

Oraliq imtixon(Fundamentals)

1. Supervised machine learning nima va qanday turlarga bo'linadi?

Javob: Supervised machine learning — bu o'qituvchi bilan o'rganish usuli bo'lib, unda model kirish ma'lumotlari (input) va ularga mos to'g'ri javoblar (label/output) asosida o'qitiladi. Modelning vazifasi yangi ma'lumot kelganda to'g'ri natijani bashorat qilishdir.

Supervised learning 2 turga bo'linadi:

1. Classification-Natija kategoriyali bo'ladi (masalan: NO/YES)

2. Regression - Natija sonli qiymat bo'ladi (masalan: uy narxini bashorat qilish)

2. Data qanday bo'lishi kerak?

Javob: Ishonchli (reliable) - noto'g'ri yoki yolg'on ma'lumot bo'lmasligi kerak

Toza (clean) - xatolar, takrorlar va bo'sh qiymatlar minimal bo'lishi kerak

Tartiblangan (structured) - analiz qilish oson bo'lishi kerak

Mavzuga aloqador (relevant) - muammoni yechishga xizmat qilishi kerak

Yetarli hajmda (sufficient amount) - model o'rganishi uchun ma'lumot ko'p bo'lishi kerak

Balanslangan (balanced) - ayniqsa classification masalalarida sinflar teng taqsimlangan bo'lishi kerak.

3. Data Preprocessing nima?

Javob: Data preprocessing — bu modelni o'qitishdan oldin ma'lumotlarni tayyorlash jarayoni bo'lib, u ma'lumotlarni tozalash, tartiblash, yetishmayotgan qiymatlarni to'ldirish, kodlash, normallashtirish va balanslash kabi bosqichlarni o'z ichiga oladi.

4. Mean va Modeni farqi nimada?

Javob: Mean(o'rtacha qiymat (faqat sonlar)) - bu sonli ma'lumotlarning o'rta arifmetik qiymati. U barcha qiymatlarni qo'shib, soniga bo'lish orqali hisoblanadi va faqat numerical ma'lumotlar uchun ishlatiladi.

Mode(eng ko'p uchragan qiymat (son va kategorik)) - bu ma'lumotlar ichida eng ko'p uchraydigan qiymat. U sonli ham, kategorik (matnli) ma'lumotlar uchun ham qo'llaniladi.

5. Regression va Classification nima va ularga idea aytin?

Javob: Regression - cheksiz(continuous) sonli qiymatni bashorat qilish masalasi.

Misol: mashinaning yurgan km miqdorini, narxini, yoqilg'i sarfini bashorat qilish.

Classification - ma'lumotlarni oldindan belgilangan chekli sinflarga ajratish masalasi. Misol: mashina urilgan/urilmagan, email spam/spam emas, bemor kasal/sog'lom.

6. Tushirib qoldirilgan qiymatlarni to'ldirish usullari va ular qanday ishlaydi?

Javob: Mean, Mode, Median, Fixed va drop usullari bor.

Mean(o'rtacha qiymat bilan) - Numerical ustunlarda ishlatiladi, barcha mavjud qiymatlar o'rtacha arifmetigi hisoblanib, yo'q qiymatlar shu bilan to'ldiriladi.

Median(o'rtadagi qiymat bilan) - Numerical ma'lumotlar uchun, qiymatlar tartiblanib, o'rtadagi qiymat bilan to'ldiriladi, outlier(keskin katta/kichik qiymatlar) bo'lsa, Mean'dan yaxshiroq.

Mode(eng ko'p uchragan qiymat bilan) - Categorical va ba'zan numerical ma'lumotlar uchun, eng ko'p takrorlangan qiymat bilan to'ldiriladi

Fixed value(oldindan belgilangan qiymat) - Foydalanuvchi tanlagan bitta qiymat bilan to'ldiriladi.

Drop(o'chirish) - Yo'q qiymatlar juda kam bo'lsa, qator(row) yoki ustun(column)

butunlay o'chirib tashlanadi.

7. ML va AI da eng muhim element?

Javob: ML va AI'da eng muhim element — data, chunki model ma'lumotsiz o'rganolmaydi.

8. Median qanday ishlaydi?

Javob: Median — bu numerical qiymatlar tartiblangandan so'ng o'rtada joylashgan qiymat bo'lib, yo'q(missing) qiymatlarni to'ldirishda aynan shu qiymat bilan almashtiriladi. Agar qiymatlar soni juft bo'lsa, o'rtadagi ikki qiymatning o'rtacha arifmetigi olinadi.