|  |  |
| --- | --- |
|  | **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**

**Отчет**

|  |  |
| --- | --- |
| **по лабораторной работе №** | 3 |

**Название:** Интеллектуальный анализ данных об объекте мониторинга

**Дисциплина:** Дистанционный мониторинг сложных систем и процессов

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Студент | ИУ6-12М |  |  | С.В. Астахов | |
|  | (Группа) |  | (Подпись, дата) | | (И.О. Фамилия) |
|  |  |  |  | |  |
| Преподаватель |  |  |  | | Ю.А. Вишневская |
|  |  |  | (Подпись, дата) | | (И.О. Фамилия) |

Москва, 2023

**Введение**

**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.

**Ход выполнения**

**Исходные данные:** набор эталонов и текущий (распознаваемый) вектор. В качестве эталонов выбрать три образа, например, группу из трех символов (рисунок 1).

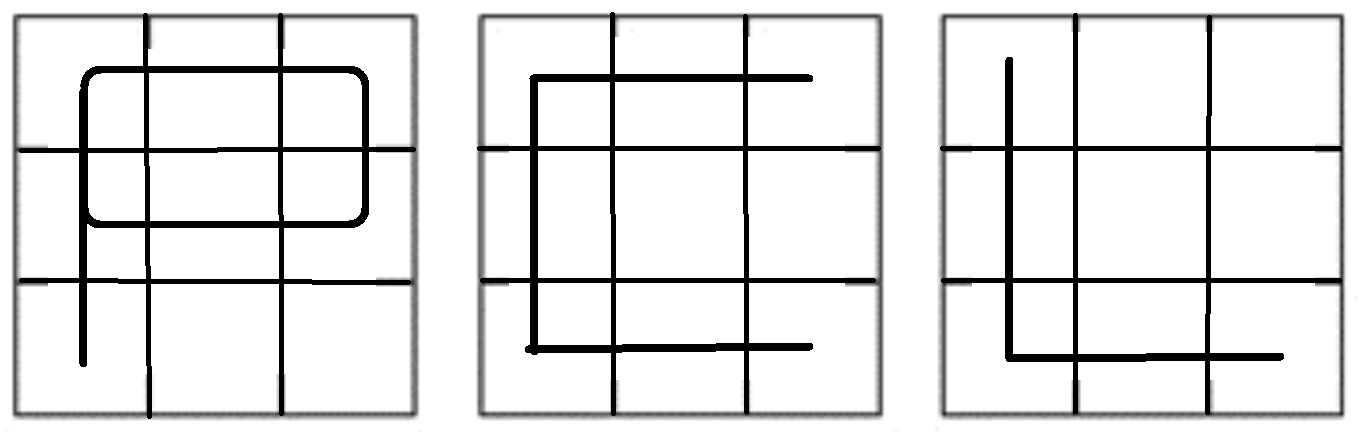


Рисунок 1 — эталонные символы

Здесь эталонные символы вписаны в матрицу 3×3. Таким образом, вектор признаков имеет длину, равную 9. В качестве тестового символа был выбран символ, представленный на рисунке 2.

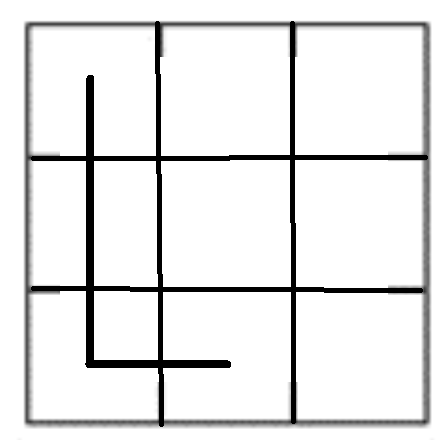


Рисунок 2 — тестовый символ

Для удобства изложения, основной ход выполнения работы описан в комментариях к листингам. Консольный вывод программы приведен в комментариях с началом “(stdout)”. Решение выполнено на языке python. В лиистинге 1 приведен исходный код перевода векторов признаков в десятичную систему

