|  |  |  |
| --- | --- | --- |
| Лабораторная работа № 5. Деревья решений | 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 помощью метода дерева решений, построить само дерево, произвести интерпретацию полученного результата.

Код:

import pandas as pd

from sklearn.tree import DecisionTreeClassifier

from sklearn.tree import export\_graphviz

import graphviz

# 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']

# Создание модели дерева решений

clf = DecisionTreeClassifier(random\_state=42)

# Тренировка модели

clf.fit(X, y)

# Создание файла .dot для построения графа дерева решений

dot\_file = export\_graphviz(clf, out\_file='tree.dot', feature\_names=X.columns, class\_names=y.unique(), filled=True)

# Отображение графа дерева решений

with open('tree.dot') as f:

    dot\_graph = f.read()

graphviz.Source(dot\_graph)

Результат:

