



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

КАФЕДРА КОМПЬЮТЕРНЫЕ СИСТЕМЫ И СЕТИ (ИУ6)

НАПРАВЛЕНИЕ ПОДГОТОВКИ 09.04.01 Информатика и вычислительная техника

МАГИСТЕРСКАЯ ПРОГРАММА 09.04.01/07 Интеллектуальные системы анализа,
обработки и интерпретации больших данных

О Т Ч Е Т

по лабораторной работе № 3

Название: Интеллектуальный анализ данных об объекте мониторинга

Дисциплина: Дистанционный мониторинг сложных систем и процессов

Студент

ИУ6-12М

(Группа)

(Подпись,
дата)

Д.С. Каткова

(И.О. Фамилия)

Преподаватель

(Подпись,
дата)

Ю.А. Вишневская

(И.О. Фамилия)

Москва, 2023

Цель работы – Разработка и исследование алгоритма принятия решения в условиях неопределенности.

Задание: При выполнении лабораторной работы необходимо формализовать задачу принятия решения по распознаванию текущего образа в системе мониторинга, разработать алгоритм распознавания и классификации образа, реализовать интеллектуальный подход при принятии решения в условиях неопределенности. Затем необходимо составить отчет по лабораторной работе с полученными результатами.

Исходные данные: набор эталонов и текущий (распознаваемый) вектор.

В качестве эталонов выбрать три образа, например, группу из трех символов (рис. 1).

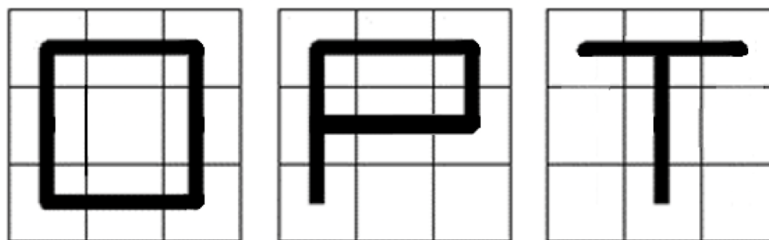


Рис 1 - Эталонные символы

Здесь эталонные символы вписаны в матрицу 3×3 . Таким образом, вектор признаков имеет длину, равную 9.

В качестве текущего вектора выбрать другое написание одного из символов.

Далее предположим, что синергетическая модель имеет следующие значения параметров: $M = 3$; $C = 3$; $B = 1 - C = -2$; $\lambda_1 = \lambda_2 = \lambda_3 = 1$.

Порядок выполнения работы:

1. Выбрать идентификационные признаки для эталонных векторов. Такими признаками может быть, например, наличие горизонтальных и/или вертикальных линий в ячейке матрицы;
2. Выполнить предварительную обработку (кодирование признаков), например: наличие линии – 1, отсутствие – 0;
3. Составить эталонные векторы, содержащие признаки объектов;
4. Рассчитать сопряженные векторы V_k^+ ;

5. Вычислить начальные значения параметров порядка;
6. Решать уравнение классификации;
7. Построить графики изменения значений параметров порядка и по максимальному значению классифицировать текущий вектор.

Ход работы

Эталонные символы на рисунке 1 при переводе в вектора дают значения:

$$O = [11, 10, 11, 01, 00, 01, 11, 10, 11]$$

$$P = [11, 10, 11, 11, 10, 11, 01, 00, 00]$$

$$T = [10, 11, 10, 00, 01, 00, 00, 01, 00]$$

Т.к. данные вектора являются комбинацией двух цифр, то принимаем, что значения векторов сформированы в двоичной системе. Для работы с векторами в вектора-эталонны были переведены поэлементно в двоичную систему.

$$V1' = [3, 2, 3, 1, 0, 1, 3, 2, 3]$$

$$V2' = [3, 2, 3, 3, 2, 3, 1, 0, 0]$$

$$V3' = [2, 3, 2, 0, 1, 0, 0, 1, 0]$$

Далее производится нормализация и централизация векторов (рисунок 2), в итоге формируются следующие вектора.

