



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ

Информатика и системы управления (ИУ)

КАФЕДРА

Информационная безопасность (ИУ8)

Отчёт по лабораторной работе № 6

«Использование модели CART в задаче двухфакторной биометрической аутентификации»

Вариант: 1

Студент:

Александров Алексей Николаевич, группа ИУ8-94
(5 курс)

(подпись, дата)

Преподаватель:

профессор кафедры ИУ8
Басараб Михаил Алексеевич

(подпись, дата)

Москва, 2023 г.

Цель работы

Исследовать модель системы двухфакторной бинарной классификации «Свой- Чужой» с использованием аппарата деревьев регрессии и классификации.

Ход работы

В лабораторной работе № 2 были сгенерированы синтетические выборки классов «Свой- Чужой» с коррелированными параметрами (см. рисунок 1). Углы поворота выборок: 22.3° и 76.4° соответственно.

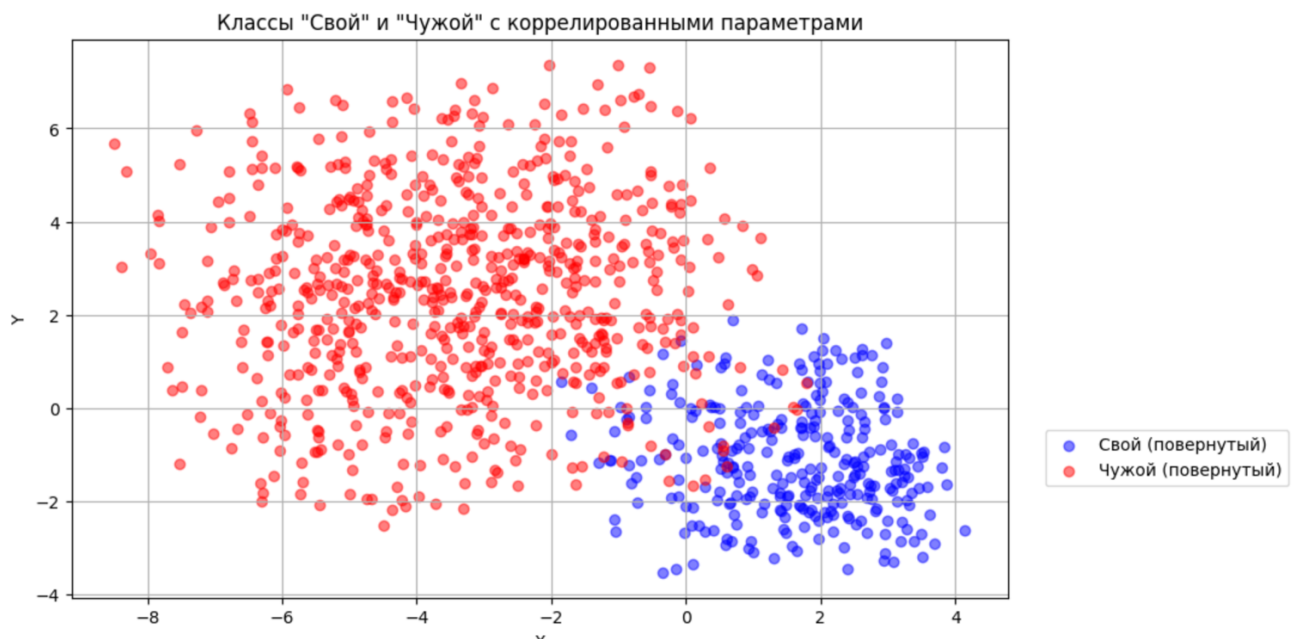


Рисунок 1 — Гистограммы классов «Свой»-«Чужой» с коррелированными параметрами

В данной работе на основе этих выборок была обучена модель двухфакторного бинарного классификатора на основе CART. По варианту была выбран энтропийный критерий информативности, который можно вычислить по формуле ниже. На рисунке 2 преведены границы классификации (пороги), а на рисунке 3 – дерево принятия решений глубины 2 для полученного классификатора.

$$E = -(p_1 * \log(p_1) + p_2 * \log(p_2))$$



Рисунок 2 — Распределения классов «Свой», «Чужой» и построенные пороги для дерева CART глубины 2

Матрица ошибок:

	Чужой	Свой
Чужой	683	17
Свой	35	265

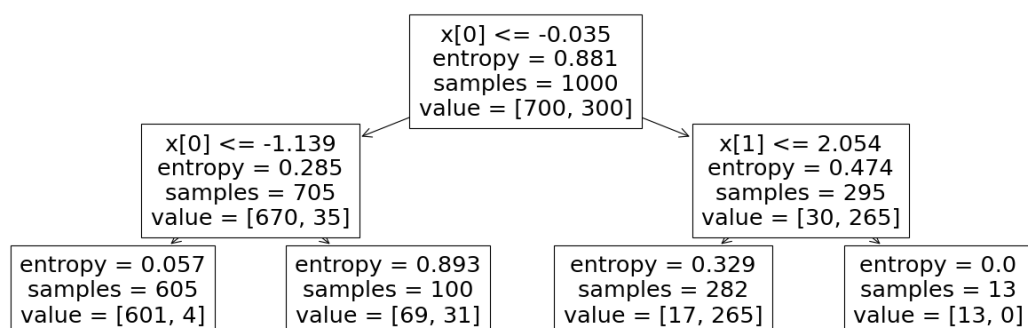


Рисунок 3 — Дерево CART глубины 2 для полученной модели классификатора

Листинг обучения модели:

```
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import confusion_matrix
import seaborn as sns

# Создание обучающей выборки и меток
X_train = np.concatenate((np.column_stack((samples_own_x_rotated2, samples_own_y_rotated2)),
                             np.column_stack((samples_foreign_x_rotated2, samples_foreign_y_rotated2))))
y_train = np.concatenate((np.ones(samples_own_x_rotated2.shape[0]),
                             np.zeros(samples_foreign_x_rotated2.shape[0])))

# Обучение дерева решений
clf = DecisionTreeClassifier(
    max_depth=2,
    criterion='entropy'
)
model = clf.fit(X_train, y_train)

# Предсказание меток для обучающей выборки
y_pred = model.predict(X_train)

# Построение матрицы ошибок
conf_matrix = confusion_matrix(y_train, y_pred)

# Построение матрицы ошибок с использованием seaborn для лучшей визуализации
plt.figure(figsize=(6, 4))
sns.heatmap(conf_matrix, annot=True, fmt='d', cmap='Blues', xticklabels=['Чужой', 'Свой'],
            yticklabels=['Чужой', 'Свой'])
plt.xlabel('Предсказанный класс')
plt.ylabel('Истинный класс')
plt.title('Матрица ошибок')
plt.show()

# Построение графика с разделением выборок прямыми
plt.figure(figsize=(10, 6))

# Разделение области на сетку
xx, yy = np.meshgrid(np.linspace(X_train[:, 0].min() - 1, X_train[:, 0].max() + 1, 100),
                     np.linspace(X_train[:, 1].min() - 1, X_train[:, 1].max() + 1, 100))

# Предсказание для каждой точки сетки
Z = model.predict(np.c_[xx.ravel(), yy.ravel()])
Z = Z.reshape(xx.shape)

# Рисование контуров решающей области
plt.contourf(xx, yy, Z, alpha=0.3, cmap='coolwarm')

# Рисование точек выборки
plt.scatter(samples_own_x_rotated2, samples_own_y_rotated2, label='Свой (повернутый)', color='blue',
            alpha=0.5)
```

```
plt.scatter(samples_foreign_x_rotated2, samples_foreign_y_rotated2, label='Чужой (повернутый)',
            color='red', alpha=0.5)

plt.xlabel('X')
plt.ylabel('Y')
plt.legend(loc=(1.04, 0.2))
plt.title('Классификация методом CART с границами решений')
plt.grid(True)
plt.show()
```

Вывод

В ходе работы была исследована модель системы классификации «Свой-Чужой» с использованием CART является хорошим алгоритмом дерева решений для задач классификации и регрессии. Деревья решений являются интуитивно понятными и интерпретируемыми моделями, что делает их популярным выбором в машинном обучении.