

# 1-oylik report

## 1.AI nima?

**AI(Artificial Intelligence)** – bu kompyuter yoki dasturga inson kabi fikrlash, o`rganish va qaror qabul qilish qobiliyatini beradigan texnologiyadir. Bu kompyuterni aqlli ishlashi degani.

## 2.Machine Learning nima?

Bu **AI(Sun`iy intellekt)**ning bir qismi bo`lib, kompyuterga ma`lumotlar orqali o`rganish va tajriba asosida yaxshilanish imkonini beradi. **Machine Learning** – bu kompyuterga ma`lumotlarni o`qitib, ma`lumotlar orqali o`zi mustaqil o`rganishni o`rgatish.

## 3.Supervised Machine Learning va uning turlari...

**Supervised Machine Learning** – bu kompyuterni javobi oldindan ma`lum bo`lgan ma`lumotlar bilan o`rgatish usulidir. Biz kompyuterga savol ham, uning to`g`ri javobini ham beramiz. Kompyuter shu misollar orqali o`rganadi va keyin yangi ma`lumot kelsa, o`zi javob topa oladi.

### *Supervised Machine Learning turlari:*

**Supervised ML** asosan 2turga bo`linadi.

### **1.Classification (Tasniflash)**

Bu yerda natija tur(kategoriya) bo`ladi. Ya`ni javob 0/1(nom) ko`rinishida chiqadi.

Masalan: Email→Spam/Spam emas, Talaba→O`tdi/Yiqildi, Rasm→Mushuk/It.

**Classification** – 2turga bo`linadi:

**Binary Classification** – bu yerda faqat 2ta javob varianti bo`ladi.

Masalan: Ha/Yo`q, 0/1, Spam/Spam emas.(Ya`ni javob ikkita bo`ladi holos)

**Multi-class Classification** – bu yerda 2tadan ko`p javob varianti bo`ladi.

Masalan: Rasm→Mushuk/It/Qush, Baholash→A/B/C/D, Meva→Olma/Banan/Uzum.

(Ya`ni javob ikkitadan ko`p qaytariluvchi bo`ladi)

## 2. Regression

Bu yerda natija son(raqam) ko`rinishida chiqadi.

Masalan: Uy narxi → 120.000dollar, Talaba bali → 85, Harorat → 18C.

(Ya`ni regressionda javob aniq son bo`ladi)

## 4. Data Preprocessing?

**Data Preprocessing** – bu ma`lumotlarni modelga berishdan oldin uni tozalash va tayyorlash jarayoni. *Data preprocessing* – bu modelni o`qitishdan oldin ma`lumotlarni tayyorlash jarayoni bo`lib, u ma`lumotlarni tozalaydi, tartiblash, yetishmayotgani qiymatlarni to`ldirish, kodlash, normallashtirish va balanslash kabi bosqichlarni o`z ichiga oladi.

Oddiy qilib aytganda, Xom(raw) ma`lumotlarni modelga bersak, u yaxshi ishlamaydi. Shuning uchun avval ma`lumotni tartibga keltiramiz, xatolarini tuzatamiz va model tushadigan shaklga o`tkazamiz.

Data preprocessingning asosiy bosqichlari:

1. Missing Values (Mean, Mode, Median, Fixed, Drop)
2. Encoding (One-hot, Label, Ordinal, Target, Frequency)
3. Scaling (MinMaxScaler, StandardScaler, RobustScaler)

Malumotlarni tushirib qoldirish (ustunlar-agarda ko`p bo`lsa, qatorlar-agarda kam bo`lsa)

## 5. Algorithm

Algorithm – bu muammoni yechish uchun yozilgan aniq va ketma-ket qadamlar to`plami, ya`ni biror ishni qanday bajarishni ko`rsatib beradigan yo`l-yo`riq.

Algorithm family:

Linear (Chiziqli) algoritmlar

1. Linear Regression – bu regression algorithm va natija son(raqam) bo`ladi. Ya`ni u ma`lumotlar orasidan to`g`ri chiziqli topadi va shu chiziqli yordamida qiymatni taxmin qiladi.

Masalan: Uy narxini taxmin qilish, Talabaning balini aniqlash.

2. Logistic Regression – bu classification algorithm va natija kategoriya (0 yoki 1) bo`ladi. Ya`ni u ehtimoliy hisoblaydi va natijani 0 yoki 1 qilib beradi.

Masalan: Spam/Spam emas, Kasallik bor/yo`q

### \_Tree-Based (Daraxt asosidagi) algorithmlar\_

1. Decision Tree - bu daraxt shaklida ishlaydi. Ya`ni savol-javob qilib qaror qabul qiladi.

Masalan: Yosh>18?, Ha→Talaba, Yo`q→O`quvchi

(tushinish qiyin, lekin ba`zan overfitting bo`lishi mumkin)

2. Random Forest – bu ko`p Decision Tree`lardan iborat. Ya`ni bitta daraxt emas, balki ko`p daraxt ishlaydi va oxirida ovoz berib natija chiqaradi. (Afzalligi: aniqroq, overfitting kamroq)

### \_Distance-Based (Masofa asosidagi) algorithmlar\_

1. KNN (K-Nearest Neighbors) – bu algorithm eng yaqin qo`shnilarga qarab qaror qiladi. Ya`ni yangi ma`lumot qaysi guruhga yaqin bo`lsa, o`sha guruhga qo`shadi.

Masalan: Agar yangi rasm mushuk rasmlariga yaqin bo`lsa → mushuk deb belgilaydi.

2. SVM (Support Vector Machine) – bu algorithm ma`lumotlarni eng yaxshi ajratuvchi chiziq yoki tekislik topadi. Ya`ni ikki guruh orasida maksimal masofa qoldirib ajratadi.

