Лабораторная работа 2

**Автоматизированное приобретение знаний**

**из баз данных**

**Цель работы:** освоить метод индуцирования знаний из баз данных и получить навыки разработки алгоритма приобретения знаний по технологии Data Mining.

**Предметная область**: 19. Поиск неисправностей легкового автомобиля.

**Задание к лабораторной работе**

Получить дополнительные знания для проектируемой экспертной системы из таблицы, содержащей сведения о зависимости некоторого целевого параметра от характеристик заданной предметной области. Источником знаний считать реляционную базу данных (таблицу), в которой собраны данные о связях между различными параметрами (значения параметров задать самостоятельно, исходя из реальных ситуаций). Сведения в таблице должны характеризовать некоторую подцель, т.е. поддерево, соответствующее одному из вопросов главного дерева решений задачи из лабораторной работы №1. Пути достижения выбранной подцели должны быть выявлены и представлены новым поддеревом решений и продукционными правилами, полученными методом индуцирования знаний из таблицы.

**Исходная таблица данных**

Таблица результатов наблюдений опытными водителями за поведением автомобиля для определения «Есть ли признаки неисправности автомобиля».

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Запах** | **Движение** | **Заводится** | **Свет** | **Тормоза** | **Неисправность** |
| 0 | Сильный | Ровное | Хорошо | Работает | Работает | Да |
| 1 | Сильный | Ровное | Хорошо | Работает | Неполадки | Да |
| 2 | Сильный | Ровное | Хорошо | Неполадки | Работает | Да |
| 3 | Сильный | Ровное | Хорошо | Неполадки | Неполадки | Да |
| 4 | Сильный | Ровное | Плохо | Работает | Работает | Да |
| 5 | Сильный | Ровное | Плохо | Работает | Неполадки | Да |
| 6 | Сильный | Ровное | Плохо | Неполадки | Работает | Да |
| 7 | Сильный | Ровное | Плохо | Неполадки | Неполадки | Да |
| 8 | Сильный | Ровное | Нет | Работает | Работает | Да |
| 9 | Сильный | Ровное | Нет | Работает | Неполадки | Да |
| 10 | Сильный | Ровное | Нет | Неполадки | Работает | Да |
| 11 | Сильный | Ровное | Нет | Неполадки | Неполадки | Да |
| 12 | Сильный | Ведет | Хорошо | Работает | Работает | Да |
| 13 | Сильный | Ведет | Хорошо | Работает | Неполадки | Да |
| 14 | Сильный | Ведет | Хорошо | Неполадки | Работает | Да |
| 15 | Сильный | Ведет | Хорошо | Неполадки | Неполадки | Да |
| 16 | Сильный | Ведет | Плохо | Работает | Работает | Да |
| 17 | Сильный | Ведет | Плохо | Работает | Неполадки | Да |
| 18 | Сильный | Ведет | Плохо | Неполадки | Работает | Да |
| 19 | Сильный | Ведет | Плохо | Неполадки | Неполадки | Да |
| 20 | Сильный | Ведет | Нет | Работает | Работает | Да |
| 21 | Сильный | Ведет | Нет | Работает | Неполадки | Да |
| 22 | Сильный | Ведет | Нет | Неполадки | Работает | Да |
| 23 | Сильный | Ведет | Нет | Неполадки | Неполадки | Да |
| 24 | Сильный | Нет | Хорошо | Работает | Работает | Да |
| 25 | Сильный | Нет | Хорошо | Работает | Неполадки | Да |
| 26 | Сильный | Нет | Хорошо | Неполадки | Работает | Да |
| 27 | Сильный | Нет | Хорошо | Неполадки | Неполадки | Да |
| 28 | Сильный | Нет | Плохо | Работает | Работает | Да |
| 29 | Сильный | Нет | Плохо | Работает | Неполадки | Да |
| 30 | Сильный | Нет | Плохо | Неполадки | Работает | Да |
| 31 | Сильный | Нет | Плохо | Неполадки | Неполадки | Да |
| 32 | Сильный | Нет | Нет | Работает | Работает | Да |
| 33 | Сильный | Нет | Нет | Работает | Неполадки | Да |
| 34 | Сильный | Нет | Нет | Неполадки | Работает | Да |
| 35 | Сильный | Нет | Нет | Неполадки | Неполадки | Да |
| 36 | Слабый | Ровное | Хорошо | Работает | Работает | Нет |
| 37 | Слабый | Ровное | Хорошо | Работает | Неполадки | Да |
| 38 | Слабый | Ровное | Хорошо | Неполадки | Работает | Да |
| 39 | Слабый | Ровное | Хорошо | Неполадки | Неполадки | Да |
| 40 | Слабый | Ровное | Плохо | Работает | Работает | Нет |
| 41 | Слабый | Ровное | Плохо | Работает | Неполадки | Да |
| 42 | Слабый | Ровное | Плохо | Неполадки | Работает | Да |
| 43 | Слабый | Ровное | Плохо | Неполадки | Неполадки | Да |
| 44 | Слабый | Ровное | Нет | Работает | Работает | Да |
| 45 | Слабый | Ровное | Нет | Работает | Неполадки | Да |
| 46 | Слабый | Ровное | Нет | Неполадки | Работает | Да |
| 47 | Слабый | Ровное | Нет | Неполадки | Неполадки | Да |
| 48 | Слабый | Ведет | Хорошо | Работает | Работает | Нет |
| 49 | Слабый | Ведет | Хорошо | Работает | Неполадки | Да |
| 50 | Слабый | Ведет | Хорошо | Неполадки | Работает | Да |
| 51 | Слабый | Ведет | Хорошо | Неполадки | Неполадки | Да |
| 52 | Слабый | Ведет | Плохо | Работает | Работает | Нет |
| 53 | Слабый | Ведет | Плохо | Работает | Неполадки | Да |
| 54 | Слабый | Ведет | Плохо | Неполадки | Работает | Да |
| 55 | Слабый | Ведет | Плохо | Неполадки | Неполадки | Да |
| 56 | Слабый | Ведет | Нет | Работает | Работает | Да |
| 57 | Слабый | Ведет | Нет | Работает | Неполадки | Да |
| 58 | Слабый | Ведет | Нет | Неполадки | Работает | Да |
| 59 | Слабый | Ведет | Нет | Неполадки | Неполадки | Да |
| 60 | Слабый | Нет | Хорошо | Работает | Работает | Да |
| 61 | Слабый | Нет | Хорошо | Работает | Неполадки | Да |
| 62 | Слабый | Нет | Хорошо | Неполадки | Работает | Да |
| 63 | Слабый | Нет | Хорошо | Неполадки | Неполадки | Да |
| 64 | Слабый | Нет | Плохо | Работает | Работает | Да |
| 65 | Слабый | Нет | Плохо | Работает | Неполадки | Да |
| 66 | Слабый | Нет | Плохо | Неполадки | Работает | Да |
| 67 | Слабый | Нет | Плохо | Неполадки | Неполадки | Да |
| 68 | Слабый | Нет | Нет | Работает | Работает | Да |
| 69 | Слабый | Нет | Нет | Работает | Неполадки | Да |
| 70 | Слабый | Нет | Нет | Неполадки | Работает | Да |
| 71 | Слабый | Нет | Нет | Неполадки | Неполадки | Да |
| 72 | Нет | Ровное | Хорошо | Работает | Работает | Нет |
| 73 | Нет | Ровное | Хорошо | Работает | Неполадки | Да |
| 74 | Нет | Ровное | Хорошо | Неполадки | Работает | Да |
| 75 | Нет | Ровное | Хорошо | Неполадки | Неполадки | Да |
| 76 | Нет | Ровное | Плохо | Работает | Работает | Нет |
| 77 | Нет | Ровное | Плохо | Работает | Неполадки | Да |
| 78 | Нет | Ровное | Плохо | Неполадки | Работает | Да |
| 79 | Нет | Ровное | Плохо | Неполадки | Неполадки | Да |
| 80 | Нет | Ровное | Нет | Работает | Работает | Да |
| 81 | Нет | Ровное | Нет | Работает | Неполадки | Да |
| 82 | Нет | Ровное | Нет | Неполадки | Работает | Да |
| 83 | Нет | Ровное | Нет | Неполадки | Неполадки | Да |
| 84 | Нет | Ведет | Хорошо | Работает | Работает | Нет |
| 85 | Нет | Ведет | Хорошо | Работает | Неполадки | Да |
| 86 | Нет | Ведет | Хорошо | Неполадки | Работает | Да |
| 87 | Нет | Ведет | Хорошо | Неполадки | Неполадки | Да |
| 88 | Нет | Ведет | Плохо | Работает | Работает | Нет |
| 89 | Нет | Ведет | Плохо | Работает | Неполадки | Да |
| 90 | Нет | Ведет | Плохо | Неполадки | Работает | Да |
| 91 | Нет | Ведет | Плохо | Неполадки | Неполадки | Да |
| 92 | Нет | Ведет | Нет | Работает | Работает | Да |
| 93 | Нет | Ведет | Нет | Работает | Неполадки | Да |
| 94 | Нет | Ведет | Нет | Неполадки | Работает | Да |
| 95 | Нет | Ведет | Нет | Неполадки | Неполадки | Да |
| 96 | Нет | Нет | Хорошо | Работает | Работает | Да |
| 97 | Нет | Нет | Хорошо | Работает | Неполадки | Да |
| 98 | Нет | Нет | Хорошо | Неполадки | Работает | Да |
| 99 | Нет | Нет | Хорошо | Неполадки | Неполадки | Да |
| 100 | Нет | Нет | Плохо | Работает | Работает | Да |
| 101 | Нет | Нет | Плохо | Работает | Неполадки | Да |
| 102 | Нет | Нет | Плохо | Неполадки | Работает | Да |
| 103 | Нет | Нет | Плохо | Неполадки | Неполадки | Да |
| 104 | Нет | Нет | Нет | Работает | Работает | Да |
| 105 | Нет | Нет | Нет | Работает | Неполадки | Да |
| 106 | Нет | Нет | Нет | Неполадки | Работает | Да |
| 107 | Нет | Нет | Нет | Неполадки | Неполадки | Да |

