|  |  |  |
| --- | --- | --- |
| Лабораторная работа № 6. Ансамблевый подход | Студент | Бокова О. Д. |
| Группа | ИВТ-363 |
| Преподаватель | Фокин Р.О. |
| Оценка |  |
| Дата | 17.04.2023 |

Для выбранного или произвольного набора данных из репозитория

UC Irvine Machine Learning Repository необходимо выполнить классификацию с использованием ансамблевого подхода – бэггинга или бустинга на

выбор, произвести интерпретацию полученного результата.

Код:

import pandas as pd

from sklearn.ensemble import GradientBoostingClassifier

from sklearn.model\_selection import train\_test\_split

from sklearn.metrics import accuracy\_score

import matplotlib.pyplot as plt

# Загрузка данных

df = pd.read\_csv("biofam.csv", delimiter = ';')

# Удаление столбца id

df.drop(columns=['idhous'], inplace=True)

df.drop(columns=['nat\_1\_02'], inplace=True)

df.drop(columns=['plingu02'], inplace=True)

df.drop(columns=['p02r01'], inplace=True)

df.drop(columns=['p02r04'], inplace=True)

df.drop(columns=['cspfaj'], inplace=True)

df.drop(columns=['cspmoj'], inplace=True)

# Замена значений столбца sex

df['sex'].replace({0: 'woman', 1: 'man'}, inplace=True)

# Разделение на признаки и целевую переменную

X = df.drop(columns=['sex'])

y = df['sex']

# Разделение выборки на обучающую и тестовую

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.3, random\_state=42)

# Создание модели бустинга

model = GradientBoostingClassifier(random\_state=42)

# Обучение модели

model.fit(X\_train, y\_train)

# Предсказание на тестовой выборке

y\_pred = model.predict(X\_test)

# Оценка качества модели

accuracy = accuracy\_score(y\_test, y\_pred)

print("Accuracy: {:.2f}%".format(accuracy\*100))

# График важности признаков

importance = model.feature\_importances\_

plt.bar([x for x in X.columns], importance)

plt.xticks(rotation='vertical')

plt.show()

Результат:

