

## MASHINANI O'QITISH VA O'RGANISH USULLARI

**Qodirov Kamoliddin Qodirjon o'g'li**

Alisher Navoiy nomidagi Toshkent davlat o'zbek tili  
va adabiyoti universiteti talabasi

### ANNOTATSIYA

Ushbu maqolada mashinani o'rganish ilmiy tadqiqot va sanoat sohalarida dolzarb mavzu ekanligi, muntazam ravishda yangi metodologiyalar ishlab chiqilishi, sohaning tezligi va murakkabligi nimalarda namoyon bo'layotgani haqida batafsil to'xtalangan va ilmiy asoslar keltirilgan.

**Kalit so'zlar:** Mashinali o'qitish, klasterlash, tasniflash, logistic regressiya, regressiya, python dasturlash tili, ma'lumotlar tahlili.

### ABSTRACT

In this article, the fact that machine learning is a hot topic in the fields of scientific research and industry, new methodologies are regularly developed, and the speed and complexity of the field are reflected in detail and scientific reasons are presented.

**Keywords:** machine learning, clustering, classification, logistic regression, python programming language, data analysis.

### KIRISH

Model deb ataladigan mashinani o'rganish algoritimda muammo ko'pincha biznes muammosi kontekstida yuzaga kelib, bu ma'lumotlarni aks ettiruvchi matematik ifodadir. Bundan ko'zlangan asosiy maqsad ma'lumotlardan tushunchaga o'tishdir. Masalan, agar onlayn chakana sotuvchi keyingi chorakda sotuvlarni kutmoqchi bo'lsa, ular o'tgan sotuvlar va boshqa tegishli ma'lumotlarga asoslanib, ushbu sotuvlarni bashorat qiluvchi mashinani o'rganish algoritmidan foydalanishi mumkin. Xuddi shunday: shamol tegirmoni ishlab chiqaruvchisi muhim uskunalarni vizual ravishda kuzatishi va xavfli yoriqlarni aniqlash uchun o'qitilgan algoritmlar orqali video ma'lumotlarini yetkazib berishi mumkin.

O'nta usulni umumiy ko'rinishini taklif qiladi - va siz mashinani o'rganish bo'yicha bilim va ko'nikmalingizni oshirganingizda asos yaratishingiz mumkin:

1. Regressiya
2. Tasniflash
3. Klasterlash

Ishga kirishishdan oldin muhim narsa. Keling, mashinani o'rganishning ikkita umumiy toifasini ajratamiz: **nazorat ostida va nazoratsiz**. Biz tushuntirmoqchi bo'lgan ma'lumotlarga ega bo'lganimizda, nazorat ostidagi **ML** texnikasini qo'llaymiz. Biz buni yangi kirishga asoslangan chiqishni bashorat qilish uchun oldingi kirish va chiqish ma'lumotlaridan foydalanib qilamiz. Masalan, kelasi oy xizmatga ro'yxatdan o'tadigan yangi foydalanuvchilar sonini bashorat qilmoqchi bo'lgan xizmat biznesiga yordam berish uchun nazorat ostidagi **ML** texnikasidan foydalanishingiz mumkin. Aksincha, nazoratsiz **ML prognoz** qilish uchun maqsadli o'zgaruvchidan foydalanmasdan ma'lumotlar nuqtalarini bog'lash va guruhlash usullarini ko'rib chiqadi. Boshqacha qilib aytganda, u ma'lumotlarni belgilar bo'yicha baholaydi va bir-biriga o'xshash elementlarning klasterlarini shakllantirish uchun xususiyatlardan foydalanadi. Misol uchun, siz o'xshash xususiyatlarga ega mahsulotlarni segmentlarga ajratishni xohlaydigan chakana sotuvchiga yordam berish uchun nazoratsiz o'rganish usullaridan foydalanishingiz mumkin – qaysi xususiyatlardan foydalanishni oldindan belgilamasdan.