**Листинг работы программы**

Шаг алгоритма №0

Запах Движение Заводится Свет Тормоза Неисправность

0 Сильный Ровное Хорошо Работает Работает Да

1 Сильный Ровное Хорошо Работает Неполадки Да

2 Сильный Ровное Хорошо Неполадки Работает Да

3 Сильный Ровное Хорошо Неполадки Неполадки Да

4 Сильный Ровное Плохо Работает Работает Да

5 Сильный Ровное Плохо Работает Неполадки Да

6 Сильный Ровное Плохо Неполадки Работает Да

7 Сильный Ровное Плохо Неполадки Неполадки Да

8 Сильный Ровное Нет Работает Работает Да

9 Сильный Ровное Нет Работает Неполадки Да

10 Сильный Ровное Нет Неполадки Работает Да

11 Сильный Ровное Нет Неполадки Неполадки Да

12 Сильный Ведет Хорошо Работает Работает Да

13 Сильный Ведет Хорошо Работает Неполадки Да

14 Сильный Ведет Хорошо Неполадки Работает Да

15 Сильный Ведет Хорошо Неполадки Неполадки Да

16 Сильный Ведет Плохо Работает Работает Да

17 Сильный Ведет Плохо Работает Неполадки Да

18 Сильный Ведет Плохо Неполадки Работает Да

19 Сильный Ведет Плохо Неполадки Неполадки Да

20 Сильный Ведет Нет Работает Работает Да

21 Сильный Ведет Нет Работает Неполадки Да

22 Сильный Ведет Нет Неполадки Работает Да

23 Сильный Ведет Нет Неполадки Неполадки Да

24 Сильный Нет Хорошо Работает Работает Да

25 Сильный Нет Хорошо Работает Неполадки Да

26 Сильный Нет Хорошо Неполадки Работает Да

27 Сильный Нет Хорошо Неполадки Неполадки Да

28 Сильный Нет Плохо Работает Работает Да

29 Сильный Нет Плохо Работает Неполадки Да

30 Сильный Нет Плохо Неполадки Работает Да

31 Сильный Нет Плохо Неполадки Неполадки Да

32 Сильный Нет Нет Работает Работает Да

33 Сильный Нет Нет Работает Неполадки Да

34 Сильный Нет Нет Неполадки Работает Да

35 Сильный Нет Нет Неполадки Неполадки Да

36 Слабый Ровное Хорошо Работает Работает Нет

37 Слабый Ровное Хорошо Работает Неполадки Да

38 Слабый Ровное Хорошо Неполадки Работает Да

39 Слабый Ровное Хорошо Неполадки Неполадки Да

40 Слабый Ровное Плохо Работает Работает Нет

41 Слабый Ровное Плохо Работает Неполадки Да

42 Слабый Ровное Плохо Неполадки Работает Да

43 Слабый Ровное Плохо Неполадки Неполадки Да

44 Слабый Ровное Нет Работает Работает Да

45 Слабый Ровное Нет Работает Неполадки Да

46 Слабый Ровное Нет Неполадки Работает Да

47 Слабый Ровное Нет Неполадки Неполадки Да

48 Слабый Ведет Хорошо Работает Работает Нет

49 Слабый Ведет Хорошо Работает Неполадки Да

50 Слабый Ведет Хорошо Неполадки Работает Да

51 Слабый Ведет Хорошо Неполадки Неполадки Да

52 Слабый Ведет Плохо Работает Работает Нет

53 Слабый Ведет Плохо Работает Неполадки Да

54 Слабый Ведет Плохо Неполадки Работает Да

55 Слабый Ведет Плохо Неполадки Неполадки Да

56 Слабый Ведет Нет Работает Работает Да

57 Слабый Ведет Нет Работает Неполадки Да

58 Слабый Ведет Нет Неполадки Работает Да

59 Слабый Ведет Нет Неполадки Неполадки Да

60 Слабый Нет Хорошо Работает Работает Да

61 Слабый Нет Хорошо Работает Неполадки Да

62 Слабый Нет Хорошо Неполадки Работает Да

63 Слабый Нет Хорошо Неполадки Неполадки Да

64 Слабый Нет Плохо Работает Работает Да

65 Слабый Нет Плохо Работает Неполадки Да

66 Слабый Нет Плохо Неполадки Работает Да

67 Слабый Нет Плохо Неполадки Неполадки Да

68 Слабый Нет Нет Работает Работает Да

69 Слабый Нет Нет Работает Неполадки Да

70 Слабый Нет Нет Неполадки Работает Да

71 Слабый Нет Нет Неполадки Неполадки Да

72 Нет Ровное Хорошо Работает Работает Нет

73 Нет Ровное Хорошо Работает Неполадки Да

74 Нет Ровное Хорошо Неполадки Работает Да

75 Нет Ровное Хорошо Неполадки Неполадки Да

76 Нет Ровное Плохо Работает Работает Нет

77 Нет Ровное Плохо Работает Неполадки Да

78 Нет Ровное Плохо Неполадки Работает Да

79 Нет Ровное Плохо Неполадки Неполадки Да

80 Нет Ровное Нет Работает Работает Да

81 Нет Ровное Нет Работает Неполадки Да

82 Нет Ровное Нет Неполадки Работает Да

83 Нет Ровное Нет Неполадки Неполадки Да

84 Нет Ведет Хорошо Работает Работает Нет

85 Нет Ведет Хорошо Работает Неполадки Да

86 Нет Ведет Хорошо Неполадки Работает Да

87 Нет Ведет Хорошо Неполадки Неполадки Да

88 Нет Ведет Плохо Работает Работает Нет

89 Нет Ведет Плохо Работает Неполадки Да

90 Нет Ведет Плохо Неполадки Работает Да

91 Нет Ведет Плохо Неполадки Неполадки Да

92 Нет Ведет Нет Работает Работает Да

93 Нет Ведет Нет Работает Неполадки Да

94 Нет Ведет Нет Неполадки Работает Да

95 Нет Ведет Нет Неполадки Неполадки Да

96 Нет Нет Хорошо Работает Работает Да

97 Нет Нет Хорошо Работает Неполадки Да

98 Нет Нет Хорошо Неполадки Работает Да

99 Нет Нет Хорошо Неполадки Неполадки Да

100 Нет Нет Плохо Работает Работает Да

101 Нет Нет Плохо Работает Неполадки Да

102 Нет Нет Плохо Неполадки Работает Да

103 Нет Нет Плохо Неполадки Неполадки Да

104 Нет Нет Нет Работает Работает Да

105 Нет Нет Нет Работает Неполадки Да

106 Нет Нет Нет Неполадки Работает Да

107 Нет Нет Нет Неполадки Неполадки Да

Gain(Запах) = 0.0454410291882929

Gain(Движение) = 0.0454410291882929

Gain(Заводится) = 0.0454410291882929

Gain(Свет) = 0.07835329738867974

Gain(Тормоза) = 0.07835329738867974

Наибольшее Gain: Gain(Тормоза) = 0.07835329738867974

Шаг алгоритма №1

Запах Движение Заводится Свет Тормоза Неисправность

0 Сильный Ровное Хорошо Работает Работает Да

2 Сильный Ровное Хорошо Неполадки Работает Да

4 Сильный Ровное Плохо Работает Работает Да

6 Сильный Ровное Плохо Неполадки Работает Да

8 Сильный Ровное Нет Работает Работает Да

10 Сильный Ровное Нет Неполадки Работает Да

12 Сильный Ведет Хорошо Работает Работает Да

14 Сильный Ведет Хорошо Неполадки Работает Да

16 Сильный Ведет Плохо Работает Работает Да

18 Сильный Ведет Плохо Неполадки Работает Да

20 Сильный Ведет Нет Работает Работает Да

22 Сильный Ведет Нет Неполадки Работает Да

24 Сильный Нет Хорошо Работает Работает Да

26 Сильный Нет Хорошо Неполадки Работает Да

28 Сильный Нет Плохо Работает Работает Да

30 Сильный Нет Плохо Неполадки Работает Да

32 Сильный Нет Нет Работает Работает Да

34 Сильный Нет Нет Неполадки Работает Да

36 Слабый Ровное Хорошо Работает Работает Нет

38 Слабый Ровное Хорошо Неполадки Работает Да

40 Слабый Ровное Плохо Работает Работает Нет

42 Слабый Ровное Плохо Неполадки Работает Да

44 Слабый Ровное Нет Работает Работает Да

46 Слабый Ровное Нет Неполадки Работает Да

48 Слабый Ведет Хорошо Работает Работает Нет

50 Слабый Ведет Хорошо Неполадки Работает Да

52 Слабый Ведет Плохо Работает Работает Нет

54 Слабый Ведет Плохо Неполадки Работает Да

56 Слабый Ведет Нет Работает Работает Да

58 Слабый Ведет Нет Неполадки Работает Да

60 Слабый Нет Хорошо Работает Работает Да

62 Слабый Нет Хорошо Неполадки Работает Да

64 Слабый Нет Плохо Работает Работает Да

66 Слабый Нет Плохо Неполадки Работает Да

68 Слабый Нет Нет Работает Работает Да

70 Слабый Нет Нет Неполадки Работает Да

72 Нет Ровное Хорошо Работает Работает Нет

74 Нет Ровное Хорошо Неполадки Работает Да

76 Нет Ровное Плохо Работает Работает Нет

78 Нет Ровное Плохо Неполадки Работает Да

80 Нет Ровное Нет Работает Работает Да

82 Нет Ровное Нет Неполадки Работает Да

84 Нет Ведет Хорошо Работает Работает Нет

86 Нет Ведет Хорошо Неполадки Работает Да

88 Нет Ведет Плохо Работает Работает Нет

90 Нет Ведет Плохо Неполадки Работает Да

92 Нет Ведет Нет Работает Работает Да

94 Нет Ведет Нет Неполадки Работает Да

96 Нет Нет Хорошо Работает Работает Да

98 Нет Нет Хорошо Неполадки Работает Да

100 Нет Нет Плохо Работает Работает Да

102 Нет Нет Плохо Неполадки Работает Да

104 Нет Нет Нет Работает Работает Да

106 Нет Нет Нет Неполадки Работает Да

Gain(Запах) = 0.09571690562767377

Gain(Движение) = 0.09571690562767377

Gain(Заводится) = 0.09571690562767377

Gain(Свет) = 0.1668284321336886

Наибольшее Gain: Gain(Свет) = 0.1668284321336886

Шаг алгоритма №2

Подтаблица, в которой Тормоза = Работает

Запах Движение Заводится Свет Неисправность

0 Сильный Ровное Хорошо Работает Да

4 Сильный Ровное Плохо Работает Да

8 Сильный Ровное Нет Работает Да

12 Сильный Ведет Хорошо Работает Да

16 Сильный Ведет Плохо Работает Да

20 Сильный Ведет Нет Работает Да

24 Сильный Нет Хорошо Работает Да

28 Сильный Нет Плохо Работает Да

32 Сильный Нет Нет Работает Да

36 Слабый Ровное Хорошо Работает Нет

40 Слабый Ровное Плохо Работает Нет

44 Слабый Ровное Нет Работает Да

48 Слабый Ведет Хорошо Работает Нет

52 Слабый Ведет Плохо Работает Нет

56 Слабый Ведет Нет Работает Да

60 Слабый Нет Хорошо Работает Да

64 Слабый Нет Плохо Работает Да

68 Слабый Нет Нет Работает Да

72 Нет Ровное Хорошо Работает Нет

76 Нет Ровное Плохо Работает Нет

80 Нет Ровное Нет Работает Да

84 Нет Ведет Хорошо Работает Нет

88 Нет Ведет Плохо Работает Нет

92 Нет Ведет Нет Работает Да

96 Нет Нет Хорошо Работает Да

100 Нет Нет Плохо Работает Да

104 Нет Нет Нет Работает Да

Gain(Запах) = 0.21599891577398245

Gain(Движение) = 0.21599891577398245

Gain(Заводится) = 0.21599891577398245

Наибольшее Gain: Gain(Заводится) = 0.21599891577398245

Шаг алгоритма №3

Подтаблица, в которой Тормоза = Работает, Свет = Работает

Запах Движение Заводится Неисправность

0 Сильный Ровное Хорошо Да

12 Сильный Ведет Хорошо Да

24 Сильный Нет Хорошо Да

36 Слабый Ровное Хорошо Нет

48 Слабый Ведет Хорошо Нет

60 Слабый Нет Хорошо Да

72 Нет Ровное Хорошо Нет

84 Нет Ведет Хорошо Нет

96 Нет Нет Хорошо Да

Gain(Запах) = 0.3788788371352292

Gain(Движение) = 0.3788788371352292

Наибольшее Gain: Gain(Движение) = 0.3788788371352292

Шаг алгоритма №4

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Хорошо

Запах Движение Неисправность

0 Сильный Ровное Да

36 Слабый Ровное Нет

72 Нет Ровное Нет

Gain(Запах) = 0.9182958340544896

Наибольшее Gain: Gain(Запах) = 0.9182958340544896

Шаг алгоритма №5

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ровное

Запах Неисправность

0 Сильный Да

Найдена монотонная таблица:

Выведено новое правило:

IF Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ровное, Запах = Сильный THEN Неисправность = Да

Шаг алгоритма №5

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ровное

Запах Неисправность

36 Слабый Нет

Найдена монотонная таблица:

Выведено новое правило:

IF Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ровное, Запах = Слабый THEN Неисправность = Нет

Шаг алгоритма №5

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ровное

Запах Неисправность

72 Нет Нет

Найдена монотонная таблица:

Выведено новое правило:

IF Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ровное, Запах = Нет THEN Неисправность = Нет

Шаг алгоритма №4

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Хорошо

Запах Движение Неисправность

12 Сильный Ведет Да

48 Слабый Ведет Нет

84 Нет Ведет Нет

Gain(Запах) = 0.9182958340544896

Наибольшее Gain: Gain(Запах) = 0.9182958340544896

Шаг алгоритма №5

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ведет

Запах Неисправность

12 Сильный Да

Найдена монотонная таблица:

Выведено новое правило:

IF Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ведет, Запах = Сильный THEN Неисправность = Да

Шаг алгоритма №5

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ведет

Запах Неисправность

48 Слабый Нет

Найдена монотонная таблица:

Выведено новое правило:

IF Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ведет, Запах = Слабый THEN Неисправность = Нет

Шаг алгоритма №5

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ведет

Запах Неисправность

84 Нет Нет

Найдена монотонная таблица:

Выведено новое правило:

IF Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ведет, Запах = Нет THEN Неисправность = Нет

Шаг алгоритма №4

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Хорошо

Запах Движение Неисправность

24 Сильный Нет Да

60 Слабый Нет Да

96 Нет Нет Да

Найдена монотонная таблица:

Выведено новое правило:

IF Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Нет THEN Неисправность = Да

