REPORT

Muammoga yechimning asosiy qadamlari;

Data (bilim)

Model (metod)

Train (o'rgatish = 80%)

Test (imtihon = 20%)

Supervised Learning

x = input ( Feature = modelni training qiladigan ustunlar)

y = output (predict qiliniyotgan ustun = target variable (input orqali modeldan chiqadigan natija))

Supervised Learning turlari

1. Classification (single/multi-class)

2. Regression

1. Classification (Tasniflash) - Agar talab qilinadigan muammoning yechimi chekli miqdorda bo'lsa. yes/no

Natija: Oldindan belgilangan kategoriya (sinf).

Email Tasnifi:

Kiruvchi ma'lumot: Email matni.

Natija (output): "Spam" yoki "Oddiy" (2 sinf - Binary Classification).

Misol: "Ushbu email spammi yoki yo'qmi?"

Rasm Tasnifi:

Kiruvchi ma'lumot: Tasvir (masalan, rasm fayli).

Natija (output): "It", "Mushuk" yoki "Qush" (3 sinf - Multi-class Classification).

Misol: "Rasmda qaysi hayvon bor?"

Tibbiy Tashxis:

Kiruvchi ma'lumot: Tibbiy test natijalari.

Natija (output): "Diabet bor", "Diabet yo'q" (2 sinf - Binary Classification).

Misol: "Bemorda diabet bor yoki yo'q?"

Classification natijasi kategoriyalar, masalan, "Ha" yoki "Yo'q", "Qizil", "Ko'k", "Yashil".

Classificationda discrete (chekli) qiymat bo'ladi, chunki tasniflash jarayonida natija (output) doimo diskret kategoriyalardan birini tanlashni talab qiladi. Masalan: {0, 1}, yoki {"Qizil", "Ko'k", "Yashil"}.

2. Regression (Regressiya)

Natija: Sonli qiymat (uzluksiz qiymat).Agar talab qilinadigan muammoning yechimi chekli bo'lmasa.

Uy Narxi Prognozi:

Kiruvchi ma'lumot: Uyning maydoni, xonalar soni, joylashuvi.

Natija (output): Uy narxi ($200,000, $300,000 va h.k.).

Misol: "Uyning narxi qancha?"

Havo Harorati Prognozi:

Kiruvchi ma'lumot: Havo bosimi, namlik, shamol tezligi.

Natija (output): Harorat (masalan, 23.5°C, 30.1°C).

Misol: "Ertaga havo harorati qancha bo'ladi?"

Mashina Tezligi Prognozi:

Kiruvchi ma'lumot: Mashina vazni, dvigatel quvvati, yo'l qiyaligi.

Natija (output): Mashina tezligi (60.2 km/soat, 80.6 km/soat).

Misol: "Mashina maksimal qancha tezlikka erishadi?"

Regression natijada sonli qiymatlar ( continues value = cheksiz) prognoz qilinadi.

Regressiyada ko'proq predict qilsak, Classificationda ko'proq sinflaymiz.

Train

Accuracy -> 75 % ( mashina ypodalab olgan bo'ladi agar katta testing bilan trainingni aniqligi har xil chiqsa)

Testing

Accuracy -> 50 % ( yaxshi model bo'lishi uchun bir birga yaqin qiymay va 90 % dan yuqori bo'lishi kerak aniqlik darajasi)

Bias - tarafkashlik

**Classification uchun asosiy modellar**

| **Model** | **Output turi** | **Misol** |
| --- | --- | --- |
| **Logistic Regression** | Diskret kategoriya | Emailni "Spam" yoki "Oddiy" deb tasniflash. |
| **Decision Tree** | Diskret kategoriya | Kredit olish layoqatini aniqlash: "Ha" yoki "Yo'q". |
| **Random Forest** | Diskret kategoriya (bir nechta sinf) | Kasallik tashxisi: "Sog'lom", "Gripp", "Diabet". |
| **Gradient Boosting** | Diskret kategoriya (bir nechta sinf) | Mijoz churnini prognozlash: "Ketadi" yoki "Qoladi". |
| **Support Vector Machine (SVM)** | Diskret kategoriya | Tasvirdagi ob'ektni aniqlash: "It" yoki "Mushuk". |
| **Neural Networks** | Diskret kategoriya (multi-class) | Rasmda ob'ektni tasniflash: "It", "Mushuk", yoki "Qush". |

**Regression uchun asosiy modellar**

| **Model** | **Output turi** | **Misol** |
| --- | --- | --- |
| **Linear Regression** | Uzluksiz qiymat (continuous value) | Uy narxi prognozi: $200,000 yoki $350,000. |
| **Polynomial Regression** | Uzluksiz qiymat | Yo‘l va vaqt o‘rtasidagi bog‘liqlikni prognoz qilish. |
| **Decision Tree** | Uzluksiz qiymat | Harorat prognozi: 23.5°C yoki 30.2°C. |
| **Random Forest** | Uzluksiz qiymat | Reklama xarajatlari va daromad o‘rtasidagi bog‘liqlikni prognoz qilish. |
| **Gradient Boosting** | Uzluksiz qiymat | Marketing kampaniyasining daromadini prognoz qilish. |
| **Neural Networks** | Uzluksiz qiymat | Energiya iste'molini prognoz qilish. |

VAZIFALAR

Table. Talaba universities kira olish yoki kira olmasligi (Classification)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Name | Name of University | Last year exam score | School  grades | How many times to submit | IELTS  score | Result |
| Alex | Massachusetts Institute of Technology | 220 | 4 | 2 | 6.5 | pass |
| Ali | Oxford University | 235 | 3 | 1 | 7.0 | fail |
| Alise | Harvard University | 266 | 5 | 1 | 6.0 | pass |
| Ferede | Qahira Al-Azhar | 247 | 5 | 2 | 5.5 | pass |

Table 2. Non savdosidan qoladigan daromadni taxmin qilish (Regression)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Type of Nan | Weight | Location | Labor cost | Season | Cost price | Selling  cost | Profit |
| Kulcha | |  | | --- | | 800 gr | | Tashkent | |  | | --- | | 3000 UZS | | |  | | --- | | qish | | |  | | --- | | 2000 som | | 2500 | |  | | --- | | 500 | |
| Patir | 1 kg | Samarqand | 2.500 | Yoz | 3000 som | 3800 | 800 |
| Tashkent nan | 1.2 kg | Tashkent | 3.500 | kuz | 9000  So’m | 12000 | 3000 |
| Samarqand nan | 1.5 kg | Samarqand | 3,500 UZS | bahor | 7000  So’m | 10000 | 3000 |