Централизованные векторы:									
[0.	1.	1.	-1.	-2.	-1.	1.	0.	1.]
[1.11111111	0.11111111	1.11111111	1.11111111	0.11111111	1.11111111	1.11111111	-0.88888889	-1.88888889]
[1.	2.	1.	-1.	0.	-1.	-1.	0.	-1.]
Нормализованные векторы:									
[0.	0.125	0.125	-0.125	-0.25	-0.125	0.125	0.	0.125]
[0.119	0.012	0.119	0.119	0.012	0.119	-0.095	-0.202	-0.202]
[0.125	0.25	0.125	-0.125	0.	-0.125	-0.125	0.	-0.125]

Рисунок 2 – Нормализация и централизация векторов

Полученные вектора должны удовлетворять двум условиям (рисунок 3):

- 1) $\sum_{j=1}^N v_{kj} = 0$ сумма всех компонентов вектора k равна нулю;
- 2) $\sum_i^N |v_{ki}| = 1$ сумма модулей значений компонентов равна нулю.

```
Проверка векторов на удовлетворение условиям
Для 1го вектора:
0.0
1.0
Для 2го вектора:
0.0
1.0
Для 3го вектора:
0.0
1.0
```

Рисунок 3 – Проверка векторов на удовлетворение условиям

Далее было проведено вычисление сопряженных векторов. Для этого вводим эталонную матрицу, состоящую из векторов-эталонов (рисунок 4).

```
Эталонная матрица:
[[ 0.    0.119  0.125]
 [ 0.125  0.012  0.25 ]
 [ 0.125  0.119  0.125]
 [-0.125  0.119 -0.125]
 [-0.25   0.012  0.   ]
 [-0.125  0.119 -0.125]
 [ 0.125 -0.095 -0.125]
 [ 0.    -0.202  0.   ]
 [ 0.125 -0.202 -0.125]]
```

Рисунок 4 – Эталонная матрица

Далее была найдена матрица W – произведение эталонной матрицы на ее транспонированную (рисунок 5).

```
Произведение эталонной матрицы на ее транспонированную (W)
[[ 0.156 -0.054  0.047]
 [-0.054  0.148  0.04 ]
 [ 0.047  0.04   0.156]]
```

Рисунок 5 – Нахождение матрицы W

После была найдена обратная матрица A (рисунок 6).

```
Обратная матрица (A)
[[ 9.017  4.3   -3.811]
 [ 4.3    9.316 -3.686]
 [-3.811 -3.686  8.491]]
```

Рисунок 6 – Нахождение обратной матрицы A

Используя матрицу A находим сопряженные векторы (рисунок 7).

Присоединенные векторы:									
[0.036	0.226	1.163	-0.139	-2.203	-0.139	1.194	-0.87	0.733]
[0.648	-0.273	1.186	1.032	-0.964	1.032	0.111	-1.885	-0.887]
[0.623	1.603	0.146	-1.024	0.909	-1.024	-1.187	0.746	-0.792]

Рисунок 7 – Сопряженные векторы

В качестве произвольного вектора q взят образ «О» другого написания.

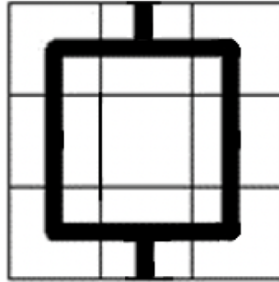


Рисунок 8 – Тестовый символ

$$q = [11, 11, 11, 01, 00, 01, 11, 11, 11]$$

$$q = [3, 3, 3, 1, 0, 1, 3, 3, 3]$$

Для текущего тестируемого вектора q были вычислены его параметры порядка (рисунок 9).

Параметры порядка:								
7.165317315225907								
-1.237050578915293								
1.3685035257247309								

Рисунок 9 – Параметры порядка текущего вектора

$$p_1 = 1.291; p_2 = -0.352; p_3 = 0.521$$

Для каждого описанного выше случая было решено дифференциальное уравнение:

$$\frac{dp_k(t)}{dt} = \lambda_k p_k - B p_k \sum_{i \neq k}^M p_i^2 - C p_k \sum_{i=1}^M p_i^2,$$

где $M = 3$, $B = 1 - C = -2$, $C = 3$, $\lambda_1 = \lambda_2 = \lambda_3 = 1$. Начальное состояние векторов параметров порядка задается следующим образом: $p_E = [7.1653, -1.237, 1.3685]$

