Рубежный контроль №2

Коновалов Максим ИУ5-63Б

Вариант 6

In [1]:

```
import pandas as pd
import warnings
warnings.filterwarnings("ignore")
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np

from sklearn.metrics import mean_absolute_error, mean_squared_error, median_
from sklearn.preprocessing import MinMaxScaler
from sklearn.neighbors import KNeighborsRegressor, KNeighborsClassifier
from sklearn.model_selection import train_test_split, GridSearchCV
from sklearn.impute import SimpleImputer, MissingIndicator
from sklearn.preprocessing import LabelEncoder, OneHotEncoder, MinMaxScaler,
from sklearn.model_selection import StratifiedKFold
from sklearn.model_selection import cross_val_score
```

In [2]:

```
df = pd.read_csv('Admission_Predict.csv')
```

In [3]:

```
df.head()
```

Out[3]:

	Serial No.	GRE Score	TOEFL Score	University Rating	SOP	LOR	CGPA	Research	Chance of Admit
0	1	337	118	4	4.5	4.5	9.65	1	0.92
1	2	324	107	4	4.0	4.5	8.87	1	0.76
2	3	316	104	3	3.0	3.5	8.00	1	0.72
3	4	322	110	3	3.5	2.5	8.67	1	0.80
4	5	314	103	2	2.0	3.0	8.21	0	0.65

In [4]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 400 entries, 0 to 399
Data columns (total 9 columns):
 #   Column             Non-Null Count  Dtype  
---  ------             --------------  -----  
 0   Serial No.         400 non-null    int64  
 1   GRE Score          400 non-null    int64  
 2   TOEFL Score        400 non-null    int64  
 3   University Rating  400 non-null    int64  
 4   SOP                400 non-null    float64
 5   LOR                400 non-null    float64
 6   CGPA               400 non-null    float64
 7   Research           400 non-null    int64  
 8   Chance of Admit    400 non-null    float64
dtypes: float64(4), int64(5)
memory usage: 28.2 KB
```

In [5]:

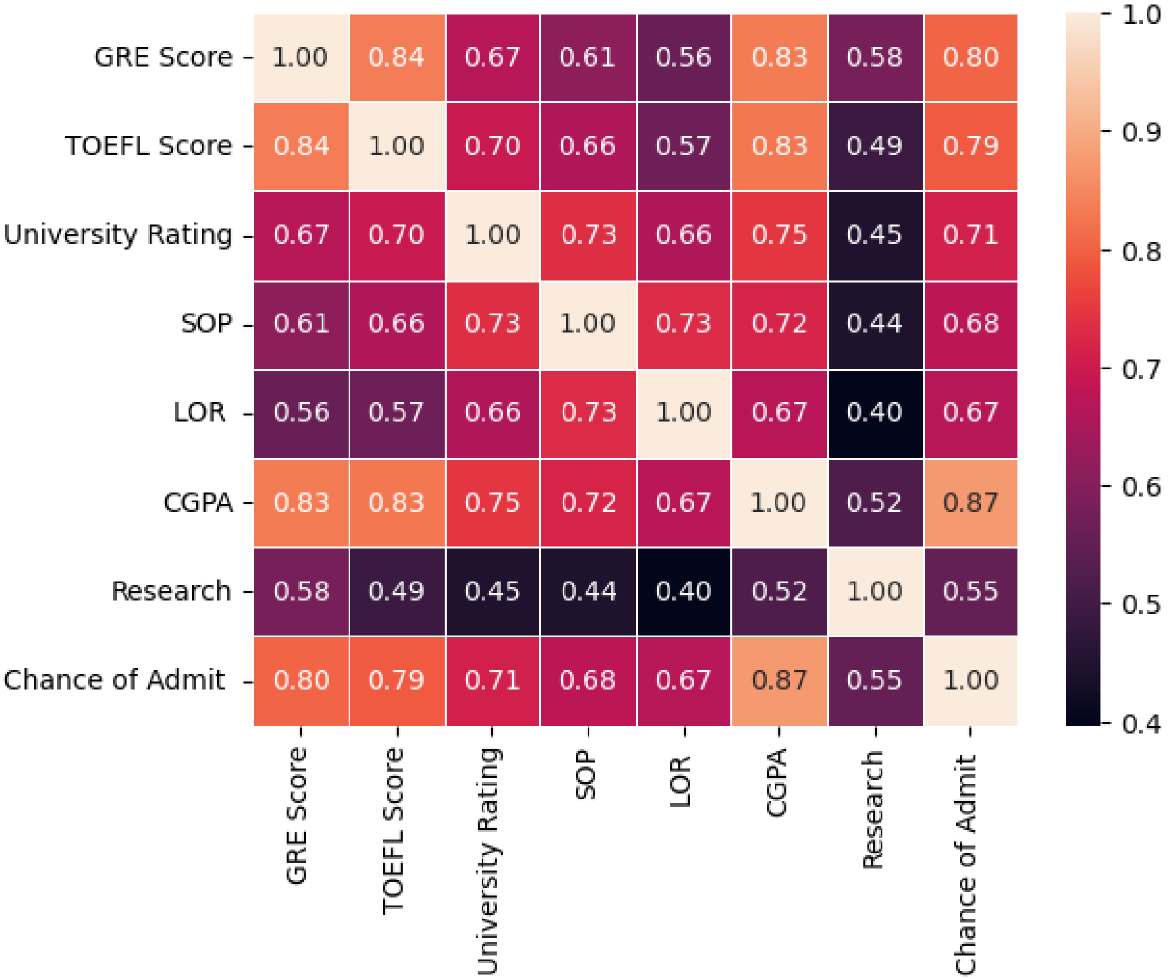
```
df = df.drop(['Serial No.'], axis=1)
df.isnull().sum()
```

Out[5]:

```
GRE Score            0
TOEFL Score          0
University Rating    0
SOP                  0
LOR                  0
CGPA                 0
Research             0
Chance of Admit      0
dtype: int64
```

In [6]:

```
corr = df.corr()
sns.heatmap(corr, linewidths=.5, annot=True, fmt=".2f")
plt.show()
```



1) С целевым признаком "Chance of Admit" наиболее коррелируют признаки "CGPA", "GRE Score" , "TOEFL Score". При построении модели машинного обучения перечисленные признаки будут наиболее информативными.

2) Стоит отметить корреляцию признаков "SOP" и "University Rating".

3) Можно построить модель машинного обучения на основе признаков "CGPA", "GRE Score", "TOEFL Score", "LOR", "Research". Первые 3 признака наиболее сильно повлияют на результат ввиду их высокой корреляции. Обученные модели позволят бакалаврам оценить свои возможности для поступления на магистратуру.

Разделение данных

Разделим данные на целевой столбец и признаки При построении предсказательных моделей исходные данные обычно разбиваются на обучающую ("training set") и контрольную ("test set") выборки. **Обучающая выборка** используется для построения математических отношений между некоторой переменной-откликом и предикторами, тогда как **контрольная (= "проверочная")** выборка служит для получения оценки прогнозных свойств модели на новых данных, т.е. данных, которые не были использованы для обучения модели.

In [7]:

```
X = df.drop(['Chance of Admit '], axis=1) #Наименования признаков
y = df['Chance of Admit '] # Значения признаков
```

In [8]:

```
# кодируем категориальные данные из строк в числа
le = LabelEncoder()
y = le.fit_transform(y)
```

In [9]:

```
X_train, X_test, y_train, y_test = train_test_split(X,y,test_size = 0.20, sh
```

In [10]:

```
# Размер обучающей выборки
X_train.shape, y_train.shape
```

Out[10]:

```
((320, 7), (320,))
```

In [11]:

```
# Размер тестовой выборки
X_test.shape, y_test.shape
```

Out[11]:

```
((80, 7), (80,))
```

дерево решений.

Оцените качество моделей с помощью трех подходящих для задачи метрик. Сравните качество полученных моделей.

In [35]:

```
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error,accuracy_score
from sklearn.svm import SVC, NuSVC, LinearSVC
from sklearn.tree import DecisionTreeClassifier, DecisionTreeRegressor, expo
from sklearn.model_selection import GridSearchCV
from sklearn import tree
from IPython.core.display import HTML
from sklearn.tree import export_text
from operator import itemgetter
```

Дерево решений: Допустим, у нас есть задача классификации и мы хотим обучить дерево решений на данных. Одним из гиперпараметров этой модели является максимальная глубина дерева. Мы можем использовать GridSearchCV для подбора оптимальной глубины дерева с помощью кросс-валидации.

In [36]:

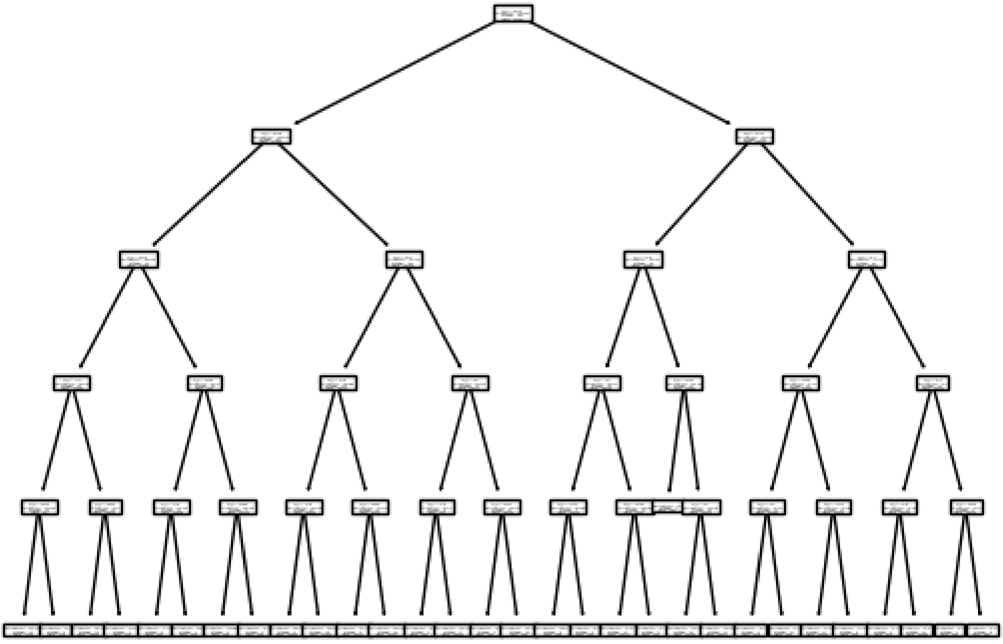
```
dt_none = DecisionTreeRegressor(max_depth=5)
dt_none.fit(X_train, y_train)
```

Out[36]:

```
DecisionTreeRegressor(max_depth=5)
```

In [37]:

```
tree.plot_tree(dt_none);
```



In [15]:

```
clf = DecisionTreeClassifier(random_state=1)
clf.fit(X_train, y_train)
```

Out[15]:

```
DecisionTreeClassifier(random_state=1)
```

In [16]:

```
def test_model(model):
    print("mean_absolute_error:",
          mean_absolute_error(y_test, model.predict(X_test)))
    print("median_absolute_error:",
          median_absolute_error(y_test, model.predict(X_test)))
    print("r2_score:",
          r2_score(y_test, model.predict(X_test)))
```

In [17]:

```
test_model(dt_none)
```

```
mean_absolute_error: 6.960165565428724
median_absolute_error: 5.19047619047619
r2_score: 0.5696142563467339
```

In [18]:

```
tree_rules = export_text(dt_none, feature_names=list(X.columns))
HTML('<pre>' + tree_rules + '</pre>')
```

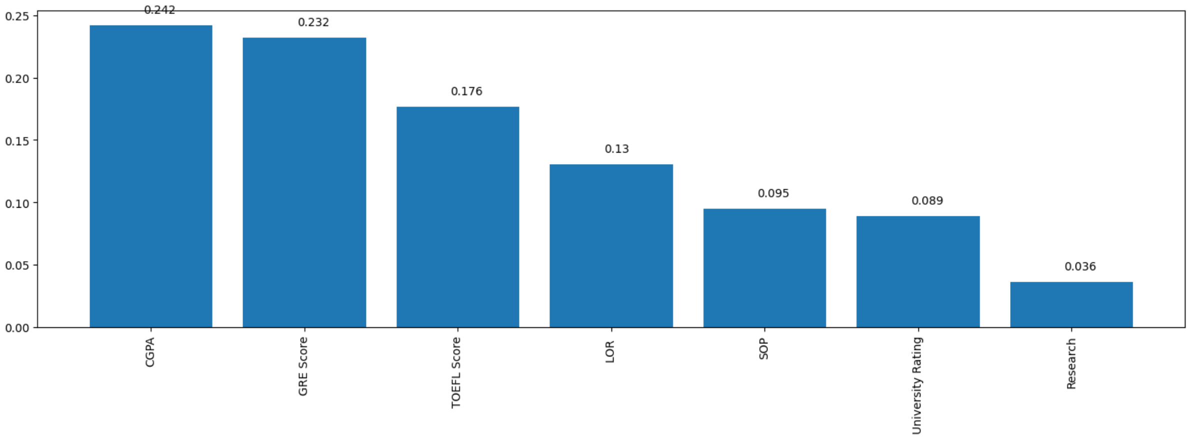
```
Out[18]: |--- CGPA <= 8.85
|   |--- CGPA <= 8.03
|   |   |--- GRE Score <= 307.50
|   |   |   |--- CGPA <= 7.67
|   |   |   |   |--- GRE Score <= 303.50
|   |   |   |   |   |--- GRE Score <= 298.50
|   |   |   |   |   |   |--- CGPA <= 7.64
|   |   |   |   |   |   |   |--- LOR  <= 2.75
|   |   |   |   |   |   |   |   |--- LOR  <= 1.75
|   |   |   |   |   |   |   |   |   |--- value: [11.00]
|   |   |   |   |   |   |   |   |--- LOR  >  1.75
|   |   |   |   |   |   |   |   |   |--- Research <= 0.50
|   |   |   |   |   |   |   |   |   |   |--- value: [7.75]
|   |   |   |   |   |   |   |   |   |--- Research >  0.50
|   |   |   |   |   |   |   |   |   |   |--- value: [6.00]
|   |   |   |   |   |   |   |--- LOR  >  2.75
|   |   |   |   |   |   |   |   |--- CGPA <= 7.37
|   |   |   |   |   |   |   |   |   |--- value: [9.00]
|   |   |   |   |   |   |   |   |--- CGPA >  7.37
|   |   |   |   |   |   |   |   |   |--- value: [15.00]
|   |   |   |   |   |   |--- CGPA >  7.64
|   |   |   |   |   |   |   |--- value: [19.00]
|   |   |   |   |   |--- GRE Score >  298.50
|   |   |   |   |   |   |--- TOEFL Score <= 98.00
|   |   |   |   |   |   |   |--- value: [2.00]
|   |   |   |   |   |   |--- TOEFL Score >  98.00
|   |   |   |   |   |   |   |--- value: [1.00]
|   |   |   |   |--- GRE Score >  303.50
|   |   |   |   |   |--- CGPA <= 7.58
|   |   |   |   |   |   |--- value: [14.00]
|   |   |   |   |   |--- CGPA >  7.58
|   |   |   |   |   |   |--- value: [20.00]
|   |   |   |--- CGPA >  7.67
|   |   |   |   |--- GRE Score <= 300.50
|   |   |   |   |   |--- SOP <= 3.50
|   |   |   |   |   |   |--- SOP <= 1.75
|   |   |   |   |   |   |   |--- GRE Score <= 298.50
|   |   |   |   |   |   |   |   |--- GRE Score <= 297.50
|   |   |   |   |   |   |   |   |   |--- value: [14.00]
|   |   |   |   |   |   |   |   |--- GRE Score >  297.50
|   |   |   |   |   |   |   |   |   |--- value: [16.00]
|   |   |   |   |   |   |   |--- GRE Score >  298.50
|   |   |   |   |   |   |   |   |--- University Rating <= 1.50
|   |   |   |   |   |   |   |   |   |--- SOP <= 1.25
|   |   |   |   |   |   |   |   |   |   |--- value: [20.00]
|   |   |   |   |   |   |   |   |   |--- SOP >  1.25
|   |   |   |   |   |   |   |   |   |   |--- value: [21.00]
|   |   |   |   |   |   |   |   |--- University Rating >  1.50
|   |   |   |   |   |   |   |   |   |--- CGPA <= 7.87
|   |   |   |   |   |   |   |   |   |   |--- value: [16.00]
|   |   |   |   |   |   |   |   |   |--- CGPA >  7.87
|   |   |   |   |   |   |   |   |   |   |--- value: [18.00]
|   |   |   |   |   |   |--- SOP >  1.75
|   |   |   |   |   |   |   |--- LOR  <= 3.25
|   |   |   |   |   |   |   |   |--- LOR  <= 2.75
|   |   |   |   |   |   |   |   |   |--- value: [23.00]
|   |   |   |   |   |   |   |   |--- LOR  >  2.75
|   |   |   |   |   |   |   |   |   |--- value: [22.00]
|   |   |   |   |   |   |   |--- LOR  >  3.25
|   |   |   |   |   |   |   |   |--- GRE Score <= 296.00
|   |   |   |   |   |   |   |   |   |--- value: [26.00]
|   |   |   |   |   |   |   |   |--- GRE Score >  296.00
|   |   |   |   |   |   |   |   |   |--- value: [30.00]
|   |   |   |   |   |--- SOP >  3.50
|   |   |   |   |   |   |--- value: [0.00]
|   |   |   |   |--- GRE Score >  300.50
|   |   |   |   |   |--- TOEFL Score <= 108.50
|   |   |   |   |   |   |--- SOP <= 1.25
|   |   |   |   |   |   |   |--- value: [18.00]
|   |   |   |   |   |   |--- SOP >  1.25
|   |   |   |   |   |   |   |--- SOP <= 1.75
|   |   |   |   |   |   |   |   |--- value: [4.00]
|   |   |   |   |   |   |   |--- SOP >  1.75
|   |   |   |   |   |   |   |   |--- TOEFL Score <= 98.50
|   |   |   |   |   |   |   |   |   |--- value: [6.00]
|   |   |   |   |   |   |   |   |--- TOEFL Score >  98.50
|   |   |   |   |   |   |   |   |   |--- University Rating <= 1.50
|   |   |   |   |   |   |   |   |   |   |--- value: [12.00]
|   |   |   |   |   |   |   |   |   |--- University Rating >  1.50
|   |   |   |   |   |   |   |   |   |   |--- value: [10.00]
|   |   |   |   |   |--- TOEFL Score >  108.50
|   |   |   |   |   |   |--- value: [24.00]
|   |   |--- GRE Score >  307.50
|   |   |   |--- GRE Score <= 315.50
|   |   |   |   |--- SOP <= 2.25
|   |   |   |   |   |--- CGPA <= 7.98
|   |   |   |   |   |   |--- GRE Score <= 311.00
|   |   |   |   |   |   |   |--- value: [16.00]
|   |   |   |   |   |   |--- GRE Score >  311.00
|   |   |   |   |   |   |   |--- LOR  <= 3.00
|   |   |   |   |   |   |   |   |--- TOEFL Score <= 101.50
|   |   |   |   |   |   |   |   |   |--- value: [20.00]
|   |   |   |   |   |   |   |   |--- TOEFL Score >  101.50
|   |   |   |   |   |   |   |   |   |--- value: [19.00]
|   |   |   |   |   |   |   |--- LOR  >  3.00
|   |   |   |   |   |   |   |   |--- value: [18.00]
|   |   |   |   |   |--- CGPA >  7.98
|   |   |   |   |   |   |--- University Rating <= 2.00
|   |   |   |   |   |   |   |--- value: [14.00]
|   |   |   |   |   |   |--- University Rating >  2.00
|   |   |   |   |   |   |   |--- value: [12.00]
|   |   |   |   |--- SOP >  2.25
|   |   |   |   |   |--- TOEFL Score <= 108.50
|   |   |   |   |   |   |--- LOR  <= 3.50
|   |   |   |   |   |   |   |--- GRE Score <= 310.00
|   |   |   |   |   |   |   |   |--- value: [24.00]
|   |   |   |   |   |   |   |--- GRE Score >  310.00
|   |   |   |   |   |   |   |   |--- CGPA <= 7.89
|   |   |   |   |   |   |   |   |   |--- value: [25.00]
|   |   |   |   |   |   |   |   |--- CGPA >  7.89
|   |   |   |   |   |   |   |   |   |--- TOEFL Score <= 103.00
|   |   |   |   |   |   |   |   |   |   |--- value: [27.00]
|   |   |   |   |   |   |   |   |   |--- TOEFL Score >  103.00
|   |   |   |   |   |   |   |   |   |   |--- value: [26.00]
|   |   |   |   |   |   |--- LOR  >  3.50
|   |   |   |   |   |   |   |--- SOP <= 3.25
|   |   |   |   |   |   |   |   |--- value: [30.00]
|   |   |   |   |   |   |   |--- SOP >  3.25
|   |   |   |   |   |   |   |   |--- value: [28.00]
|   |   |   |   |   |--- TOEFL Score >  108.50
|   |   |   |   |   |   |--- value: [8.00]
|   |   |   |--- GRE Score >  315.50
|   |   |   |   |--- University Rating <= 3.50
|   |   |   |   |   |--- CGPA <= 7.96
|   |   |   |   |   |   |--- value: [26.00]
|   |   |   |   |   |--- CGPA >  7.96
|   |   |   |   |   |   |--- SOP <= 3.50
|   |   |   |   |   |   |   |--- GRE Score <= 319.00
|   |   |   |   |   |   |   |   |--- value: [34.00]
|   |   |   |   |   |   |   |--- GRE Score >  319.00
|   |   |   |   |   |   |   |   |--- value: [32.00]
|   |   |   |   |   |   |--- SOP >  3.50
|   |   |   |   |   |   |   |--- value: [27.00]
|   |   |   |   |--- University Rating >  3.50
|   |   |   |   |   |--- value: [40.00]
|   |--- CGPA >  8.03
|   |   |--- CGPA <= 8.63
|   |   |   |--- LOR  <= 3.25
|   |   |   |   |--- LOR  <= 2.25
|   |   |   |   |   |--- SOP <= 2.25
|   |   |   |   |   |   |--- value: [4.00]
|   |   |   |   |   |--- SOP >  2.25
|   |   |   |   |   |   |--- CGPA <= 8.47
|   |   |   |   |   |   |   |--- GRE Score <= 321.00
|   |   |   |   |   |   |   |   |--- Research <= 0.50
|   |   |   |   |   |   |   |   |   |--- CGPA <= 8.33
|   |   |   |   |   |   |   |   |   |   |--- value: [27.33]
|   |   |   |   |   |   |   |   |   |--- CGPA >  8.33
|   |   |   |   |   |   |   |   |   |   |--- value: [24.00]
|   |   |   |   |   |   |   |   |--- Research >  0.50
|   |   |   |   |   |   |   |   |   |--- value: [23.00]
|   |   |   |   |   |   |   |--- GRE Score >  321.00
|   |   |   |   |   |   |   |   |--- value: [32.00]
|   |   |   |   |   |   |--- CGPA >  8.47
|   |   |   |   |   |   |   |--- value: [15.00]
|   |   |   |   |--- LOR  >  2.25
|   |   |   |   |   |--- CGPA <= 8.20
|   |   |   |   |   |   |--- CGPA <= 8.19
|   |   |   |   |   |   |   |--- University Rating <= 2.50
|   |   |   |   |   |   |   |   |--- SOP <= 1.75
|   |   |   |   |   |   |   |   |   |--- value: [33.00]
|   |   |   |   |   |   |   |   |--- SOP >  1.75
|   |   |   |   |   |   |   |   |   |--- GRE Score <= 308.50
|   |   |   |   |   |   |   |   |   |   |--- value: [25.75]
|   |   |   |   |   |   |   |   |   |--- GRE Score >  308.50
|   |   |   |   |   |   |   |   |   |   |--- value: [20.67]
|   |   |   |   |   |   |   |--- University Rating >  2.50
|   |   |   |   |   |   |   |   |--- CGPA <= 8.07
|   |   |   |   |   |   |   |   |   |--- value: [29.00]
|   |   |   |   |   |   |   |   |--- CGPA >  8.07
|   |   |   |   |   |   |   |   |   |--- CGPA <= 8.14
|   |   |   |   |   |   |   |   |   |   |--- value: [34.00]
|   |   |   |   |   |   |   |   |   |--- CGPA >  8.14
|   |   |   |   |   |   |   |   |   |   |--- value: [33.00]
|   |   |   |   |   |   |--- CGPA >  8.19
|   |   |   |   |   |   |   |--- GRE Score <= 305.50
|   |   |   |   |   |   |   |   |--- value: [16.00]
|   |   |   |   |   |   |   |--- GRE Score >  305.50
|   |   |   |   |   |   |   |   |--- SOP <= 3.25
|   |   |   |   |   |   |   |   |   |--- value: [11.00]
|   |   |   |   |   |   |   |   |--- SOP >  3.25
|   |   |   |   |   |   |   |   |   |--- value: [9.00]
|   |   |   |   |   |--- CGPA >  8.20
|   |   |   |   |   |   |--- GRE Score <= 322.50
|   |   |   |   |   |   |   |--- University Rating <= 3.50
|   |   |   |   |   |   |   |   |--- SOP <= 3.75
|   |   |   |   |   |   |   |   |   |--- Research <= 0.50
|   |   |   |   |   |   |   |   |   |   |--- value: [29.84]
|   |   |   |   |   |   |   |   |   |--- Research >  0.50
|   |   |   |   |   |   |   |   |   |   |--- value: [33.00]
|   |   |   |   |   |   |   |   |--- SOP >  3.75
|   |   |   |   |   |   |   |   |   |--- value: [18.00]
|   |   |   |   |   |   |   |--- University Rating >  3.50
|   |   |   |   |   |   |   |   |--- CGPA <= 8.54
|   |   |   |   |   |   |   |   |   |--- CGPA <= 8.47
|   |   |   |   |   |   |   |   |   |   |--- value: [19.00]
|   |   |   |   |   |   |   |   |   |--- CGPA >  8.47
|   |   |   |   |   |   |   |   |   |   |--- value: [23.00]
|   |   |   |   |   |   |   |   |--- CGPA >  8.54
|   |   |   |   |   |   |   |   |   |--- value: [32.00]
|   |   |   |   |   |   |--- GRE Score >  322.50
|   |   |   |   |   |   |   |--- value: [7.00]
|   |   |   |--- LOR  >  3.25
|   |   |   |   |--- GRE Score <= 306.50
|   |   |   |   |   |--- CGPA <= 8.58
|   |   |   |   |   |   |--- TOEFL Score <= 105.50
|   |   |   |   |   |   |   |--- GRE Score <= 301.50
|   |   |   |   |   |   |   |   |--- CGPA <= 8.33
|   |   |   |   |   |   |   |   |   |--- value: [30.00]
|   |   |   |   |   |   |   |   |--- CGPA >  8.33
|   |   |   |   |   |   |   |   |   |--- value: [31.00]
|   |   |   |   |   |   |   |--- GRE Score >  301.50
|   |   |   |   |   |   |   |   |--- LOR  <= 3.75
|   |   |   |   |   |   |   |   |   |--- value: [26.00]
|   |   |   |   |   |   |   |   |--- LOR  >  3.75
|   |   |   |   |   |   |   |   |   |--- LOR  <= 4.50
|   |   |   |   |   |   |   |   |   |   |--- value: [28.00]
|   |   |   |   |   |   |   |   |   |--- LOR  >  4.50
|   |   |   |   |   |   |   |   |   |   |--- value: [27.00]
|   |   |   |   |   |   |--- TOEFL Score >  105.50
|   |   |   |   |   |   |   |--- Research <= 0.50
|   |   |   |   |   |   |   |   |--- GRE Score <= 302.50
|   |   |   |   |   |   |   |   |   |--- value: [26.00]
|   |   |   |   |   |   |   |   |--- GRE Score >  302.50
|   |   |   |   |   |   |   |   |   |--- value: [25.00]
|   |   |   |   |   |   |   |--- Research >  0.50
|   |   |   |   |   |   |   |   |--- value: [24.00]
|   |   |   |   |   |--- CGPA >  8.58
|   |   |   |   |   |   |--- value: [18.00]
|   |   |   |   |--- GRE Score >  306.50
|   |   |   |   |   |--- University Rating <= 3.50
|   |   |   |   |   |   |--- TOEFL Score <= 99.50
|   |   |   |   |   |   |   |--- value: [38.00]
|   |   |   |   |   |   |--- TOEFL Score >  99.50
|   |   |   |   |   |   |   |--- TOEFL Score <= 108.50
|   |   |   |   |   |   |   |   |--- GRE Score <= 307.50
|   |   |   |   |   |   |   |   |   |--- value: [40.00]
|   |   |   |   |   |   |   |   |--- GRE Score >  307.50
|   |   |   |   |   |   |   |   |   |--- TOEFL Score <= 107.50
|   |   |   |   |   |   |   |   |   |   |--- value: [30.81]
|   |   |   |   |   |   |   |   |   |--- TOEFL Score >  107.50
|   |   |   |   |   |   |   |   |   |   |--- value: [24.50]
|   |   |   |   |   |   |   |--- TOEFL Score >  108.50
|   |   |   |   |   |   |   |   |--- University Rating <= 2.50
|   |   |   |   |   |   |   |   |   |--- TOEFL Score <= 109.50
|   |   |   |   |   |   |   |   |   |   |--- value: [33.00]
|   |   |   |   |   |   |   |   |   |--- TOEFL Score >  109.50
|   |   |   |   |   |   |   |   |   |   |--- value: [34.00]
|   |   |   |   |   |   |   |   |--- University Rating >  2.50
|   |   |   |   |   |   |   |   |   |--- value: [37.00]
|   |   |   |   |   |--- University Rating >  3.50
|   |   |   |   |   |   |--- CGPA <= 8.48
|   |   |   |   |   |   |   |--- CGPA <= 8.36
|   |   |   |   |   |   |   |   |--- value: [39.00]
|   |   |   |   |   |   |   |--- CGPA >  8.36
|   |   |   |   |   |   |   |   |--- SOP <= 3.50
|   |   |   |   |   |   |   |   |   |--- value: [40.00]
|   |   |   |   |   |   |   |   |--- SOP >  3.50
|   |   |   |   |   |   |   |   |   |--- value: [41.00]
|   |   |   |   |   |   |--- CGPA >  8.48
|   |   |   |   |   |   |   |--- value: [33.00]
|   |   |--- CGPA >  8.63
|   |   |   |--- University Rating <= 2.50
