

Министерство образования Республики Беларусь
Учреждение образования
«Брестский Государственный технический университет»
Кафедра ИИТ

Отчёт по лабораторной работе №2

Выполнил:
Студент 3 курса
Группы АС-65
Романюк Д. А.
Проверил:
Крощенко А. А.

Брест 2025

Цель: Изучить применение линейной и логистической регрессии для решения практических задач. Научиться обучать модели, оценивать их качество с помощью соответствующих метрик и интерпретировать результаты.

Вариант 5

• Регрессия (Предсказание итоговой оценки студента) 1. Student Performance 2. Предсказать итоговую оценку (G3) 3. Задания: загрузите данные, выбрав в класса качестве признаков studytime, failures, G1, G2; обучите модель линейной регрессии; рассчитайте MAE и R^2 ; визуализируйте зависимость G3 от G2 (оценка предыдущий период) с линией регрессии.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_absolute_error, r2_score

data = pd.read_csv("C:\\ОМО\\2lab\\student-mat.csv", sep=';')

features = ['studytime', 'failures', 'G1', 'G2']
target = 'G3'

X = data[features]
y = data[target]

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)

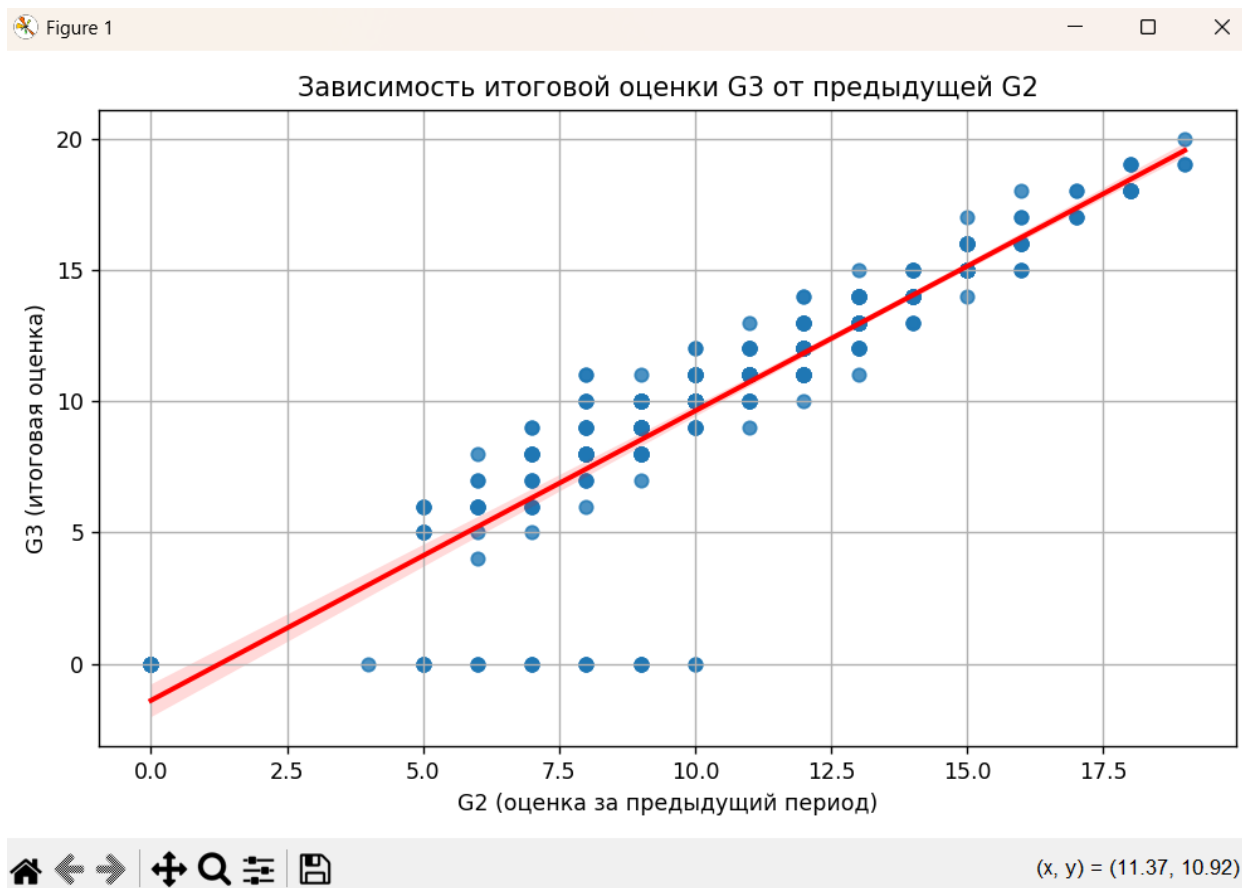
model = LinearRegression()
model.fit(X_train, y_train)

y_pred = model.predict(X_test)

mae = mean_absolute_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)

print("Mean Absolute Error (MAE):", round(mae, 3))
print("R2 Score:", round(r2, 3))

plt.figure(figsize=(8, 5))
sns.regplot(x=data['G2'], y=data['G3'], line_kws={"color": "red"})
plt.title('Зависимость итоговой оценки G3 от предыдущей G2')
plt.xlabel('G2 (оценка за предыдущий период)')
plt.ylabel('G3 (итоговая оценка)')
plt.grid(True)
plt.tight_layout()
plt.show()
```



• Классификация (Определение съедобности грибов) 1. Mushroom Classification за 2. Определить, является ли гриб ядовитым (class = 'p') или съедобным ('e') 3. Задания: загрузите данные. Все признаки являются категориальными, примените к ним One-Hot Encoding; обучите модель логистической регрессии; рассчитайте Accuracy, Precision и Recall; постройте матрицу ошибок.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import OneHotEncoder
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score, precision_score, recall_score,
confusion_matrix

url = "https://archive.ics.uci.edu/ml/machine-learning-
databases/mushroom/agaricus-lepiota.data"
columns = [
    'class', 'cap-shape', 'cap-surface', 'cap-color', 'bruises', 'odor', 'gill-
attachment',
    'gill-spacing', 'gill-size', 'gill-color', 'stalk-shape', 'stalk-root',
```

```

        'stalk-surface-above-ring', 'stalk-surface-below-ring', 'stalk-color-
above-ring',
        'stalk-color-below-ring', 'veil-type', 'veil-color', 'ring-number', 'ring-
type',
        'spore-print-color', 'population', 'habitat'
    ]
    data = pd.read_csv(url, header=None, names=columns)

    X = data.drop('class', axis=1)
    y = data['class'].map({'e': 0, 'p': 1})

    encoder = OneHotEncoder(sparse_output=False)
    X_encoded = encoder.fit_transform(X)

    X_train, X_test, y_train, y_test = train_test_split(X_encoded, y,
test_size=0.2, random_state=42)

    model = LogisticRegression(max_iter=200)
    model.fit(X_train, y_train)

    y_pred = model.predict(X_test)

    acc = accuracy_score(y_test, y_pred)
    prec = precision_score(y_test, y_pred)
    rec = recall_score(y_test, y_pred)

    print("Accuracy:", round(acc, 3))
    print("Precision:", round(prec, 3))
    print("Recall:", round(rec, 3))

    cm = confusion_matrix(y_test, y_pred)

    plt.figure(figsize=(6, 4))
    sns.heatmap(cm, annot=True, fmt='d', cmap='Blues',
                xticklabels=['Съедобный', 'Ядовитый'],
                yticklabels=['Съедобный', 'Ядовитый'])
    plt.xlabel('Предсказано')
    plt.ylabel('Истинное значение')
    plt.title('Матрица ошибок')
    plt.tight_layout()
    plt.show()

```



Вывод: Изучил применение линейной и логистической регрессии для решения практических задач. Научился обучать модели, оценивать их качество с помощью соответствующих метрик и интерпретировать результаты.