



## RANDOM FOREST MODELI VA MODELNING GIPERPARAMETRLARI

---

*Assistant Salimov Jamshid Obid o'g'li*

*Jizzakh branch of National University of Uzbekistan*

*ORCID 0000-0002-9276-9719*

*[jamshidsalimov8@gmail.com](mailto:jamshidsalimov8@gmail.com)*

*Tel: +998904571257*

*к.ф.м.н Алдай Мақтагул*

*Eurasian National University named after L.N. Gumilyov*

*Student Abdiyev Samandar Ilhom o'g'li*

*Jizzakh branch of National University of Uzbekistan*

*[abdiyevsamandar0098@gmail.com](mailto:abdiyevsamandar0098@gmail.com)*

*Tel: +998930350098*

**Аннотация:** Mashinali o'rganish algoritmlari ma'lumotlardan qimmatli tushunchalarni olishda, biznes va tadqiqotchilarga ongli qarorlar qabul qilishda muhim rol o'ynaydi. Bunday algoritmlardan biri tasniflash vazifalari uchun keng qo'llaniladigan Random Forest. Random Forest tasnifi ko'plab Qarorlar daraxtini qo'radi va yangi obyekt eng ko'p ovoz tuplagan klassga tegishli deb qaraladi. bu esa uni murakkab muammolarni hal qilish uchun intuitiv va kuchli vositaga aylantiradi. Ushbu maqolada qaror Random Forest algoritmi yordamida ma'lum bir kasallikka chalingan bemorga qaysi dori mos kelishini aniqlaydigan model yaratilgan. Bu muammo multi-sinf tasnifi (multiclass classification) yordamida hisoblanadi. Shu bilan bir qatorda, funktsiya, ko'plab modular transformerlar va giperparametrlar orqali Random Forest algoritmi amalga oshirilgan.

**Аннотация:** Алгоритмы машинного обучения играют важную роль в извлечении ценной информации из данных, помогая предприятиям и исследователям принимать обоснованные решения. Одним из таких алгоритмов является Random Forest, который широко используется для задач классификации. Классификатор случайного леса поддерживает несколько деревьев решений, и новый объект считается принадлежащим классу, набравшему наибольшее количество голосов. что делает его интуитивно понятным и мощным инструментом для решения сложных проблем. В данной статье создана модель, определяющая, какой препарат подходит пациенту с определенным заболеванием, с помощью алгоритма Random Forest. Эта проблема рассчитывается с использованием многоклассовой классификации.



Кроме того, алгоритм Random Forest реализован посредством функции, множества модульных преобразователей и гиперпараметров.

**Annotation:** Machine learning algorithms play an important role in extracting valuable insights from data, helping businesses and researchers make informed decisions. One such algorithm is Random Forest, which is widely used for classification tasks. The Random Forest classifier maintains multiple Decision Trees, and a new object is considered to belong to the class with the highest number of votes. which makes it an intuitive and powerful tool for solving complex problems. In this article, a model was created that determines which drug is suitable for a patient with a certain disease using the Random Forest algorithm. This problem is calculated using multiclass classification. In addition, the Random Forest algorithm is implemented through the function, many modular transformers and hyperparameters.

**Key words:** Random Forest, Dataset, testset, trainset, hyperparameters, classification, prediction, transformers.

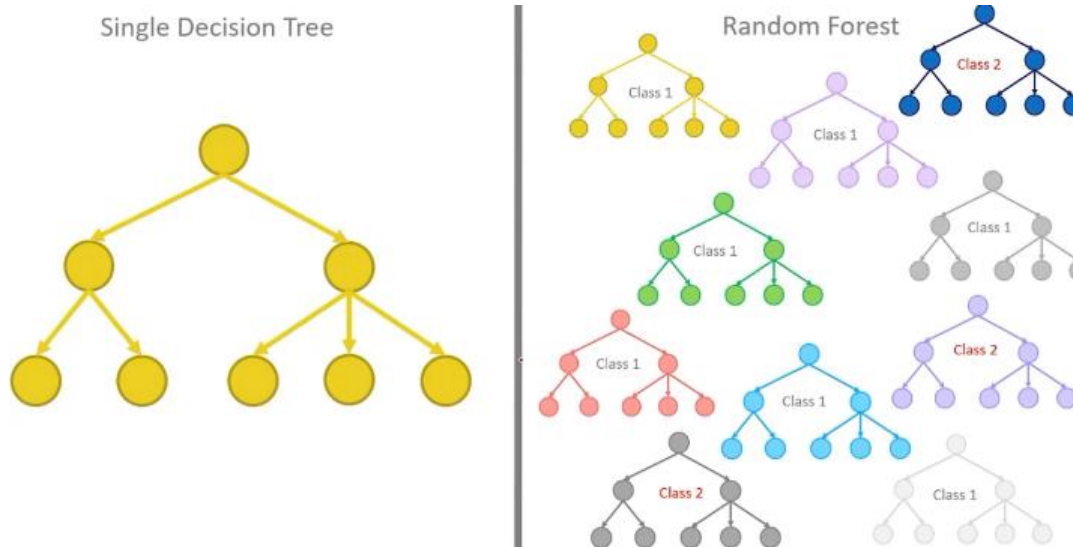
**Random Forest** -bu moslashuvchan, ishlatish uchun qulay bo'lgan mashinani o'rganish algoritmi, hatto giper parametrlarni sozlamasdan ham, ko'pincha ajoyib natija beradi. Shuningdek, u soddaligi va xilma-xilligi tufayli eng ko'p ishlatiladigan algoritmlardan biridir (u klassifikatsiya uchun ham, regressiya vazifalari uchun ham ishlatilishi mumkin).

### **Random Forest nima?**

Random Forest -bu Nazorat ostida o'rganish algoritmi. U quradigan" o'rmon" - bu Decision Treelar ansambli, odatda bagging usuli bilan o'rgatilgan. Bagging usulining umumiy g'oyasi shundaki, o'quv modellarining kombinatsiyasi umumiy natijani oshiradi.

Oddiy qilib aytganda: Random Forest bir nechta Decision Treelarini quradi va aniqroq va barqaror bashorat qilish uchun ularni birlashtiradi.

Random Forest da esa bir-nechta Decision Tree lardan iborat bo'ladi so'ngra har bir Decision Tree dan chiqqan natijaga qarab qaysi biri ko'proq ovoz olsa o'sha natija sifatida qaraladi.



### Random Forestdan foydalanish holatlari

- Moliya sohasida ishonchli qarzdorlar va potentsial firibgarlarni aniqlaydi
- Sogʻliqni saqlashda tibbiyot komponentlari va bemor ma'lumotlarini tekshiradi

• Mijozlarga elektron tijoratda mahsulotlar yoqadimi yoki yoʻqligini oʻlchaydi  
Masalan, moliya sohasida mijozlarni qarzlarni oʻz vaqtida toʻlash ehtimoli koʻproq yoki bank xizmatlaridan tez-tez foydalanish uchun foydalaniladi. Ushbu domenda u bankni aldash uchun firibgarlarni aniqlash uchun ham ishlatiladi. Savdoda algoritm aksiyalarning kelajakdagi xatti-harakatlarini aniqlash uchun ishlatilishi mumkin.

Random Forest quydagi kabi amalga oshiriladi.

**Random Forest modeli yordamida Bemorga dori tavsiya qilish tizimini yaratish.**

**Vazifa:** Tasavvur qiling, siz tadqiqot uchun ma'lumot to'playotgan tibbiy tadqiqotchisiz. Siz bir xil kasallikdan aziyat chekkan bemorlar haqida ma'lumot to'pladingiz. Davolash kursi davomida bemorlarga 5 xil doridan biri ijobiy ta'sir qildi.

**Maqsad:** xuddi shu kasallik bilan kasallangan kelajakdagi bemorga qaysi dori mos kelishi mumkinligini aniqlaydigan model yaratish. Ushbu muammo ko'p toifali klassifikator (multiclass classification) yordamida amalga oshiriladi.

