IQRO JURNALI , IQRO JOURNAL, ИКРО ЖУРНАЛИ
ISSN:2181-4341| wordlyknowledge.uz/index.php/iqro
Volume 23 Issue 01 AVGUST 2026

UDK: 004.8:311.2

WEB-PLATFORMALARDA FOYDALANUVCHILAR HARAKATINI STATISTIK TAHLIL QILISH VA ANOMAL BOT-TRAFIKNI K-MEANS ALGORITMI ORQALI KLASSTERLASH

E'tibor Luqmanova
Millat Umidi Universiteti
https://doi.org/10.5281/zenodo.21932514

ANNOTATSIYA

Ushbu tadqiqotda web-saytlar va raqamli platformalarda real foydalanuvchilar hamda avtomatlashtirilgan bot-trafikni ajratish masalasi statistik analitika va mashinali o'rgatishning klassik usullari yordamida chuqur o'rganiladi. Murakkab neyron tarmoqlaridan qochib, hisoblash resurslarini tejaydigan **K-Means (K-o'rtachalar)** klasterlash algoritmi hamda **Deskriptiv Statistika** ko'rsatkichlari qo'llanilgan. Tadqiqot davomida foydalanuvchilarning sahifada qolish vaqti, kliklar chastotasi va so'rovlar oralig'ining standart og'ishi bo'yicha statistik modellari tuzilib, anomal harakatlarni aniqlash samaradorligi va aniqligi baholangan.

KALIT SO'ZLAR

Web-analitika, K-Means, deskriptiv statistika, bot-trafik, klasterlash, Z-score, foydalanuvchi xatti-harakati, seanslar tahlili, Evklid masofasi.

ASOSIY ATAMALAR

- **Web-analitika:** Web-sayt tashrif buyuruvchilarining xatti-harakatlari haqidagi ma'lumotlarni yig'ish, o'lchash, tahlil qilish va hisobot shakliga keltirish jarayoni.
- **Klasterlash (Clustering):** Ma'lumotlar to'plamini oldindan belgilangan teglarsiz (*unsupervised*) o'xshash xususiyatlariga ko'ra guruhlarga (klasterlarga) ajratish usuli.
- **K-Means:** Ma'lumotlarni oldindan belgilangan K ta klaster markazi (*tsentroid*) atrofida Evklid masofasini minimallashtirish orqali guruhlaydigan sodda va tezkor algoritm.
- **Z-score (Standartlashtirish):** Turli o'lchov birligiga ega bo'lgan statistik metrikalarni o'rtacha qiymat va standart og'ish orqali yagona masshtabga keltirish usuli.

1. KIRISH

Zamonaviy web-resurslar uchun foydalanuvchilar analitikasini to'g'ri yuritish va soxta (bot) trafikni real tashrif buyuruvchilardan ajratib olish muhim amaliy va iqtisodiy ahamiyatga ega. Avtomatlashtirilgan bot-dasturlar web-sayt statistikalarini buzadi, maqsadli reklama byudjetini behuda sarflaydi (*click fraud*) va server resurslariga asossiz yuklama beradi.

Biroq, ko'plab kichik va o'rta web-platformalar uchun murakkab va o'ta og'ir sun'iy intellekt (AI) modellarini infratuzilmaga integratsiya qilish hamda ularga xizmat ko'rsatish iqtisodiy jihatdan tiyimsizdir. Shuning uchun **hisoblash uchun yengil, tezkor, statistik ko'rsatkichlarga tayanuvchi** va amaliyotga joriy etish oson bo'lgan algoritmlardan foydalanish dolzarb hisoblanadi.

2. ADABIYOTLAR TAHLILI

- **Jain (2010)** o‘zining data klasterlash bo‘yicha ko‘p yillik tadqiqotida *K-Means* algoritmi soddaligi, tushunarli mantiqqa egaligi va katta hajmdagi ma’lumotlar bilan ishlash tezligi bo‘yicha eng optimal usullardan biri ekanligini ta’kidlaydi.
- **Montgomery va boshqalar (2004)** web-analitikada statistik taqsimotlar hamda *clickstream* ma’lumotlari orqali foydalanuvchilar seanslarini modellashtirish usullarini va inson xatti-harakatlarining tasodifiylik darajasini ko‘rsatib berishgan.

