

Министерство науки и высшего образования Российской Федерации  
Муромский институт (филиал)  
Федерального государственного бюджетного образовательного учреждения  
высшего образования  
«Владимирский государственный университет  
имени Александра Григорьевича и Николая Григорьевича Столетовых»

Факультет \_\_\_\_\_ ИТР \_\_\_\_\_

Кафедра \_\_\_\_\_ ПИН \_\_\_\_\_

## ЛАБОРАТОРНАЯ РАБОТА №5

По \_\_\_\_\_ Технологиям машинного обучения \_\_\_\_\_

Тема \_\_\_\_\_ Деревья решений \_\_\_\_\_

Руководитель

Захаров А.А.

\_\_\_\_\_  
(фамилия, инициалы)

\_\_\_\_\_  
(подпись)

\_\_\_\_\_  
(дата)

Студент \_\_\_\_\_ ПИН - 121 \_\_\_\_\_

\_\_\_\_\_  
(группа)

Ермилов М.В.

\_\_\_\_\_  
(фамилия, инициалы)

\_\_\_\_\_  
(подпись)

\_\_\_\_\_  
(дата)

Муром 2024

## Лабораторная работа №5

**Цель работы:** изучить метод деревьев решений, освоить его практическую реализацию.

### Ход работы:

Импорт библиотек:

```
from sklearn.datasets import load_iris
from sklearn.tree import DecisionTreeClassifier
from sklearn import tree
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
```

Загрузка, изучение и преобразование набора данных Iris:

Данный набор данных предназначен для построения модели классификации. Данные о 150 экземплярах ириса, по 50 экземпляров из трёх видов – Ирис щетиный (Iris setosa), Ирис виргинский (Iris virginica) и Ирис разноцветный (Iris versicolor). Для каждого экземпляра измерялись четыре характеристики (в сантиметрах):

- 1) длина наружной доли околоцветника (sepal length);
- 2) ширина наружной доли околоцветника (sepal width);
- 3) длина внутренней доли околоцветника (petal length);
- 4) ширина внутренней доли околоцветника (petal width).

```
data = load_iris() # Загрузка набора данных Iris
data # Просмотр содержимого данных
data.target_names # Получение списка имен целевых классов
df = pd.DataFrame(data.data)
df.head()
df.columns = data.feature_names
df.head()
targets = data.target_names
print(targets)
df['Species'] = data.target
df.head()
```

|           |      |              |         |      |                    |                 |      |        |
|-----------|------|--------------|---------|------|--------------------|-----------------|------|--------|
|           |      |              |         |      | МИВУ 09.03.04 - 05 |                 |      |        |
| Изм.      | Лист | № докум.     | Подпись | Дата |                    |                 |      |        |
| Разраб.   |      | Ермилов М.В. |         |      | Деревья решений    | Лит.            | Лист | Листов |
| Провер.   |      | Захаров А.А. |         |      |                    |                 | 2    | 4      |
| Реценз.   |      |              |         |      |                    | МИ ВлГУ ПИН-121 |      |        |
| Н. Контр. |      |              |         |      |                    |                 |      |        |
| Утверд.   |      |              |         |      |                    |                 |      |        |

|   | sepal length (cm) | sepal width (cm) | petal length (cm) | petal width (cm) | Species |
|---|-------------------|------------------|-------------------|------------------|---------|
| 0 | 5.1               | 3.5              | 1.4               | 0.2              | 0       |
| 1 | 4.9               | 3.0              | 1.4               | 0.2              | 0       |
| 2 | 4.7               | 3.2              | 1.3               | 0.2              | 0       |
| 3 | 4.6               | 3.1              | 1.5               | 0.2              | 0       |
| 4 | 5.0               | 3.6              | 1.4               | 0.2              | 0       |

Рисунок 1 – полученные данные

X и Y матрицы:

```
X = data.data # Признаки объектов
y = data.target # Метки классов
```

Разделение на обучающий и тестовый наборы данных:

```
from sklearn.model_selection import train_test_split
X_train,X_test,y_train,y_test = train_test_split(X,y,test_size=0.2,random_state=43)
cls = DecisionTreeClassifier()
cls.fit(X_train, y_train) #Обучение дерева решений на обучающих данных
```

Визуализация дерева решений:

```
fig = plt.figure(figsize=(25,20))
_ = tree.plot_tree(cls,
feature_names=iris.feature_names,
class_names=iris.target_names,
filled=True) # Визуализация обученного дерева решений cls с подписями и цветами
```

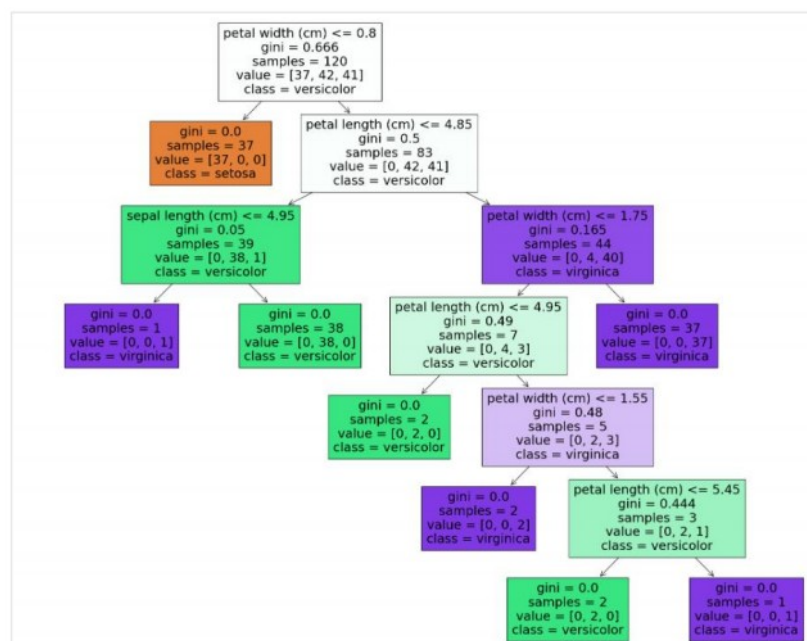


Рисунок 2 – дерево решений

Оценка модели:

```
predictions = cls.predict(X_test)
data.target_names[predictions]
from sklearn.metrics import accuracy_score
result = accuracy_score(y_test, predictions)
print("Точность модели : ", result)
```

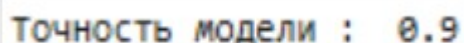


Рисунок 3 – точность модели

Получение результата для некоторых данных:

```
X_pred = [4.6, 3.1, 1.5, 0.1]
y_pred = cls.predict([X_pred])
print("Prediction is: {}".format(targets[y_pred]))
```




Рисунок 4 – полученный результат - Ирис щетинистый (Iris setosa)

**Вывод:** В ходе выполнения данной лабораторной работы были изучены методы деревьев решений на примере его практической реализации.

|      |      |          |         |      |                    |      |
|------|------|----------|---------|------|--------------------|------|
|      |      |          |         |      | МИВУ 09.02.03 – 05 | Лист |
| Изм. | Лист | № докум. | Подпись | Дата |                    | 4    |