#### 1-qadam

# kerakli kutubxonalarni chaqirib olamiz

```
import numpy as np
```

```
import pandas as pd
```

```
from sklearn.tree import DecisionTreeClassifier
```



```

from sklearn.ensemble import RandomForestClassifier
from sklearn.preprocessing import LabelEncoder
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report
from sklearn.metrics import confusion_matrix
from sklearn import tree
from matplotlib import pyplot as plt
from sklearn.model_selection import cross_val_predict
from sklearn.metrics import classification_report
import seaborn as sns
import matplotlib.pyplot as plt

```

### 2-qadam

```
# kerakli Dataset ni yuklab olish
```

```
df=pd.read_csv("https://raw.githubusercontent.com/JamshidSalimov/Ai-
Fayls/master/drug200.csv")
```

```
df
```

### Natija:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
0	23	F	HIGH	HIGH	25.355	drugY
1	47	M	LOW	HIGH	13.093	drugC
2	47	M	LOW	HIGH	10.114	drugC
3	28	F	NORMAL	HIGH	7.798	drugX
4	61	F	LOW	HIGH	18.043	drugY
...	...	...	...	...	...	...
195	56	F	LOW	HIGH	11.567	drugC
196	16	M	LOW	HIGH	12.006	drugC
197	52	M	NORMAL	HIGH	9.894	drugX
198	23	M	NORMAL	NORMAL	14.020	drugX
199	40	F	LOW	NORMAL	11.349	drugX

200 rows × 6 columns

### 3-qadam



```
# Matnli ustunlarga ishlov berish
```

```
encoder = LabelEncoder()
```

```
df['Sex'] = encoder.fit_transform(df['Sex'].values)
```

```
df['BP'] = encoder.fit_transform(df['BP'].values)
```

```
df['Cholesterol'] = encoder.fit_transform(df['Cholesterol'].values)
```

```
df.sample(10)
```

**Natija:**

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
149	22	1	1	0	8.151	drugC
196	16	1	1	0	12.006	drugC
38	39	0	2	1	9.709	drugX
72	24	0	2	0	10.605	drugX
16	69	1	1	1	11.455	drugX
33	65	0	0	1	31.876	drugY
177	25	1	2	0	19.011	drugY
121	15	1	0	1	17.206	drugY
171	45	1	1	1	10.017	drugX
54	68	0	0	1	10.189	drugB

**4-qadam.**

```
# X va y ni ajratib olish
```

```
X = df[['Age','Sex','BP','Cholesterol','Na_to_K']].values
```

```
y = df['Drug'].values
```

```
print(X[0:5])
```

```
print(y[0:5])
```

**Natija:**

```
[[23.  0.  0.  0. 25.355]
```

```
[47.  1.  1.  0. 13.093]
```



```
[47.  1.  1.  0. 10.114]
[28.  0.  2.  0.  7.798]
[61.  0.  1.  0. 18.043]]
['drugY' 'drugC' 'drugC' 'drugX' 'drugY']
```

### 5-qadam

```
# malumotlarni Train va Test uchun ajratib olish (Train/test split)
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.4,
random_state=20)
```

### 6-qadam

```
# Model yaratish va modelni o'qitish
```

```
Random_forest_model=RandomForestClassifier(n_estimators=20)
```

```
Random_forest_model.fit(X_train,y_train)
```

```
RandomForestClassifier
RandomForestClassifier(n_estimators=20)
```

### 7-qadam

```
# bashorat qiymatlarni topish
```

```
y_predict=Random_forest_model.predict(X_test)
```

```
# Modelni baholash
```

```
print(classification_report(y_test, y_predict))
```

**Natija:**

	precision	recall	f1-score	support
drugA	1.00	1.00	1.00	7
drugB	0.80	1.00	0.89	4
drugC	1.00	1.00	1.00	9
drugX	1.00	0.95	0.98	21
drugY	0.97	0.97	0.97	39
accuracy		0.97		80
macro avg	0.95	0.99	0.97	80
weighted avg	0.98	0.97	0.98	80