|   |   |   |   |--- LOR  <= 2.75
|   |   |   |   |   |--- GRE Score <= 322.50
|   |   |   |   |   |   |--- value: [43.00]
|   |   |   |   |   |--- GRE Score >  322.50
|   |   |   |   |   |   |--- value: [37.00]
|   |   |   |   |--- LOR  >  2.75
|   |   |   |   |   |--- Research <= 0.50
|   |   |   |   |   |   |--- LOR  <= 3.50
|   |   |   |   |   |   |   |--- value: [28.00]
|   |   |   |   |   |   |--- LOR  >  3.50
|   |   |   |   |   |   |   |--- value: [29.00]
|   |   |   |   |   |--- Research >  0.50
|   |   |   |   |   |   |--- value: [19.00]
|   |   |   |--- University Rating >  2.50
|   |   |   |   |--- SOP <= 2.25
|   |   |   |   |   |--- value: [26.00]
|   |   |   |   |--- SOP >  2.25
|   |   |   |   |   |--- Research <= 0.50
|   |   |   |   |   |   |--- GRE Score <= 324.50
|   |   |   |   |   |   |   |--- GRE Score <= 318.50
|   |   |   |   |   |   |   |   |--- GRE Score <= 316.50
|   |   |   |   |   |   |   |   |   |--- CGPA <= 8.64
|   |   |   |   |   |   |   |   |   |   |--- value: [30.00]
|   |   |   |   |   |   |   |   |   |--- CGPA >  8.64
|   |   |   |   |   |   |   |   |   |   |--- value: [36.27]
|   |   |   |   |   |   |   |   |--- GRE Score >  316.50
|   |   |   |   |   |   |   |   |   |--- TOEFL Score <= 111.00
|   |   |   |   |   |   |   |   |   |   |--- value: [26.50]
|   |   |   |   |   |   |   |   |   |--- TOEFL Score >  111.00
|   |   |   |   |   |   |   |   |   |   |--- value: [33.00]
|   |   |   |   |   |   |   |--- GRE Score >  318.50
|   |   |   |   |   |   |   |   |--- SOP <= 4.25
|   |   |   |   |   |   |   |   |   |--- value: [41.00]
|   |   |   |   |   |   |   |   |--- SOP >  4.25
|   |   |   |   |   |   |   |   |   |--- value: [38.00]
|   |   |   |   |   |   |--- GRE Score >  324.50
|   |   |   |   |   |   |   |--- value: [14.00]
|   |   |   |   |   |--- Research >  0.50
|   |   |   |   |   |   |--- GRE Score <= 311.50
|   |   |   |   |   |   |   |--- value: [30.00]
|   |   |   |   |   |   |--- GRE Score >  311.50
|   |   |   |   |   |   |   |--- GRE Score <= 315.00
|   |   |   |   |   |   |   |   |--- TOEFL Score <= 107.50
|   |   |   |   |   |   |   |   |   |--- value: [43.00]
|   |   |   |   |   |   |   |   |--- TOEFL Score >  107.50
|   |   |   |   |   |   |   |   |   |--- value: [44.00]
|   |   |   |   |   |   |   |--- GRE Score >  315.00
|   |   |   |   |   |   |   |   |--- CGPA <= 8.72
|   |   |   |   |   |   |   |   |   |--- LOR  <= 3.50
|   |   |   |   |   |   |   |   |   |   |--- value: [42.50]
|   |   |   |   |   |   |   |   |   |--- LOR  >  3.50
|   |   |   |   |   |   |   |   |   |   |--- value: [35.50]
|   |   |   |   |   |   |   |   |--- CGPA >  8.72
|   |   |   |   |   |   |   |   |   |--- LOR  <= 1.75
|   |   |   |   |   |   |   |   |   |   |--- value: [32.00]
|   |   |   |   |   |   |   |   |   |--- LOR  >  1.75
|   |   |   |   |   |   |   |   |   |   |--- value: [36.10]
|--- CGPA >  8.85
|   |--- CGPA <= 9.22
|   |   |--- SOP <= 3.75
|   |   |   |--- Research <= 0.50
|   |   |   |   |--- TOEFL Score <= 111.00
|   |   |   |   |   |--- GRE Score <= 317.00
|   |   |   |   |   |   |--- University Rating <= 2.50
|   |   |   |   |   |   |   |--- value: [32.00]
|   |   |   |   |   |   |--- University Rating >  2.50
|   |   |   |   |   |   |   |--- value: [36.00]
|   |   |   |   |   |--- GRE Score >  317.00
|   |   |   |   |   |   |--- CGPA <= 9.08
|   |   |   |   |   |   |   |--- GRE Score <= 322.50
|   |   |   |   |   |   |   |   |--- value: [44.00]
|   |   |   |   |   |   |   |--- GRE Score >  322.50
|   |   |   |   |   |   |   |   |--- value: [42.00]
|   |   |   |   |   |   |--- CGPA >  9.08
|   |   |   |   |   |   |   |--- value: [40.00]
|   |   |   |   |--- TOEFL Score >  111.00
|   |   |   |   |   |--- SOP <= 3.25
|   |   |   |   |   |   |--- value: [23.00]
|   |   |   |   |   |--- SOP >  3.25
|   |   |   |   |   |   |--- value: [17.00]
|   |   |   |--- Research >  0.50
|   |   |   |   |--- TOEFL Score <= 109.50
|   |   |   |   |   |--- value: [30.00]
|   |   |   |   |--- TOEFL Score >  109.50
|   |   |   |   |   |--- CGPA <= 9.08
|   |   |   |   |   |   |--- CGPA <= 9.03
|   |   |   |   |   |   |   |--- CGPA <= 8.97
|   |   |   |   |   |   |   |   |--- CGPA <= 8.86
|   |   |   |   |   |   |   |   |   |--- value: [44.00]
|   |   |   |   |   |   |   |   |--- CGPA >  8.86
|   |   |   |   |   |   |   |   |   |--- CGPA <= 8.93
|   |   |   |   |   |   |   |   |   |   |--- value: [42.00]
|   |   |   |   |   |   |   |   |   |--- CGPA >  8.93
|   |   |   |   |   |   |   |   |   |   |--- value: [41.00]
|   |   |   |   |   |   |   |--- CGPA >  8.97
|   |   |   |   |   |   |   |   |--- CGPA <= 8.99
|   |   |   |   |   |   |   |   |   |--- value: [46.00]
|   |   |   |   |   |   |   |   |--- CGPA >  8.99
|   |   |   |   |   |   |   |   |   |--- value: [44.00]
|   |   |   |   |   |   |--- CGPA >  9.03
|   |   |   |   |   |   |   |--- CGPA <= 9.05
|   |   |   |   |   |   |   |   |--- value: [38.00]
|   |   |   |   |   |   |   |--- CGPA >  9.05
|   |   |   |   |   |   |   |   |--- value: [36.00]
|   |   |   |   |   |--- CGPA >  9.08
|   |   |   |   |   |   |--- SOP <= 3.25
|   |   |   |   |   |   |   |--- value: [45.00]
|   |   |   |   |   |   |--- SOP >  3.25
|   |   |   |   |   |   |   |--- value: [46.00]
|   |   |--- SOP >  3.75
|   |   |   |--- GRE Score <= 310.50
|   |   |   |   |--- value: [28.00]
|   |   |   |--- GRE Score >  310.50
|   |   |   |   |--- University Rating <= 4.50
|   |   |   |   |   |--- TOEFL Score <= 116.50
|   |   |   |   |   |   |--- CGPA <= 9.11
|   |   |   |   |   |   |   |--- GRE Score <= 322.00
|   |   |   |   |   |   |   |   |--- GRE Score <= 313.50
|   |   |   |   |   |   |   |   |   |--- LOR  <= 4.00
|   |   |   |   |   |   |   |   |   |   |--- value: [41.00]
|   |   |   |   |   |   |   |   |   |--- LOR  >  4.00
|   |   |   |   |   |   |   |   |   |   |--- value: [40.00]
|   |   |   |   |   |   |   |   |--- GRE Score >  313.50
|   |   |   |   |   |   |   |   |   |--- TOEFL Score <= 109.50
|   |   |   |   |   |   |   |   |   |   |--- value: [45.00]
|   |   |   |   |   |   |   |   |   |--- TOEFL Score >  109.50
|   |   |   |   |   |   |   |   |   |   |--- value: [48.00]
|   |   |   |   |   |   |   |--- GRE Score >  322.00
|   |   |   |   |   |   |   |   |--- CGPA <= 8.88
|   |   |   |   |   |   |   |   |   |--- value: [38.00]
|   |   |   |   |   |   |   |   |--- CGPA >  8.88
|   |   |   |   |   |   |   |   |   |--- CGPA <= 9.04
|   |   |   |   |   |   |   |   |   |   |--- value: [43.00]
|   |   |   |   |   |   |   |   |   |--- CGPA >  9.04
|   |   |   |   |   |   |   |   |   |   |--- value: [41.00]
|   |   |   |   |   |   |--- CGPA >  9.11
