

## **2-qism.9-dars**

Bir biriga yaqin featurelar overfitting (complex)

Featurelar yetarli bo'lmasa under fitting

Trading line bilan accuracy line ni orasida farqni qanchaligidan nimani hisoblasak bo'ladi juda oddiy - underfitting

juda murakkab —overfitting

## **Advanced supervised learning techniques**

Supervised Learningdagi faqat Regression va Classification modellari bilan chegaralanmay, murakkab vazifalar uchun jarayonni yanada kuchliroq va murakkabroq qilish uchun ishlab chiqilgan.

### **Turlari**

#### **1. Random Forest**

- Bir nechta decision treelardan tashkil topgan
- Har bir tree turli qismlarni predict qiladi, keyin umumiy qaror qabul qilinadi.
- Foydasi: Katta va murakkab ma'lumotlar bilan ishlay oladi, haddan tashqari moslashishdan (overfitting) saqlaydi.

#### **2. Gradient Boosting ( XGBost)**

- Model o'z xatolarini bosqichma bosqich tuzatadi.
- Har bir yangi model avalgi modelning xatolarini o'rganib, yaxshilanadi.
- Qo'llanishi: Kredit ballni aniqlash yoki xaridorlarning xatti-harakatlarini predict qilish.

#### **3. Support Vector Machines ( SVM)**

- Bu model ma'lumotlarni turkumga ajratish uchun chegarani ( boundary) topishga harakat qiladi.
- Foydasi: Kichikroq va murakkab ma'lumotlarda juda yaxshi ishlaydi.

#### **4. Ensemble Methods (Yig'ma Usullari)**

- Bir nechta modellarning natijasini birlashtirib, eng yaxshi bashoratni beradi.
- Masalan: Bagging, Boosting va Stacking usullari

#### **5. Deep Neural Networks**

- Ma'lumotlarni chuqur xususiyatlarini o'rganadi.
- Qo'llanishi: Rasm tanish, ovoz aniqlash, tabiiy tilni qayta ishlash (NLP)

### **QO'LLANISHI**

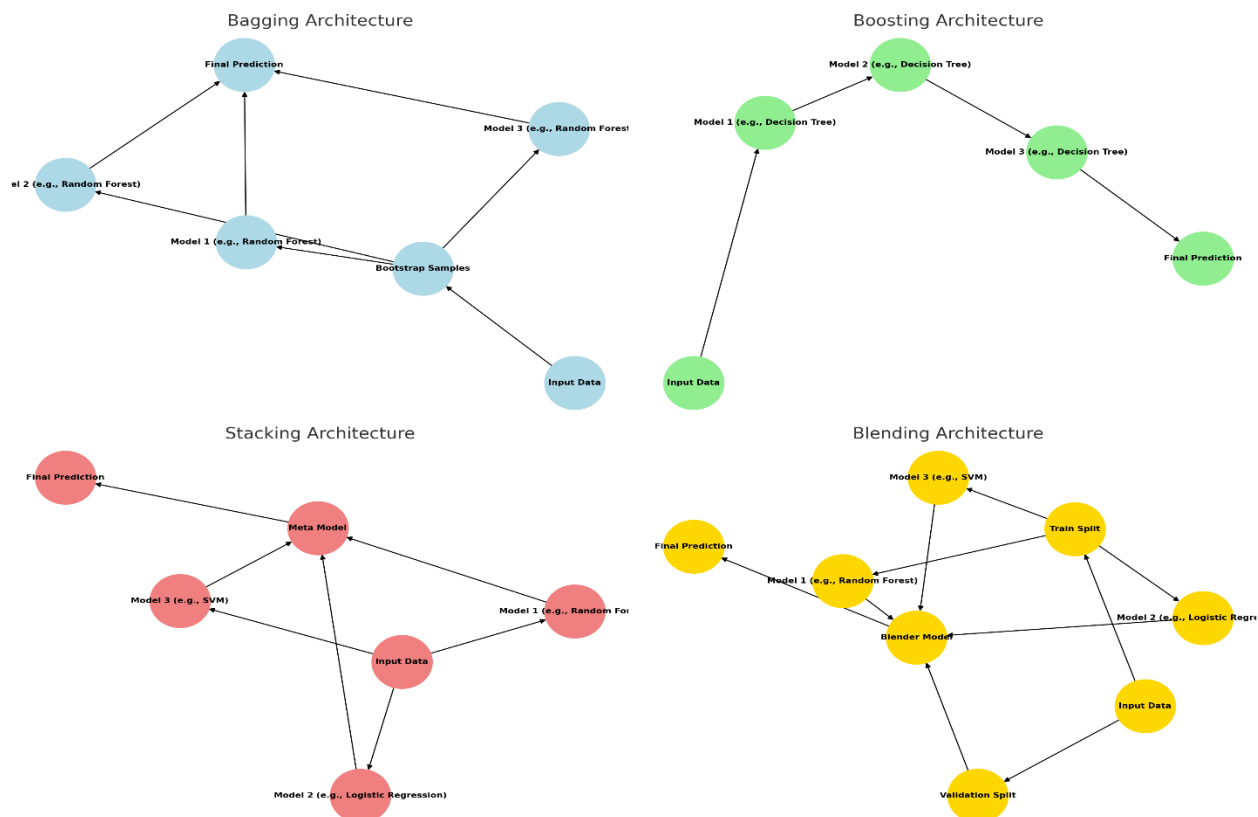
- katta miqdordagi ma'lumotlar
- Bashorat yanada aniqroq qilish kerak bo'lsa
- Real dunyo muammolarida qo'llanishi samaradorligi