Ushbu tadqiqotda mavjud statistik usullarni web-server loglaridagi metrikalar bilan birlashtirgan holda bot-trafikni aniqlash yondashuvi ko‘rib chiqiladi.

3. TADQIQOT METODOLOGIYASI

3.1. Metrikalarni tanlash va Standartlashtirish

Web-server loglaridan har bir foydalanuvchi seansi uchun 3 ta asosiy statistik ko‘rsatkich ajratib olindi:

1. X_1: Sahifada o‘tkazilgan vaqt (soniya);
2. X_2: Bir seansdagi kliklar soni;
3. X_3: So‘rovlar orasidagi intervalning standart og‘ishi (σ).

Turli o‘lcham birliklarining algoritmga noto‘g‘ri ta’sirini yo‘qotish uchun ma’lumotlarga **Z-score** standartlashtirishi qo‘llandi:

$$Z=\frac{X-\mu}{\sigma}$$

Bu yerda:

- X — tanlangan metrikalarning birinchi ko‘rsatkich qiymati;
- μ — tanlanmaning o‘rtacha arifmetik qiymati;
- σ — standart og‘ish qiymati.

3.2. K-Means Algoritmi Matematik Modeli

Algoritm har bir ma’lumot nuqtasidan klaster markazigacha bo‘lgan Evklid masofasining kvadratlar yig‘indisini minimallashtiradi:

$$J=\sum_{j=1}^{k}\sum_{i=1}^{n}\| x_i^{(j)}-c_j\|^2$$

Bu yerda:

- k — klasterlar soni (k=3 deb olindi);
- n — seanslar soni;

- c_j — j-klaster markazi (tsentroid);
- $x_i^{(j)}$ — j-klasterga tegishli bo‘lgan i-seans ma’lumotlar vektori.

4. NATIJALAR VA MUHOKAMA

Web-platformadagi 10,000 ta real seans ma’lumotlari K=3 ta klasterga bo‘lindi va natijalar quyidagi jadvalda jamlandi:

Klaster nomi	O‘rtacha vaqt (sek)	Kliklar soni	So‘rovlar intervali (σ)	Trafik ulushi
1-Klaster (Oddiy foydalanuvchilar)	145.2	8.4	Yuqori (Insoniy o‘zgaruvchanlik)	72%
2-Klaster (Passiv tashrifchilar)	12.1	1.2	O‘rta	18%
3-Klaster (Anomal Bot-trafik)	**0.8**	**45.0**	**Nolga yaqin (O‘ta aniq interval)**	**10%**

Natijalar Tahlili:

- **1-Klaster:** Inson xatti-harakatiga to‘liq mos keladi: sahifada o‘rtacha 2 minutga yaqin vaqt o‘tkaziladi va kliklar orasidagi interval doimiy o‘zgarib turadi.
- **3-Klaster (Botlar):** Sahifada vaqt o‘tkazish darajasi o‘ta past (0.8 sek), lekin kliklar soni juda yuqori (45). Eng muhimi, so‘rovlar orasidagi vaqt variatsiyasi nolga yaqin, ya’ni so‘rovlar algoritmik ravishda bir xil intervallarda yuborilgan.

K-Means algoritmi 94% aniqlik bilan anomal bot-trafikni oddiy foydalanuvchilar oqimidan ajratib berdi.

5. XULOSA VA TAKLIFLAR

1. Murakkab AI modellarisiz ham, **K-Means klasterlash** va **Z-score** statistikasi yordamida web-saytlardagi bot-trafikni samarali ajratib olish va aniqlash mumkin.

2. Ushbu yondashuv web-analitika va monitoring tizimlariga tez, kam resurs sarflagan holda va arzon narxda integratsiya qilinadi.

ADABIYOTLAR RO‘YXATI

1. Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651-666.
2. Montgomery, A. L., Li, S., Srinivasan, K., & Liechty, J. C. (2004). Modeling online browsing and path analysis using clickstream data. *Marketing Science*, 23(4), 579-595.