### 8-qadam

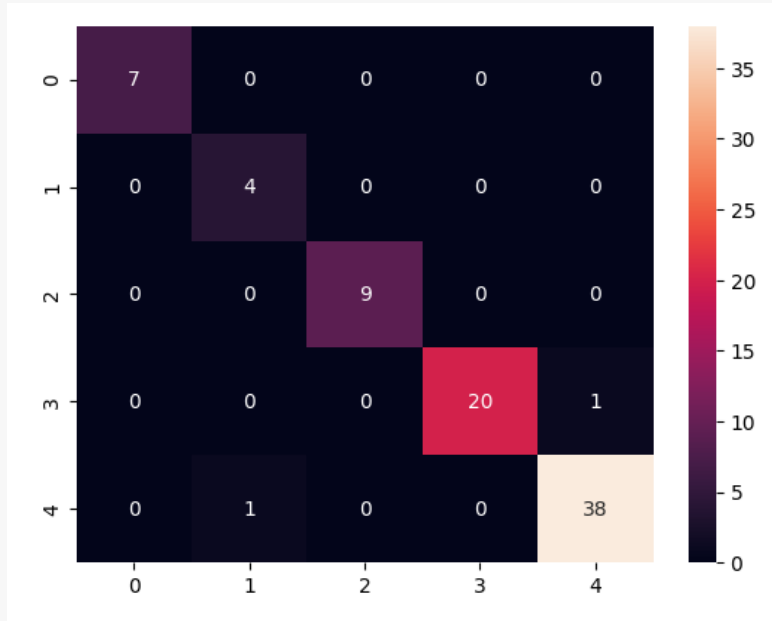
```
# Confusion matrix chizish
```

```
sns.heatmap(confusion_matrix(y_test, y_predict), annot=True)
```

```
plt.show()
```



**Natija:**



**9-qadam**

**# Cross-validation yordamida baholash**

```
predict = cross_val_predict(estimator = Random_forest_model, X = X, y = y,
cv = 5)
```

```
print("Classification Report: \n",classification_report(y, predict))
```

**Natija:**

Classification Report:

	precision	recall	f1-score	support
drugA	0.88	1.00	0.94	23
drugB	1.00	0.81	0.90	16
drugC	1.00	1.00	1.00	16
drugX	1.00	0.98	0.99	54
drugY	0.99	1.00	0.99	91
accuracy		0.98		200
macro avg	0.97	0.96	0.96	200
weighted avg	0.98	0.98	0.98	200

Random Forest algoritmining [geperparametrlari](#) haqida tiliq ma'lumot



## References

1. Amrullayevich K. A., Obid o'g'li S. J. ELEKTRON TALIM MUHITIDA TALABALARDA AXBOROT BILAN ISHLASH KOMPETENTLIKNI SHAKLLANTIRISH //International Journal of Contemporary Scientific and Technical Research. – 2022. – С. 641-645.
2. Obid o'g' A. S. J. et al. Numpy Library Capabilities. Vectorized Calculation In Numpy Va Type Of Information //Eurasian Research Bulletin. – 2022. – Т. 15. – С. 132-137.
3. Javlon X. et al. Классификатор движения рук с использованием биомиметического распознавания образов с помощью сверточных нейронных сетей с методом динамического порога для извлечения движения с использованием датчиков EF //Journal of new century innovations. – 2022. – Т. 19. – №. 6. – С. 352-357.
4. Фитратович В. и др. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ МНОГОФАЗНОЙ ФИЛЬТРАЦИИ В НЕФТЕГАЗОВОМ ПЛАСТЕ ПРИ ЕГО ЗАВОДНЕНИИ //INTERNATIONAL CONFERENCES ON LEARNING AND TEACHING. – 2022. – Т. 1. – №. 4. – С. 520-525.
5. Jamshid S. ENTROPY EVALUATION CRITERION IN DECISION TREE ALGORITHM EVALUATION //International Journal of Contemporary Scientific and Technical Research. – 2023. – С. 236-239.
6. Салимов Ж., Абулаева А. Классификации дерева в машинном обучении и гиперпараметрах //Информатика и инженерные технологии. – 2023. – Т. 1. – №. 1. – С. 71-79.
7. Obid o'g'li S. J., Nodir o'g'li X. A., Jasurjonovich B. J. SUPERVISED LEARNING REGRESSION ALGORITHM SIMPLE LINEAR REGRESSION //Academia Science Repository. – 2023. – Т. 4. – №. 04. – С. 69-76.
8. Javlon X. Salimov Jamshid Obid ogli." Классификатор движения рук с использованием биомиметического распознавания образов с помощью сверточных нейронных сетей с методом динамического порога для извлечения движения с использованием датчиков EF." //Journal of new century innovations. – 2022. – Т. 19. – С. 352-357.