## **Ensemble Methods (Yig'ma Usullar)**

Ensemble usullari bir nechta modelni birlashtirib, yakuniy natijani yaxshilashga qaratilgan yondashuvdir. Har xil usullar (bagging, boosting, stacking) o'zlarining ishlash tamoyili va qo'llanilish sohasiga ega.

<b>Xususiyat</b>	<b>Bagging</b>	<b>Boosting</b>	<b>Stacking</b>	<b>Blending</b>
<b>Modellar mustaqilligi</b>	Ha, mustaqil ishlaydi. Har bir model alohida o'qitiladi.	Yo'q, bog'liq. Keyingi model avvalgi modelning xatolarini tuzatadi.	Ha, lekin natijalar meta-model bilan birlashtiriladi.	Ha, train va validation datasetlardan foydalanadi, lekin natijalar blender modelga kiritiladi.
<b>Xatolar bilan ishlash</b>	Xatolarni kamaytiradi, barqaror natija beradi.	Xatolarni bosqichma-bosqich tuzatadi, e'tiborni noto'g'ri natijalarga qaratadi.	Modellar natijalarini uyg'unlashtiradi va meta-model bilan yakuniy bashoratni chiqaradi.	Validation dataset natijalari asosida meta-model yakuniy bashoratni qiladi.
<b>Overfitting xavfi</b>	Past, chunki tasodifiy namunalardan foydalanadi.	Yuqori, model ma'lumotga haddan tashqari moslashib ketishi mumkin.	O'rtacha, lekin meta-model bilan moslashuvni boshqarish mumkin.	O'rtacha, lekin validation dataset asosida moslashuv boshqariladi.
<b>Ma'lumotlar ishlash</b>	Tasodifiy namunalar. Train dataset tasodifiy bo'laklarga (bootstrap sampling) bo'linadi. Har bir model bir xil ma'lumot bilan ishlamaydi.	Train dataset butun holda ishlatiladi. Modellar ketma-ket o'qitilib, avvalgi model xatolarini tuzatadi.	Train dataset bir nechta model uchun ishlatiladi. Validation dataset natijalari meta-modelga kiritiladi.	Train dataset train va validation qismlariga bo'linadi. Validation natijalari meta-model orqali birlashtiriladi.

Xususiyat	Bagging	Boosting	Stacking	Blending
Misol	Random Forest. Tasodifiy daraxtlar yig'masi ma'lumotni mustaqil o'qitadi va ovoz berish orqali natija chiqaradi.	Gradient Boosting. Avvalgi model 200 ta noto'g'ri ma'lumot qilgan bo'lsa, keyingi model bu xatolarni tuzatadi.	Siz Random Forest, Logistic Regression, va SVM ishlatdingiz. Ularning natijalari regressiya modeli bilan birlashtiriladi.	Train va Validation natijalari Blender model orqali yakuniy bashoratga birlashtiriladi.
Kutubxona	scikit-learn	XGBoost, LightGBM, CatBoost	scikit-learn	scikit-learn



## Umumiy taqqoslash

Usul	Train uchun ishlatiladigan ma'lumotlar	Validation uchun ishlatiladigan ma'lumotlar	Natijani yaratish jarayoni
<b>Bagging</b>	Har bir model tasodifiy 800 ta yozuv ishlatadi	Yo'q	Mustaqil modellar natijalari ovoz berish yoki o'rtacha qilish orqali birlashtiriladi.
<b>Boosting</b>	Har bir model barcha 1000 ta yozuv ishlatadi	Yo'q	Modellar xatolarni bosqichma-bosqich tuzatadi.
<b>Stacking</b>	800 ta yozuv	200 ta yozuv	Har bir model validation natijalarini meta-modelga kiritadi.
<b>Blending</b>	700 ta yozuv	300 ta yozuv	Validation natijalari blender modelga kiritilib, yakuniy natija qilinadi.

1. Bagging (Bootstrap Aggregating) Bagging so'zi Bootstrapping (tasodifiy namunalar olish) va Aggregating (natijalarni birlashtirish) so'zlaridan olingan.

2. Boosting so'zi "yaxshilash" degan ma'noni anglatadi, chunki u modelning aniqligini xatolarni tuzatish orqali yaxshilaydi.

3. Stacking – yig'ish, Modellar natijalarini yig'ib, meta-model bilan birlashtirish.

4. Blending - Aralashtirish, Train va validation natijalarini birlashtirib yakuniy bashorat qilish.