(Klassifikatsiya va Regressiyada ishlatiladi bo`ladi)

### \_Ensemble (Bir nechta modelni birlashtirish)\_

Ensemble – bu bir nechta modelni qo`shib kuchli model qilish

1. Random Forest – (Decision Tree`larning kombinatsiyasi)

2. Gradient Boosting – bu modelning xatosini bosqichma-bosqich kamaytirib boradi. Ya`ni har yangi model oldingisining xatosini tuzatishga harakat qiladi.

Masalan: XGBoost, LightGBM kabi mashhur variantlari ham bor.

## **6. Model**

Model – bu ma`lumotlardan o`rganib keyin yangi ma`lumot kelganda natija beradigan tizim. Ya`ni Model – bu o`rgatilgan (trained) algorithm.

Masalan: Tasavvur qilamiz, biz talabalar bahosini taxmin qilmoqchimiz; 1. Algorithmni tanlaymiz (misol; Linear Regression), 2. Uni ma`lumot bilan o`rgatamiz (training qilamiz), 3. O`rgatilgan holati - bu model bo`ladi. Keyin yangi talaba ma`lumoti kelsa, model uning taxminiy balini aytadi.

Algorithm va Model farqi:

Algorithm - bu o`rganish usuli (Decision Tree - algorithm)

Model - bu o`rganib bo`lgan natija (O`qitilib chiqilgan Decision Tree – Model)

Machine Learning`da model nima qiladi?

Model→Ma`lumotdan o`rganadi, Pattern(qonuniy) topadi, Yangi ma`lumotga javob beradi.

Masalan: Spamni aniqlaydi, Uy narxini hisoblaydi, Rasmni tanib beradi.

Xulosa: Model – bu ma`lumotlar bilan o`rgatilgan va keyin mustaqil ishlay oladigan tizim

(Algorithm – o`qituvchi, Model – o`qib chiqqan o`quvchi)

## 7.LinearRegression va LogisticRegression

Linear Regression:

Linear Regression – bu regression algorithm, ya`ni natija son(raqam) bo`ladi. U ma`lumotlar orasidan eng mos to`g`ri chiziqni topadi va shu chiziq orqali qiymatni taxmin qiladi.

Misollar: Uy narxini aniqlash, Talabanning balini taxmin qilish, Oylik daromadni hisoblash.

Masalan: Yosh oshgani sari maosh ham oshadi degan bog`lanishni topadi va chiziq chizadi.

Natija doim son bo`ladi(masalan; 85, 12000, 3.5 va hokazo)

Logistic Regression

Logistic Regression – bu classification algorithm, ya`ni natija katagoriya bo`ladi. U ehtimollik hisoblaydi va natijani 0 yoki 1 qilib beradi.

Masalan: Spam/Spam emas, Kasallik boy/yo`q, O`tdi/Yiqildi.

Logistic Regression natijani 0 va 1 orasida ehtimollik sifatida chiqaradi(Misol: 0.8)

Agar 0.5dan katta bo`lsa → 1;

Kichik bo`lsa → 0;

Ko`pincha Binary Classification uchun yaxshi ishlatiladi.

Asosiy farq:

Linear Regression→Regression uchun, Natija son, To`g`ri chiziq

Logistic Regression→Classification uchun, Natija 0/1, S-shaklli(sigmoid) funksiya

## 8.Prediction

;;Prediction – bu model o`rganib bo`lgandan keyin, yangi ma`lumotga taxminiy natija berish jarayoni. Ya`ni modelni o`qitamiz(training qilamiz), keyin u yangi ma`lumotni ko`rib, javob aytadi. Shu javob - PREDICTION.

Masalan: Siz modelga talabalar haqida ma`lumot berdingiz va u bahoni o`rganib oldi. Endi yangi talaba ma`lumoti keldi.(Yosh:20, Davomati:90%, Uy vazifasi:yaxshi → Model aytadi: 85 ball oladi.) Mana shu 85 - bu prediction.

Prediction turlari:

1.Regression Prediction – natija son bo`ladi.(Masalan:Uy narxi-123.000\$, Harorat-18°C, Talaba bali-90)

2.Classification Prediction – natija kategoriya bo`ladi.(Masalan:Spam/Spam emas, Mushuk/It, O`tdi/Yiqildi). Ba`zan model avval ehtimollik beradi(masalan 0.7), keyin natija tanlaydi.

Prediction – bu modelning yangi ma`lumotga bergan javobi.

(Training – o`rganish, Prediction – taxmin qilish)

## 9.Evaluation

Evaluation – bu model qanchalik yaxshi ishlayotganini tekshirish jarayoni. Ya`ni modelni o`qitib bo`lgandan keyin uning natijasi to`g`rimi yoki yo`qmi shuni tekshiramiz. Shu jarayon - evaluation deyiladi.

Nima uchun kerak?

Agar modelni tekshirmasak, u yaxshi ishlayaptimi yoki yo`qmi bilmaymiz. Shuning uchun test ma`lumotlar yordamida modelni baholaymiz.

Evaluation turlari:

1.Classification uchun baholash

Bu yerda natija kategoriya bo`ladi(masalan: Spam/Spam emas)

2.Regression uchun baholash

Bu yerda natija son bo`ladi.

Mae(Mean Absolute Error) – O`rtacha xato miqdori

MSE(Mean Squared Error) – Xatolarni kvadrat qilib o`rtacha olish

R<sup>2</sup>\_Score – bu regression model qanchalik yaxshi tushintira olayotganini ko`rsatadi

(Model o`rgandi → Endi tekshiramiz → Qanchalik yaxshi ishlayapti shuni bilamiz.)