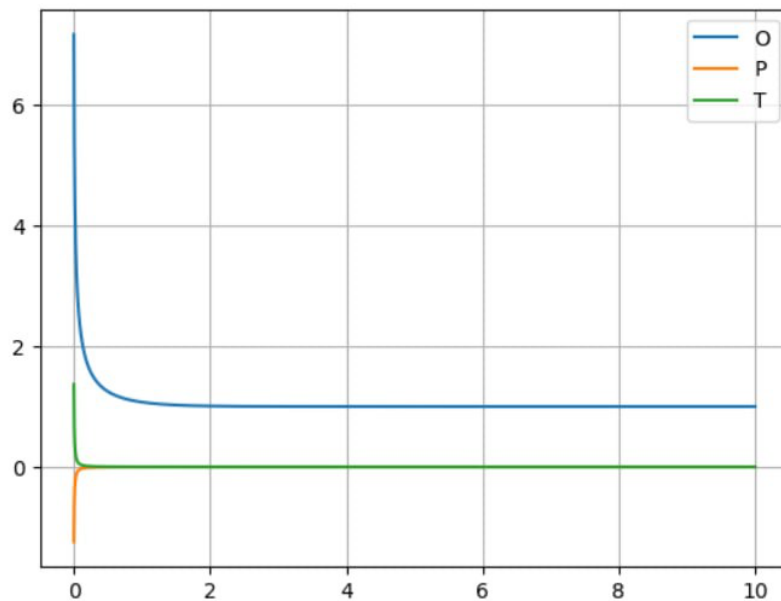


Рисунок 10 - Графическое решение системы дифференциальных уравнений

Первый параметр порядка имеет максимальное значение и после динамического изменения системы значение эталона v_1 стремится к 1, а остальные – к 0, следовательно, тестируемый вектор q больше всего похож на эталон $v_1 = "O"$.

Исходный код программы на python приведен в листинге 1.

Листинг 1 – Исходный код программы

```
import numpy as np
import math
import matplotlib.pyplot as plt
from scipy.integrate import odeint

def F(y,t):
    y0, y1, y2 = y

    dy0dt = y0 - y0**3 - 3 * y0 * (y1**2 + y2**2)
    dy1dt = y1 - y1**3 - 3 * y1 * (y0**2 + y2**2)
    dy2dt = y2 - y2**3 - 3 * y2 * (y0**2 + y1**2)

    return [dy0dt, dy1dt, dy2dt]

v1 = np.array([2, 3, 3, 1, 0, 1, 3, 2, 3])
v2 = np.array([3, 2, 3, 3, 2, 3, 1, 0, 0])
v3 = np.array([2, 3, 2, 0, 1, 0, 0, 1, 0])

print('Централизованные векторы:')
v1_centrelize = v1 - np.sum(v1)/v1.size
v2_centrelize = v2 - np.sum(v2)/v2.size
v3_centrelize = v3 - np.sum(v3)/v3.size
print(v1_centrelize)
```

```

print(v2_centrelize)
print(v3_centrelize)

print('Нормализованные векторы:')
v1_normolize = (v1_centrelize / np.sum(list(map(lambda x: abs(x) ,
v1_centrelize))))
v2_normolize = (v2_centrelize / np.sum(list(map(lambda x: abs(x) ,
v2_centrelize))))
v3_normolize = (v3_centrelize / np.sum(list(map(lambda x: abs(x) ,
v3_centrelize))))
print(v1_normolize.round(3))
print(v2_normolize.round(3))
print(v3_normolize.round(3))

print('Проверка векторов на удовлетворение условиям')
print('Для 1го вектора:')
print(np.sum(v1_normolize).round(3))
print(np.sum(list(map(lambda x: abs(x) , v1_normolize))).round(3))
print('Для 2го вектора:')
print(np.sum(v2_normolize).round(3))
print(np.sum(list(map(lambda x: abs(x) , v2_normolize))).round(3))
print('Для 3го вектора:')
print(np.sum(v3_normolize).round(3))
print(np.sum(list(map(lambda x: abs(x) , v3_normolize))).round(3))

print('Эталонная матрица:')
etalon_matrix = np.array([v1_normolize, v2_normolize, v3_normolize]).T
print(etalon_matrix.round(3))

print('Произведение эталонной матрицы на ее транспонированную (W)')
W = etalon_matrix.T@etalon_matrix
print(W.round(3))

print('Обратная матрица (A)')
A = np.linalg.inv(W)
print(A.round(3))

print('Присоединенные векторы:')
v_adj = np.empty([len(etalon_matrix),len(etalon_matrix[0])])
for i in range(0,len(etalon_matrix)) :
    v_adj[i] = A@etalon_matrix[i]

print(v_adj[:,0].round(3))
print(v_adj[:,1].round(3))
print(v_adj[:,2].round(3))

print('Параметры порядка:')
q = np.array([3, 3, 3, 1, 0, 1, 3, 3, 3])
p1 = np.array(v_adj[:,0]).dot(q)
p2 = np.array(v_adj[:,1]).dot(q)
p3 = np.array(v_adj[:,2]).dot(q)

print(p1)
print(p2)
print(p3)

t = np.linspace(0, 10, 10000)
y0 = [p1, p2, p3]