Листинг 1 — перевод векторов признаков в десятичную систему

|  |
| --- |
| import numpy as np  from scipy.integrate import odeint  import matplotlib.pyplot as plt  # вектора признаков  # 01 - вертикальная линия  # 10 - горизонтальная линия  # 11 - пересечение  P = [      11, 10, 11,      11, 10, 11,       1,  0,  0  ]  C = [      11, 10, 10,       1,  0,  0,      11, 10, 10  ]  L = [       1,  0,  0,       1,  0,  0,      11, 10,  10  ]  # перевод кодов признаков в десятичную систему  v1 = list(map(lambda x: int(str(x), 2), P))  v2 = list(map(lambda x: int(str(x), 2), C))  v3 = list(map(lambda x: int(str(x), 2), L))  print("вектора в десятичной системе:\n\n" + str(v1) + "\n" + str(v2) + "\n" + str(v3))  # (stdout) вектора в десятичной системе:  #  # (stdout) [3, 2, 3, 3, 2, 3, 1, 0, 0]  # (stdout) [3, 2, 2, 1, 0, 0, 3, 2, 2]  # (stdout) [1, 0, 0, 1, 0, 0, 3, 2, 2] |

В листинге 2 демонстрируется процесс централизации и нормализации векторов

Листинг 2 — централизация и нормализация

|  |
| --- |
| # централизация векторов  avg1 = sum(v1) / len(v1)  avg2 = sum(v2) / len(v2)  avg3 = sum(v3) / len(v3)  v1 = list(map(lambda x: x - avg1, v1))  v2 = list(map(lambda x: x - avg2, v2))  v3 = list(map(lambda x: x - avg3, v3))  # нормализация векторов  ln1 = sum(list(map(lambda x: x\*x, v1)))  v1 = list(map(lambda x: x / (ln1 \*\* 0.5), v1))  ln2 = sum(list(map(lambda x: x\*x, v2)))  v2 = list(map(lambda x: x / (ln2 \*\* 0.5), v2))  ln3 = sum(list(map(lambda x: x\*x, v3)))  v3 = list(map(lambda x: x / (ln3 \*\* 0.5), v3))  # проверка  ln1 = sum(list(map(lambda x: x\*x, v1)))  ln2 = sum(list(map(lambda x: x\*x, v2)))  ln3 = sum(list(map(lambda x: x\*x, v3)))  print(f"Длины векторов: {ln1}, {ln2}, {ln3}")  print(f"Сумма компонентов векторов: {sum(v1)}, {sum(v2)}, {sum(v3)}")  print("\nНормализованные вектора:\n\n"+str(v1)+"\n\n"+str(v2)+"\n\n"+str(v3))  ””” (stdout)  Длины векторов: 1.0, 0.9999999999999998, 1.0  Сумма компонентов векторов: 0.0, -2.498001805406602e-16, 0.0  Нормализованные вектора:  [0.30949223029508643, 0.030949223029508657, 0.30949223029508643, 0.30949223029508643, 0.030949223029508657, 0.30949223029508643, -0.24759378423606915, -0.5261367915016469, -0.5261367915016469]  [0.42163702135578385, 0.10540925533894595, 0.10540925533894595, -0.21081851067789198, -0.5270462766947299, -0.5270462766947299, 0.42163702135578385, 0.10540925533894595, 0.10540925533894595]  [0.0, -0.31622776601683794, -0.31622776601683794, 0.0, -0.31622776601683794, -0.31622776601683794, 0.6324555320336759, 0.31622776601683794, 0.31622776601683794]  ””” |

**Вывод:** в ходе лабораторной работы были изучены методы выявления закономерностей в разнородных данных и реализован один из них.

**Контрольные вопросы**

***1. С какой целью проводится кодирование информационных признаков?***

Кодирование информации происходит для уменьшение излишка информации, для удобства работы с ним в последующем, что позволяет обрабатывать информацию быстрее и точнее (при правильном кодировании).

***2. Как можно определить логические закономерности в данных?***

* последовательность;
* ассоциация;
* классификация;
* прогнозирование.

***3. Укажите методы выявления логических закономерностей.***

* стохастический локальный поиск – поиск в случайной выборке;
* алгоритм КОРА (взвешенное голосование правил) – строит набор конъюнктивных закономерностей;
* алгоритм ТЕМП (поиск в ширину);
* генетический алгоритм (поиск правил).

***4. С какой целью проводится интеграция и структурирование данных при мониторинге?***

Основная задача — упрощение понимания основных элементов, из которых состоит весь массив информации, а также логики взаимосвязанности этих элементов.

***5. Укажите интеллектуальные методы, применяемые для анализа Big Data.***

* machine learning;
* data mining;
* краудсорсинг;
* нейросети;
* предиктивный и статистический анализ;
* имитационные модели.