlar: mashinaviy o'qitish, kasbiy yo'naltirish, bitiruvchilar, sun'iy intellekt, ta'lim tahlili.

Annotation: This paper proposes an intelligent model based on machine learning algorithms to determine the future academic and professional orientations of high school graduates using academic performance, interests, and psychometric indicators. The study analyzes students' subject grades, test results, interest surveys, and socio-demographic factors. Decision Tree, Random Forest, Logistic Regression, and K-Nearest Neighbors algorithms were applied. Experimental results demonstrate that the Random Forest algorithm achieved the highest prediction accuracy. The proposed model can be used to automate and improve career guidance processes in educational institutions.

Keywords: machine learning, career guidance, graduates, artificial intelligence, educational data mining.

Аннотация: В статье предлагается интеллектуальная модель на основе алгоритмов машинного обучения для определения будущих профессиональных направлений выпускников школ на основе их академических показателей, интересов и психометрических характеристик. В исследовании использовались алгоритмы дерева решений, случайного леса, логистической регрессии и метода ближайших соседей. Экспериментальные результаты показали, что алгоритм случайного леса обеспечивает наибольшую точность. Предложенная модель может быть использована для автоматизации процессов профориентации в образовательных учреждениях.





Ключевые слова: машинное обучение, профориентация, выпускники, искусственный интеллект, анализ образовательных данных.

Kirish

Bugungi kunda ta'lim tizimida bitiruvchilarning kelajak kasbiy yo'nalishini to'g'ri tanlash muhim muammolardan biri hisoblanadi. Noto'g'ri tanlangan yo'nalish o'quvchilarning akademik muvaffaqiyatsizligi, motivatsiyaning pasayishi va mehnat bozorida malakali kadrlar yetishmovchiligiga olib keladi.

Mashina o'rganish (ML) - bu ma'lumotlardan o'rganishga qodir algoritmlarni yaratish bilan shug'ullanadigan sun'iy intellektning bir sohasi. Ushbu algoritmlar aniq belgilangan qoidalar bilan dasturlashtirilish o'rniga, ma'lumotlardagi naqshlarni aniqlaydi va ulardan tajriba bilan yaxshilanadigan bashoratlar yoki qarorlar qabul qilish uchun foydalanadi.

Raqamli texnologiyalar va sun'iy intellektning rivojlanishi ta'lim jarayonida katta hajmdagi ma'lumotlarni tahlil qilish imkonini bermoqda. Mashinaviy o'qitish algoritmlari yordamida o'quvchilarning o'tmishdagi natijalari asosida ularning kelajakdagi qiziqishlari va muvaffaqiyat ehtimolini oldindan bashorat qilish mumkin.

Mazkur tadqiqotning asosiy maqsadi - maktab bitiruvchilarining fanlarga bo'lgan moyilligi va bilim darajasi asosida optimal yo'nalishni aniqlovchi mashinaviy o'qitish modelini ishlab chiqishdir.

1. Materiallar va usullar

1.1. Ma'lumotlar to'plami

Modelni o'qitish uchun quyidagi atributlardan foydalanildi:

Matematika, fizika, informatika, ona tili fanlari bo'yicha baholar

Test natijalari (DTM yoki ichki testlar)

Qiziqishlar so'rovnomasi natijalari

Psixologik test ko'rsatkichlari

Jins, yashash hududi (shahar/qishloq)

Ma'lumotlar tozalandi, normallashtirildi va yetishmayotgan qiymatlar o'rtacha qiymat bilan to'ldirildi.

1.2. Ma'lumotlarni normallashtirish

Min-Max normallashtirish formulasi:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}$$

2. Algoritmlar va matematik asoslar

2.1. Logistik regressiya

Logistik regressiya ehtimollikni quyidagi funksiya orqali hisoblaydi:

$$P(y = 1|x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n)}}$$

2.2. K-eng yaqin qo'shnilar (KNN)

Masofa Evklid formulasi orqali hisoblanadi:





$$d(x, y) = \sqrt{\sum_{i=0}^n (x_i - y_i)^2}$$

2.3. Qaror daraxti (Decision Tree)

Ma'lumotlarni bo'lishda Entropiya ishlatiladi:

$$Entropy(S) = \sum p_i - \log_2 p_i$$

2.4. Random Forest algoritmi

Random Forest bir nechta qaror daraxtlari asosida ishlaydi:

$$\hat{y} = \text{majority vote}(T_1(x), T_2(x), \dots, T_n(x))$$

3. Natijalar va muhokama

Mashinada o'rganish modellarini yaratishda ularning qanchalik yaxshi ishlashini tushunish muhimdir. Baholash metrikalari bizga modellarimiz samaradorligini o'lchashga yordam beradi. Tasniflash muammosini hal qilishdan qat'i nazar, uzluksiz qiymatlarni bashorat qilishdan yoki ma'lumotlarni klasterlashdan qat'i nazar, to'g'ri baholash metrikasini tanlash bizga modelning maqsadlarimizga qanchalik mos kelishini baholash imkonini beradi. Ushbu maqolada biz keng tarqalgan baholash metrikalarini ko'rib chiqamiz va modelimiz uchun to'g'ri metrikani qanday tanlashni muhokama qilamiz.

Algoritmlar samaradorligi quyidagi ko'rsatkichlar bilan baholandi:

Aniqlik (Accuracy)

Aniqlik tasniflash modelining ishlashini baholash uchun ishlatiladigan asosiy metrikadir. U bizga model tomonidan qilingan to'g'ri bashoratlarning barcha bashoratlar orasidan nisbatini ko'rsatadi.

Aniqlik tezkor tasvirni taqdim etsa-da, nomutanosib ma'lumotlar to'plamlari holatlarida chalg'ituvchi bo'lishi mumkin. Masalan, 90% A sinf va 10% B sinfiga ega ma'lumotlar to'plamida faqat A sinfini bashorat qiladigan model baribir 90% aniqlikka erishadi, ammo u B sinfining biron bir misolini aniqlay olmaydi.

Aniqlik yaxshi, lekin u yuqori aniqlikka erishish haqida noto'g'ri ijobiy tuyg'u beradi. Muammo kichik sinf namunalarining noto'g'ri tasniflanishi juda yuqori bo'lishi ehtimoli tufayli yuzaga keladi.

To'liqlik (Recall)

F1 hisobi

F1 balli aniqlik va eslab qolishning garmonik o'rtacha qiymatidir. Bu aniqlik va eslab qolish o'rtasida muvozanatga muhtoj bo'lganimizda foydalidir, chunki u ikkalasini ham bitta songa birlashtiradi. Yuqori F1 balli model ikkala ko'rsatkich bo'yicha ham yaxshi ishlashini anglatadi. Uning diapazoni [0,1].

Pastroq eslab qolish va yuqori aniqlik bizga katta aniqlikni beradi, ammo keyin u ko'p sonli misollarni o'tkazib yuboradi. F1 ball qanchalik ko'p bo'lsa, unumdorlik shuncha yuqori bo'ladi. Buni matematik tarzda quyidagicha ifodalash mumkin:

$$F1 \text{ ballari} = 2 * \frac{\text{Aniqlik} * \text{Eslatib o'tish}}{\text{Aniqlik} + \text{Eslatib o'tish}}$$





F1-score

№	Algoritm	Aniqlik (%)
1	Logistik regressiya	78
2	KNN	81
3	Decision Tree	84
4	Random Forest	91

Natijalar shuni ko'rsatadiki, Random Forest algoritmi murakkab bog'liqliklarni yaxshi o'rganadi va eng yuqori aniqlikni beradi.

4. Xulosa

Mazkur tadqiqotda mashinaviy o'qitish algoritmlariga asoslangan kasbiy yo'nalishlarni aniqlash modeli ishlab chiqildi. Tadqiqot natijalari shuni ko'rsatadiki, sun'iy intellekt asosidagi yondashuvlar ta'lim tizimida kasbiy yo'naltirish samaradorligini sezilarli darajada oshiradi. Kelgusida modelga neyron tarmoqlar va chuqur o'qitish algoritmlarini integratsiya qilish rejalashtirilmoqda.

Foydalanilgan adabiyotlar

- 1.Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- 2.Mitchell, T. (1997). *Machine Learning*. McGraw-Hill.
- 3.Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- 4.Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques*. Elsevier.
- 5.OECD. (2020). *Artificial Intelligence in Education*. OECD Publishing.