```

```
s = odeint(F, y0, t)
plt.plot(t, s)
plt.legend(['O', 'P', 'T'])
plt.grid()
plt.show()
```

Вывод: в результате данной лабораторной работы был разработан и исследован алгоритм принятия решения в условиях неопределенности, а именно, была формализована задача принятия решения по распознаванию текущего образа в системе мониторинга, разработан алгоритм распознавания и классификации образа.

Контрольные вопросы

1. Что понимают под процессом принятия решений в системах мониторинга?

Под процессом принятия решения в системах мониторинга понимают процесс изменения или выполнения действий необходимых, например, для возвращения системы в нормальные условия функционирования. Т.е. процесс принятия мер на изменение состояния объекта мониторинга.

2. Почему решения в системах мониторинга принимают в условиях неопределенности? Поясните ответ.

Оценка состояния сложной системы, особенно при дистанционном мониторинге, зачастую выполняется в условиях неопределенности. Это связано как с наличием неполной информации об объекте, так и с искажениями при передаче собранных данных. Кроме того, информационные потоки могут быть искажены за счет несанкционированных воздействий на информацию, в том числе и целенаправленного характера, и на основе искаженных данных могут быть приняты неверные решения.

3. Что понимают под неопределенностью информации?

Неопределенность информации означает неопределенность или непредсказуемость результата сообщения. Чем больше неопределенность, тем меньше информационная ценность. Для измерения неопределенности используются различные показатели, которые позволяют оценить, насколько конкретное сообщение является неопределенным.

4. Поясните технологию Data Mining, ее цели и задачи?

Data Mining – это процесс, цель которого – обнаружить новые значимые корреляции, образцы и тенденции в результате просеивания большого объема хранимых данных с использованием методик распознавания образцов, а также, применение статистических и математических методов. К основным задачам Data Mining относятся:

- 1) классификация;
- 2) регрессия;

- 3) кластеризация;
- 4) поиск ассоциативных правил;
- 5) поиск последовательности;
- 6) прогнозирование.

5. Укажите основные этапы интеллектуального анализа данных.

В интеллектуальном анализе данных выделяются три этапа:

- 1) предварительное исследование;
- 2) построение модели и ее проверка;
- 3) широкое применение.

6. Поясните задачу классификации объектов. Чем она отличается от задачи кластеризации?

Задача классификации объектов - задача разделения множества объектов на классы. Классификация позволяет организовывать наборы данных всех видов.

Кластеризация – это разбиение множества объектов на подмножества (кластеры) по заданному критерию.

Главное отличие между кластеризацией и классификацией состоит в том, что при классификации каждый объект относят к одному из заранее известных классов, при кластеризации, объекты распределяются по классам в процессе, то есть классы заранее не predetermined, а определяются результатом.

7. С какой целью организуют защиту данных в системах мониторинга?

В рамках мероприятий по мониторингу защита получаемых, хранимых, обрабатываемых и передаваемых данных организуется для мониторинга от несанкционированного воздействия. Это необходимо для получения достоверных данных об объекте мониторинга и для правильного принятия решения.