### Regressiya

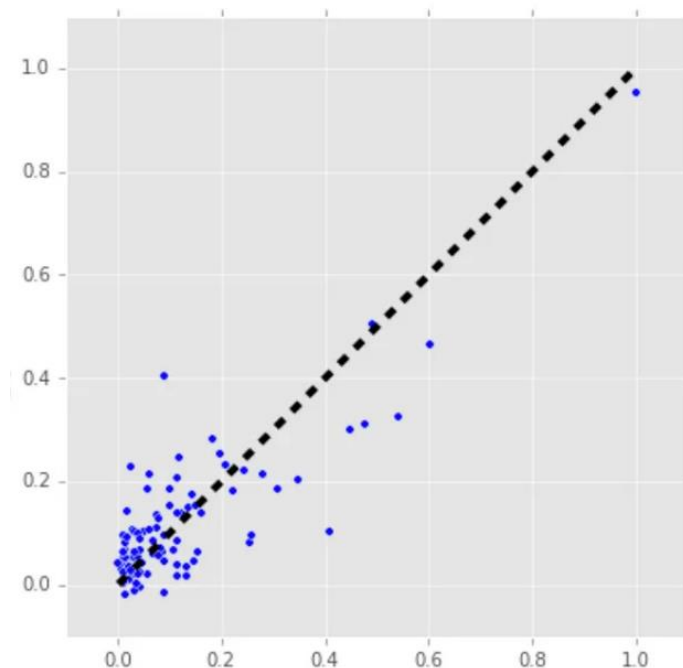
Regressiya usullari nazorat ostidagi **ML** toifasiga kiradi. Ular oldingi ma'lumotlar to'plamiga asoslangan ma'lum bir raqamli qiymatni bashorat qilish yoki tushuntirishga yordam beradi, masalan, shunga o'xshash mulklar uchun oldingi narxlash ma'lumotlari asosida mulk narxini bashorat qilish.

Eng oddiy usul chiziqli regressiya bo'lib, biz ma'lumotlar to'plamini modellashtirish uchun chiziqning matematik tenglamasidan ( $y = m * x + b$ ) foydalanamiz. Biz ko'plab ma'lumotlar juftliklari ( $x, y$ ) bilan chiziqli regressiya modelini o'rgatamiz, bu barcha ma'lumotlar nuqtalari va chiziq orasidagi umumiy masofani minimallashtiradigan chiziqning joylashuvi va qiyaligini hisoblash. Boshqacha qilib aytganda, ma'lumotlardagi kuzatishlarga eng yaqin keladigan chiziq uchun qiyalik ( $m$ ) va y-kesish ( $b$ ) ni hisoblaymiz.

Keling, chiziqli regressiyaning aniqroq misolini ko'rib chiqaylik. Men bir marta binoning yili, qavatlar sonini, kvadrat metrnlarni va tiqilib qolgan devor uskunalari sonini yig'ish orqali ma'lum binolarning energiya sarfini (kVt / soat) taxmin qilish uchun chiziqli regressiyadan foydalanganman. Bir nechta kirish (yili, kvadrat fut va boshqalar) bo'lganligi sababli, men ko'p o'zgaruvchan chiziqli regressiyadan foydalandim. Prinsip oddiy birma-bir chiziqli regressiya bilan bir xil edi, ammo bu holda men yaratgan "chiziq" o'zgaruvchilar soniga asoslangan ko'p o'lchovli fazoda sodir bo'ldi.

Quyidagi chizma chiziqli regressiya modeli binoning haqiqiy energiya sarfiga qanchalik mos kelishini ko'rsatadi. Endi tasavvur qiling-a, sizda binoning xususiyatlariga (yosh, kvadrat fut va h.k.) kirish imkoningiz bor, lekin siz energiya sarfini bilmaysiz. Bunday holda, biz ma'lum bir binoning energiya sarfini taxminiy hisoblash uchun o'rnatilgan chiziqdan foydalanishimiz mumkin.

Iste'mol qilinadigan energiyani yakuniy bashorat qilishga hissa qo'shadigan har bir omilning og'irligini baholash uchun chiziqli regressiyadan ham foydalanishingiz mumkinligini unutmang. Misol uchun, formulaga ega bo'lganingizdan so'ng, siz yosh, o'lcham yoki bo'yinging eng muhimligini aniqlashingiz mumkin.



Binoning energiya iste'moli (kVt/soat) ning chiziqli regressiya modeli.

Regressiya usullari oddiydan (chiziqli regressiya kabi) murakkabgacha (*masalan, muntazam chiziqli regressiya, polinomli regressiya, qaror daraxtlari va tasodifiy o'rmon regressiyalari, neyron tarmoqlar va boshqalar*) boshqaradi. Lekin xavotir olmang: oddiy chiziqli regressiyani o'rganishdan boshlang, texnikani o'zlashtiring va u yerdan davom eting.

### Tasniflash

Boshqariladigan **ML** ning yana bir klassi, tasniflash usullari sinf qiymatini bashorat qiladi yoki tushuntiradi. Masalan, ular onlayn mijoz mahsulot sotib oladimi yoki yo'qligini taxmin qilishga yordam beradi. Chiqish "**ha**" yoki "**yo'q**" bo'lishi mumkin: xaridor yoki xaridor emas. Ammo tasniflash usullari ikkita sinf bilan chegaralanmaydi. Masalan, tasniflash usuli berilgan tasvirda avtomobil yoki yuk

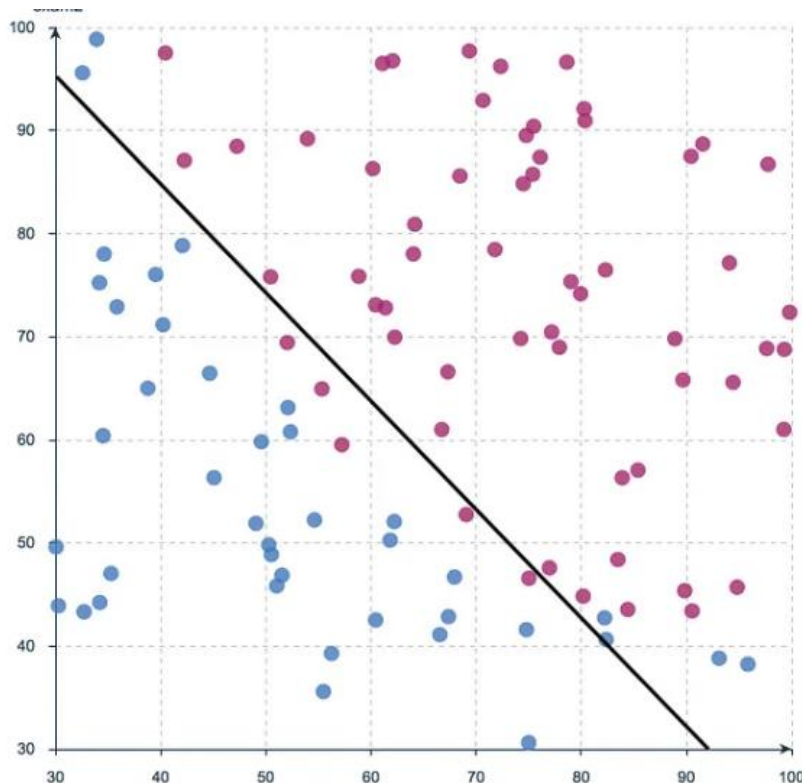
mashinasi borligini aniqlashga ko‘maklashadi. Bunday vaziyatda chiqish 3 xil qiymatga ega bo‘ladi:

- 1) rasmda avtomobil;
- 2) tasvirda yuk mashinasi;
- 3) tasvirda na avtomobil, na yuk mashinasi mavjud.

Eng oddiy tasniflash algoritmi **logistik regressiya** bo‘lib, u regressiya usuli kabi ko‘rinadi, lekin unday emas. Logistik regressiya bir yoki bir nechta ma’lumotlarga asoslanib, voqea sodir bo‘lish ehtimolini baholaydi.

Masalan, logistik regressiya talabaning ma’lum bir OTMga o‘qishga kirish ehtimolini baholash uchun talaba uchun ikkita imtixon ballini kiritishi mumkin. Baholash ehtimollik bo‘lganligi sababli, chiqish 0 dan 1 gacha bo‘lgan raqam bo‘lib, bu erda 1 to‘liq aniqlikni anglatadi. Talaba uchun, agar taxminiy ehtimollik 0,5 dan katta bo‘lsa, biz uni qabul qilinishini taxmin qilamiz. Agar taxminiy ehtimollik 0,5 dan kam bo‘lsa, biz uni rad etishini taxmin qilamiz.

Quyidagi jadvalda oldingi talabalarning ballari va ular qabul qilingan-qilmaganligi ko‘rsatilgan. Logistik regressiya bizga qaror chegarasini ifodalovchi chiziq chizish imkonini beradi.



Logistik regressiya qarori chegarasi: kollejga qabul qilinganmi yoki yo‘qmi?

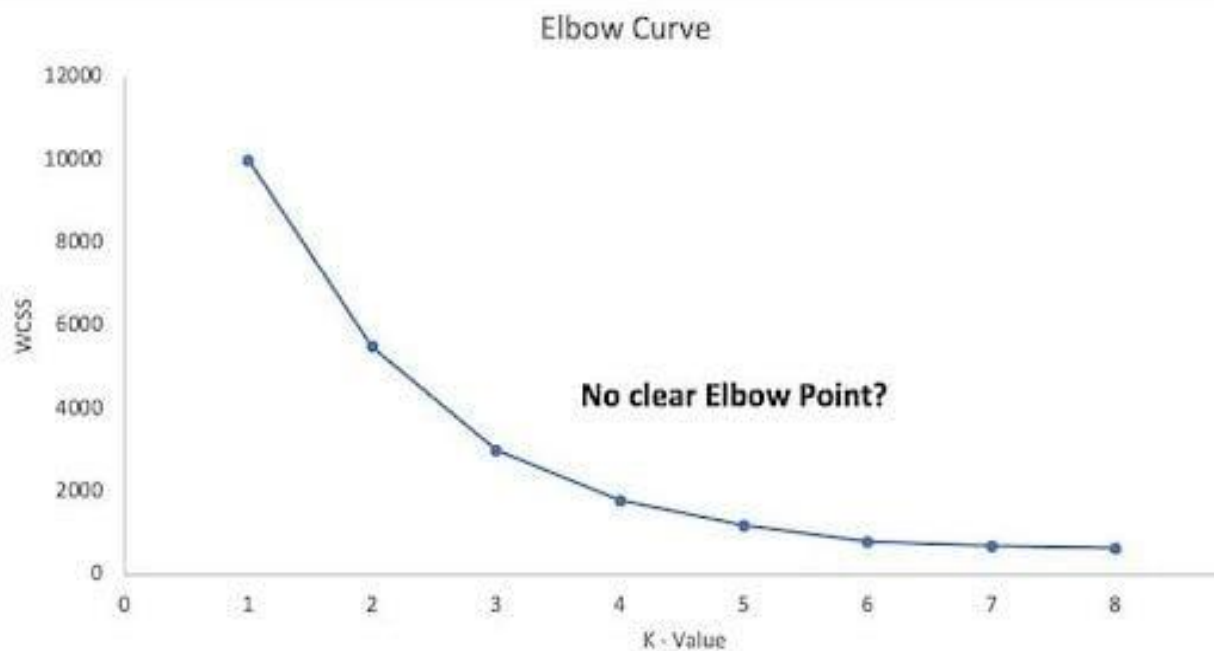
Logistik regressiya eng oddiy tasniflash modeli bo‘lganligi sababli, bu tasniflashni boshlash uchun yaxshi joy. Rivojlanayotganingizda, siz qaror daraxtlari, tasodifiy

o‘rmonlar, qo‘llab-quvvatlovchi vektor mashinalari va neyron tarmoqlari kabi chiziqli bo‘lmagan tasniflagichlarga sho‘ng‘ishingiz mumkin.

### Klasterlash

Klasterlash usullari bilan biz nazoratsiz ML toifasiga kiramiz, chunki ularning maqsadi o‘xshash xususiyatlarga ega bo‘lgan kuzatishlarni guruhlash yoki klasterlashdir. Klasterlash usullari ta’lim uchun chiqish ma’lumotlaridan foydalanmaydi, aksincha, algoritmgacha chiqishni aniqlashga imkon beradi. Klasterlash usullarida biz yechim sifatini tekshirish uchun faqat vizualizatsiyadan foydalanishimiz mumkin.

Klasterlashning eng mashhur usuli K-Means bo‘lib, bu yerda “K” foydalanuvchi yaratmoqchi bo‘lgan klasterlar sonini bildiradi. (E’tibor bering, K qiymatini tanlashning turli usullari mavjud, masalan, tirsak usuli.)

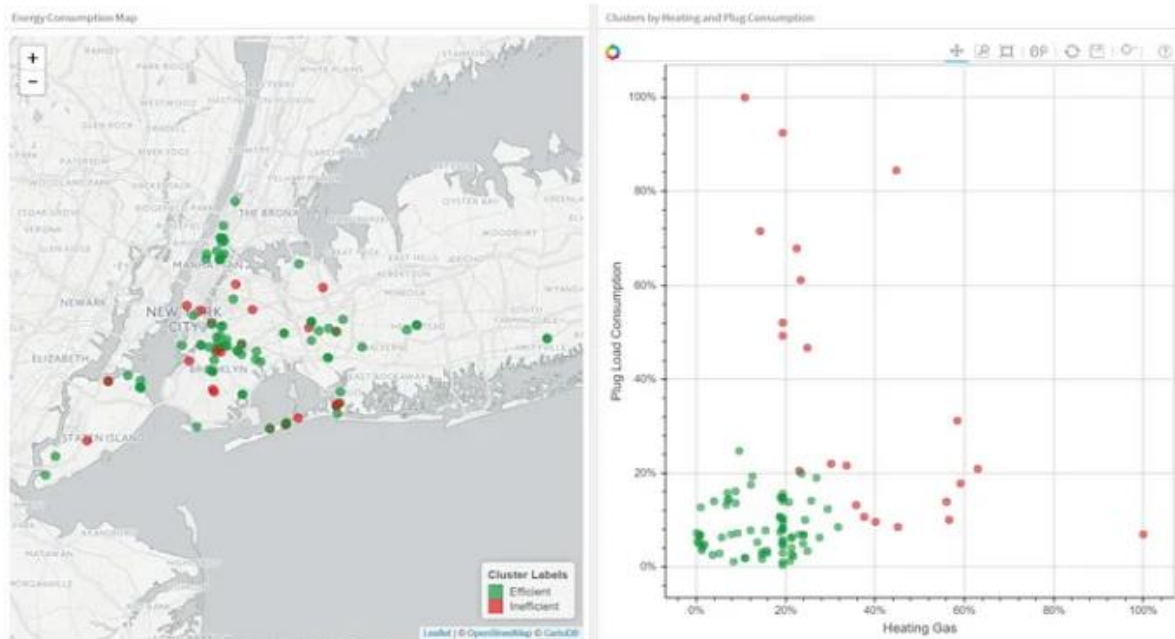


Taxminan, K-Means ma’lumotlar nuqtalari bilan nima qiladi:

1. Ma’lumotlar ichidagi K markazlarini tasodifiy tanlaydi.
2. Har bir ma’lumot nuqtasini tasodifiy yaratilgan markazlarning eng yaqiniga qo‘yadi.
3. Har bir klasterning markazini qayta hisoblab chiqadi.

Agar markazlar o‘zgarmasa (yoki juda oz o‘zgarsa), jarayon tugaydi. Aks holda, biz 2-bosqichga qaytamiz. (Agar markazlar o‘zgarishi davom etsa, cheksiz aylanaga tushib qolmaslik uchun oldindan maksimal takrorlash sonini belgilash kerak.)

Keyingi plot binolarning ma'lumotlar to'plamiga K-Meansni qo'llaydi. Plotdagi har bir ustun har bir bino uchun samaradorlikni ko'rsatadi. To'rtta o'lchov



konditsioner, rozvetkaga ulangan uskunalar (mikroto'liqli pechlar, muzlatgichlar va boshqalar), maishiy gaz va isitish gaziga tegishli. Klasterlash uchun biz  $K=2$  ni tanladik, bu klasterlardan birini samarali binolar guruhi, ikkinchisini esa samarasiz binolar guruhi sifatida izohlashni osonlashtiradi. Chap tomonda siz binolarning joylashishini va o'ng tomonda biz kirish sifatida foydalangan to'rtta o'lchamdan ikkitasini ko'rasiz: rozvetkaga ulangan uskunalar va isitish gazi.

Binolarni samarali (yashil) va samarasiz (qizil) guruhlariga ajratish.

Klasterlashni o'rganar ekansiz, siz shovqinli ilovalarning zichlikka asoslangan fazoviy klasteri (DBSCAN), o'rtacha siljish klasteri, aglomerativ iyerarxik klasterlash, Gauss aralashmasi modellaridan foydalangan holda kutish-maksimizatsiya klasteri kabi juda foydali algoritmlarga duch kelasiz.

### Xulosa

Biz ushbu kichik tadqiqotimizda mashinani o'rganishning eng muhim usullaridan bir nechtasini qamrab olishga harakat qildik. Ushbu usullarni yaxshi o'rganish va har birining asoslarini to'liq tushunish yanada ilg'or algoritm va usullarni keyingi o'rganish uchun mustahkam boshlanish nuqtasi bo'lib xizmat qilishi mumkin.

### FOYDALANILGAN ADABIYOTLAR:

1. <https://uz.itpedia.nl/2018/04/05/wat-is-machine-learning/>
2. Muhammadjonov D., Muhammadjonov S. (2008). Mashinani o‘qitish va o‘rganish usullari. – B. 12.
3. Abdusalomov A. O‘quv jarayonini tashkil etish. – B. 68.
4. Karimov Sh. (2021). Mashinani o‘rgatishda xalqaro aniq tuzilgan o‘quv dasturlari. – B. 189.
5. To‘xtamurodov N. O‘qitish va o‘rganishda yangi texnologiyalar. – B. 53.
6. <https://www.linkedin.com/pulse/types-machine-learning-techniques-training-method-based-sharma>
7. <https://blog.invgate.com/machine-learning>