|   |   |   |   |   |   |   |--- TOEFL Score <= 114.50
|   |   |   |   |   |   |   |   |--- GRE Score <= 319.00
|   |   |   |   |   |   |   |   |   |--- value: [42.00]
|   |   |   |   |   |   |   |   |--- GRE Score >  319.00
|   |   |   |   |   |   |   |   |   |--- CGPA <= 9.11
|   |   |   |   |   |   |   |   |   |   |--- value: [51.00]
|   |   |   |   |   |   |   |   |   |--- CGPA >  9.11
|   |   |   |   |   |   |   |   |   |   |--- value: [46.56]
|   |   |   |   |   |   |   |--- TOEFL Score >  114.50
|   |   |   |   |   |   |   |   |--- value: [40.00]
|   |   |   |   |   |--- TOEFL Score >  116.50
|   |   |   |   |   |   |--- CGPA <= 9.13
|   |   |   |   |   |   |   |--- TOEFL Score <= 118.50
|   |   |   |   |   |   |   |   |--- value: [49.00]
|   |   |   |   |   |   |   |--- TOEFL Score >  118.50
|   |   |   |   |   |   |   |   |--- value: [48.00]
|   |   |   |   |   |   |--- CGPA >  9.13
|   |   |   |   |   |   |   |--- value: [52.00]
|   |   |   |   |--- University Rating >  4.50
|   |   |   |   |   |--- Research <= 0.50
|   |   |   |   |   |   |--- value: [40.00]
|   |   |   |   |   |--- Research >  0.50
|   |   |   |   |   |   |--- CGPA <= 9.07
|   |   |   |   |   |   |   |--- TOEFL Score <= 111.00
|   |   |   |   |   |   |   |   |--- value: [49.00]
|   |   |   |   |   |   |   |--- TOEFL Score >  111.00
|   |   |   |   |   |   |   |   |--- value: [48.00]
|   |   |   |   |   |   |--- CGPA >  9.07
|   |   |   |   |   |   |   |--- TOEFL Score <= 117.50
|   |   |   |   |   |   |   |   |--- CGPA <= 9.21
|   |   |   |   |   |   |   |   |   |--- CGPA <= 9.17
|   |   |   |   |   |   |   |   |   |   |--- value: [50.50]
|   |   |   |   |   |   |   |   |   |--- CGPA >  9.17
|   |   |   |   |   |   |   |   |   |   |--- value: [49.00]
|   |   |   |   |   |   |   |   |--- CGPA >  9.21
|   |   |   |   |   |   |   |   |   |--- value: [54.00]
|   |   |   |   |   |   |   |--- TOEFL Score >  117.50
|   |   |   |   |   |   |   |   |--- value: [54.00]
|   |--- CGPA >  9.22
|   |   |--- CGPA <= 9.47
|   |   |   |--- TOEFL Score <= 114.50
|   |   |   |   |--- LOR  <= 3.50
|   |   |   |   |   |--- value: [47.00]
|   |   |   |   |--- LOR  >  3.50
|   |   |   |   |   |--- CGPA <= 9.31
|   |   |   |   |   |   |--- LOR  <= 4.75
|   |   |   |   |   |   |   |--- GRE Score <= 322.50
|   |   |   |   |   |   |   |   |--- value: [53.00]
|   |   |   |   |   |   |   |--- GRE Score >  322.50
|   |   |   |   |   |   |   |   |--- TOEFL Score <= 112.00
|   |   |   |   |   |   |   |   |   |--- value: [50.00]
|   |   |   |   |   |   |   |   |--- TOEFL Score >  112.00
|   |   |   |   |   |   |   |   |   |--- TOEFL Score <= 113.50
|   |   |   |   |   |   |   |   |   |   |--- value: [51.00]
|   |   |   |   |   |   |   |   |   |--- TOEFL Score >  113.50
|   |   |   |   |   |   |   |   |   |   |--- value: [52.00]
|   |   |   |   |   |   |--- LOR  >  4.75
|   |   |   |   |   |   |   |--- value: [48.00]
|   |   |   |   |   |--- CGPA >  9.31
|   |   |   |   |   |   |--- University Rating <= 4.50
|   |   |   |   |   |   |   |--- value: [55.00]
|   |   |   |   |   |   |--- University Rating >  4.50
|   |   |   |   |   |   |   |--- LOR  <= 4.75
|   |   |   |   |   |   |   |   |--- LOR  <= 4.25
|   |   |   |   |   |   |   |   |   |--- value: [53.00]
|   |   |   |   |   |   |   |   |--- LOR  >  4.25
|   |   |   |   |   |   |   |   |   |--- value: [51.00]
|   |   |   |   |   |   |   |--- LOR  >  4.75
|   |   |   |   |   |   |   |   |--- value: [55.00]
|   |   |   |--- TOEFL Score >  114.50
|   |   |   |   |--- TOEFL Score <= 117.50
|   |   |   |   |   |--- LOR  <= 3.25
|   |   |   |   |   |   |--- value: [52.00]
|   |   |   |   |   |--- LOR  >  3.25
|   |   |   |   |   |   |--- SOP <= 4.25
|   |   |   |   |   |   |   |--- SOP <= 3.75
|   |   |   |   |   |   |   |   |--- value: [53.00]
|   |   |   |   |   |   |   |--- SOP >  3.75
|   |   |   |   |   |   |   |   |--- TOEFL Score <= 115.50
|   |   |   |   |   |   |   |   |   |--- value: [54.00]
|   |   |   |   |   |   |   |   |--- TOEFL Score >  115.50
|   |   |   |   |   |   |   |   |   |--- value: [55.00]
|   |   |   |   |   |   |--- SOP >  4.25
|   |   |   |   |   |   |   |--- CGPA <= 9.40
|   |   |   |   |   |   |   |   |--- CGPA <= 9.32
|   |   |   |   |   |   |   |   |   |--- value: [56.00]
|   |   |   |   |   |   |   |   |--- CGPA >  9.32
|   |   |   |   |   |   |   |   |   |--- value: [55.00]
|   |   |   |   |   |   |   |--- CGPA >  9.40
|   |   |   |   |   |   |   |   |--- value: [56.00]
|   |   |   |   |--- TOEFL Score >  117.50
|   |   |   |   |   |--- University Rating <= 4.50
|   |   |   |   |   |   |--- CGPA <= 9.38
|   |   |   |   |   |   |   |--- value: [52.00]
|   |   |   |   |   |   |--- CGPA >  9.38
|   |   |   |   |   |   |   |--- value: [53.00]
|   |   |   |   |   |--- University Rating >  4.50
|   |   |   |   |   |   |--- SOP <= 4.75
|   |   |   |   |   |   |   |--- value: [55.00]
|   |   |   |   |   |   |--- SOP >  4.75
|   |   |   |   |   |   |   |--- value: [54.00]
|   |   |--- CGPA >  9.47
|   |   |   |--- CGPA <= 9.73
|   |   |   |   |--- LOR  <= 4.25
|   |   |   |   |   |--- SOP <= 4.25
|   |   |   |   |   |   |--- value: [52.00]
|   |   |   |   |   |--- SOP >  4.25
|   |   |   |   |   |   |--- value: [51.00]
|   |   |   |   |--- LOR  >  4.25
|   |   |   |   |   |--- TOEFL Score <= 118.50
|   |   |   |   |   |   |--- SOP <= 4.75
|   |   |   |   |   |   |   |--- CGPA <= 9.59
|   |   |   |   |   |   |   |   |--- value: [56.00]
|   |   |   |   |   |   |   |--- CGPA >  9.59
|   |   |   |   |   |   |   |   |--- CGPA <= 9.66
|   |   |   |   |   |   |   |   |   |--- value: [54.00]
|   |   |   |   |   |   |   |   |--- CGPA >  9.66
|   |   |   |   |   |   |   |   |   |--- value: [55.00]
|   |   |   |   |   |   |--- SOP >  4.75
|   |   |   |   |   |   |   |--- value: [56.00]
|   |   |   |   |   |--- TOEFL Score >  118.50
|   |   |   |   |   |   |--- SOP <= 4.75
|   |   |   |   |   |   |   |--- value: [56.00]
|   |   |   |   |   |   |--- SOP >  4.75
|   |   |   |   |   |   |   |--- CGPA <= 9.58
|   |   |   |   |   |   |   |   |--- value: [58.00]
|   |   |   |   |   |   |   |--- CGPA >  9.58
|   |   |   |   |   |   |   |   |--- value: [57.00]
|   |   |   |--- CGPA >  9.73
|   |   |   |   |--- GRE Score <= 332.00
|   |   |   |   |   |--- value: [56.00]
|   |   |   |   |--- GRE Score >  332.00
|   |   |   |   |   |--- TOEFL Score <= 117.50
|   |   |   |   |   |   |--- value: [58.00]
|   |   |   |   |   |--- TOEFL Score >  117.50
|   |   |   |   |   |   |--- GRE Score <= 333.50
|   |   |   |   |   |   |   |--- value: [58.00]
|   |   |   |   |   |   |--- GRE Score >  333.50
|   |   |   |   |   |   |   |--- value: [59.00]
```

In [19]:

```
def draw_feature_importances(tree_model, X_dataset, figsize=(18,5)):
    """
    Вывод важности признаков в виде графика
    """
    # Сортировка значений важности признаков по убыванию
    list_to_sort = list(zip(X_dataset.columns.values, tree_model.feature_imp
    sorted_list = sorted(list_to_sort, key=itemgetter(1), reverse = True)
    # Названия признаков
    labels = [x for x,_ in sorted_list]
    # Важности признаков
    data = [x for _,x in sorted_list]
    # Вывод графика
    fig, ax = plt.subplots(figsize=figsize)
    ind = np.arange(len(labels))
    plt.bar(ind, data)
    plt.xticks(ind, labels, rotation='vertical')
    # Вывод значений
    for a,b in zip(ind, data):
        plt.text(a-0.05, b+0.01, str(round(b,3)))
    plt.show()
    return labels, data
```

In [20]:

```
dt_fl, dt_fd = draw_feature_importances(clf, X_train)
```



Произведите для каждой модели подбор одного гиперпараметра с использованием GridSearchCV и кросс-валидации.

In [21]:
```
tree = DecisionTreeClassifier()

param_grid = {'max_depth': [2, 4, 6, 8, 10],
              'min_samples_split': [2, 4, 6, 8, 10],
              'min_samples_leaf': [1, 2, 3, 4, 5]}

grid_search = GridSearchCV(tree, param_grid=param_grid, cv=5)

grid_search.fit(X_train, y_train)

accuracy_tree = grid_search.best_estimator_.score(X_test,y_test)

print("Наилучшие параметры: {} ", grid_search.best_params_)
print("Оценка точности на кросс-валидации: {:.2f}".format(grid_search.best_score_
print(accuracy_tree)
```

```
Наилучшие параметры: {}  {'max_depth': 6, 'min_samples_leaf': 3, 'min_samples_
split': 4}
Оценка точности на кросс-валидации: 0.11
0.0875
```

Сравните качество полученных моделей с качеством моделей

In [22]:
```
models = [['DecisionTree :',DecisionTreeRegressor()]]
```

In [23]:
```
print('Вывод 1')
for name,model in models:
    model = model
    model.fit(X_train, y_train)
    predictions = model.predict(X_test)
    print(name, (np.sqrt(mean_squared_error(y_test, predictions))))
```

```
Вывод 1
DecisionTree : 9.419394885023134
```

In [24]:
```
models = [['DecisionTree :',DecisionTreeRegressor(max_depth = 6, min_samples
          ['Linear Regression :', LinearRegression(normalize = True)],
          ['SVC :', SVC(C = 1, kernel = 'linear')]]
```

In [25]:
```
print('Вывод 2')
for name,model in models:
    model = model
    model.fit(X_train, y_train)
    predictions = model.predict(X_test)
    print(name, (np.sqrt(mean_squared_error(y_test, predictions))))
```

```
Вывод 2
DecisionTree : 7.718895862285397
Linear Regression : 6.252859728579746
SVC : 7.565546906866681
```

Бэггинг. В этом случае однородные модели обучают на разных наборах данных и объединяют. Получают прогноз путём усреднения. Если использовать в качестве слабого ученика деревья решений, то получится случайный лес RandomForestClassifier / RandomForestRegressor.

In [26]:

```
from sklearn.ensemble import RandomForestClassifier, GradientBoostingClassif
from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import cross_val_predict
from sklearn.ensemble import RandomForestClassifier
from sklearn.ensemble import GradientBoostingClassifier
```

Модель бэггинга - случайный лес (Random Forest):

In [27]:

```
# Создаем модель случайного леса с 100 деревьями
rf_model = RandomForestClassifier(n_estimators=100)

# Обучаем модель на тренировочных данных
rf_model.fit(X_train, y_train)

# Оцениваем качество модели на тестовых данных
accuracy = rf_model.score(X_test, y_test)
print('Accuracy: {:.2f}%'.format(accuracy*100))
```

```
Accuracy: 11.25%
```

Произведите для каждой модели подбор значений одного гиперпараметра. В зависимости от используемой библиотеки можно применять функцию GridSearchCV, использовать перебор параметров в цикле, или использовать другие методы.

In [28]:

```
model = RandomForestClassifier()

param_grid = {
    'n_estimators': [200, 700],
    'max_features': ['auto', 'sqrt', 'log2']
}

grid_search = GridSearchCV(model, param_grid=param_grid, cv=5)

grid_search.fit(X_train, y_train)

accuracy_RandomForestClassifier = grid_search.best_estimator_.score(X_test,y

print("Наилучшие параметры: {} ", grid_search.best_params_)
print("Оценка точности на кросс-валидации: {:.2f}".format(grid_search.best_score_
print(accuracy_tree)
```

```
Наилучшие параметры: {}  {'max_features': 'sqrt', 'n_estimators': 200}
Оценка точности на кросс-валидации: 0.12
0.0875
```

In [29]:

```
models = [
           ['RandomForestClassifier :',RandomForestClassifier()]
          ]
```

In [30]:

```
for name,model in models:
    model = model
    model.fit(X_train, y_train)
    predictions = model.predict(X_test)
    print(name, (np.sqrt(mean_squared_error(y_test, predictions))))
```

```
RandomForestClassifier : 7.440598094239467
```

In [31]:

```
models = [['RandomForestClassifier :',RandomForestClassifier(max_features =
         ]
```

In [32]:

```
for name,model in models:
    model = model
    model.fit(X_train, y_train)
    predictions = model.predict(X_test)
    print(name, (np.sqrt(mean_squared_error(y_test, predictions))))
```

```
RandomForestClassifier : 7.402702209328699
```