Шаг алгоритма №3

Подтаблица, в которой Тормоза = Работает, Свет = Работает

Запах Движение Заводится Неисправность

4 Сильный Ровное Плохо Да

16 Сильный Ведет Плохо Да

28 Сильный Нет Плохо Да

40 Слабый Ровное Плохо Нет

52 Слабый Ведет Плохо Нет

64 Слабый Нет Плохо Да

76 Нет Ровное Плохо Нет

88 Нет Ведет Плохо Нет

100 Нет Нет Плохо Да

Gain(Запах) = 0.3788788371352292

Gain(Движение) = 0.3788788371352292

Наибольшее Gain: Gain(Движение) = 0.3788788371352292

Шаг алгоритма №4

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Плохо

Запах Движение Неисправность

4 Сильный Ровное Да

40 Слабый Ровное Нет

76 Нет Ровное Нет

Gain(Запах) = 0.9182958340544896

Наибольшее Gain: Gain(Запах) = 0.9182958340544896

Шаг алгоритма №5

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ровное

Запах Неисправность

4 Сильный Да

Найдена монотонная таблица:

Выведено новое правило:

IF Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ровное, Запах = Сильный THEN Неисправность = Да

Шаг алгоритма №5

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ровное

Запах Неисправность

40 Слабый Нет

Найдена монотонная таблица:

Выведено новое правило:

IF Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ровное, Запах = Слабый THEN Неисправность = Нет

Шаг алгоритма №5

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ровное

Запах Неисправность

76 Нет Нет

Найдена монотонная таблица:

Выведено новое правило:

IF Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ровное, Запах = Нет THEN Неисправность = Нет

Шаг алгоритма №4

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Плохо

Запах Движение Неисправность

16 Сильный Ведет Да

52 Слабый Ведет Нет

88 Нет Ведет Нет

Gain(Запах) = 0.9182958340544896

Наибольшее Gain: Gain(Запах) = 0.9182958340544896

Шаг алгоритма №5

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ведет

Запах Неисправность

16 Сильный Да

Найдена монотонная таблица:

Выведено новое правило:

IF Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ведет, Запах = Сильный THEN Неисправность = Да

Шаг алгоритма №5

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ведет

Запах Неисправность

52 Слабый Нет

Найдена монотонная таблица:

Выведено новое правило:

IF Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ведет, Запах = Слабый THEN Неисправность = Нет

Шаг алгоритма №5

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ведет

Запах Неисправность

88 Нет Нет

Найдена монотонная таблица:

Выведено новое правило:

IF Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ведет, Запах = Нет THEN Неисправность = Нет

Шаг алгоритма №4

Подтаблица, в которой Тормоза = Работает, Свет = Работает, Заводится = Плохо

Запах Движение Неисправность

28 Сильный Нет Да

64 Слабый Нет Да

100 Нет Нет Да

Найдена монотонная таблица:

Выведено новое правило:

IF Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Нет THEN Неисправность = Да

Шаг алгоритма №3

Подтаблица, в которой Тормоза = Работает, Свет = Работает

Запах Движение Заводится Неисправность

8 Сильный Ровное Нет Да

20 Сильный Ведет Нет Да

32 Сильный Нет Нет Да

44 Слабый Ровное Нет Да

56 Слабый Ведет Нет Да

68 Слабый Нет Нет Да

80 Нет Ровное Нет Да

92 Нет Ведет Нет Да

104 Нет Нет Нет Да

Найдена монотонная таблица:

Выведено новое правило:

IF Тормоза = Работает, Свет = Работает, Заводится = Нет THEN Неисправность = Да

Шаг алгоритма №2

Подтаблица, в которой Тормоза = Работает

Запах Движение Заводится Свет Неисправность

2 Сильный Ровное Хорошо Неполадки Да

6 Сильный Ровное Плохо Неполадки Да

10 Сильный Ровное Нет Неполадки Да

14 Сильный Ведет Хорошо Неполадки Да

18 Сильный Ведет Плохо Неполадки Да

22 Сильный Ведет Нет Неполадки Да

26 Сильный Нет Хорошо Неполадки Да

30 Сильный Нет Плохо Неполадки Да

34 Сильный Нет Нет Неполадки Да

38 Слабый Ровное Хорошо Неполадки Да

42 Слабый Ровное Плохо Неполадки Да

46 Слабый Ровное Нет Неполадки Да

50 Слабый Ведет Хорошо Неполадки Да

54 Слабый Ведет Плохо Неполадки Да

58 Слабый Ведет Нет Неполадки Да

62 Слабый Нет Хорошо Неполадки Да

66 Слабый Нет Плохо Неполадки Да

70 Слабый Нет Нет Неполадки Да

74 Нет Ровное Хорошо Неполадки Да

78 Нет Ровное Плохо Неполадки Да

82 Нет Ровное Нет Неполадки Да

86 Нет Ведет Хорошо Неполадки Да

90 Нет Ведет Плохо Неполадки Да

94 Нет Ведет Нет Неполадки Да

98 Нет Нет Хорошо Неполадки Да

102 Нет Нет Плохо Неполадки Да

106 Нет Нет Нет Неполадки Да

Найдена монотонная таблица:

Выведено новое правило:

IF Тормоза = Работает, Свет = Неполадки THEN Неисправность = Да

Шаг алгоритма №1

Запах Движение Заводится Свет Тормоза Неисправность

1 Сильный Ровное Хорошо Работает Неполадки Да

3 Сильный Ровное Хорошо Неполадки Неполадки Да

5 Сильный Ровное Плохо Работает Неполадки Да

7 Сильный Ровное Плохо Неполадки Неполадки Да

9 Сильный Ровное Нет Работает Неполадки Да

11 Сильный Ровное Нет Неполадки Неполадки Да

13 Сильный Ведет Хорошо Работает Неполадки Да

15 Сильный Ведет Хорошо Неполадки Неполадки Да

17 Сильный Ведет Плохо Работает Неполадки Да

19 Сильный Ведет Плохо Неполадки Неполадки Да

21 Сильный Ведет Нет Работает Неполадки Да

23 Сильный Ведет Нет Неполадки Неполадки Да

25 Сильный Нет Хорошо Работает Неполадки Да

27 Сильный Нет Хорошо Неполадки Неполадки Да

29 Сильный Нет Плохо Работает Неполадки Да

31 Сильный Нет Плохо Неполадки Неполадки Да

33 Сильный Нет Нет Работает Неполадки Да

35 Сильный Нет Нет Неполадки Неполадки Да

37 Слабый Ровное Хорошо Работает Неполадки Да

39 Слабый Ровное Хорошо Неполадки Неполадки Да

41 Слабый Ровное Плохо Работает Неполадки Да

43 Слабый Ровное Плохо Неполадки Неполадки Да

45 Слабый Ровное Нет Работает Неполадки Да

47 Слабый Ровное Нет Неполадки Неполадки Да

49 Слабый Ведет Хорошо Работает Неполадки Да

51 Слабый Ведет Хорошо Неполадки Неполадки Да

53 Слабый Ведет Плохо Работает Неполадки Да

55 Слабый Ведет Плохо Неполадки Неполадки Да

57 Слабый Ведет Нет Работает Неполадки Да

59 Слабый Ведет Нет Неполадки Неполадки Да

61 Слабый Нет Хорошо Работает Неполадки Да

63 Слабый Нет Хорошо Неполадки Неполадки Да

65 Слабый Нет Плохо Работает Неполадки Да

67 Слабый Нет Плохо Неполадки Неполадки Да

69 Слабый Нет Нет Работает Неполадки Да

71 Слабый Нет Нет Неполадки Неполадки Да

73 Нет Ровное Хорошо Работает Неполадки Да

75 Нет Ровное Хорошо Неполадки Неполадки Да

77 Нет Ровное Плохо Работает Неполадки Да

79 Нет Ровное Плохо Неполадки Неполадки Да

81 Нет Ровное Нет Работает Неполадки Да

83 Нет Ровное Нет Неполадки Неполадки Да

85 Нет Ведет Хорошо Работает Неполадки Да

87 Нет Ведет Хорошо Неполадки Неполадки Да

89 Нет Ведет Плохо Работает Неполадки Да

91 Нет Ведет Плохо Неполадки Неполадки Да

93 Нет Ведет Нет Работает Неполадки Да

95 Нет Ведет Нет Неполадки Неполадки Да

97 Нет Нет Хорошо Работает Неполадки Да

99 Нет Нет Хорошо Неполадки Неполадки Да

101 Нет Нет Плохо Работает Неполадки Да

103 Нет Нет Плохо Неполадки Неполадки Да

105 Нет Нет Нет Работает Неполадки Да

107 Нет Нет Нет Неполадки Неполадки Да

Найдена монотонная таблица:

Выведено новое правило:

IF Тормоза = Неполадки THEN Неисправность = Да

**Алгоритм индуцирования знаний**

Алгоритм С4.5:

**Шаг 1.** Для всех условных атрибутов *X*1, … *Xm* таблицы *T* вычисляем критерий разбиения Gain(*Xi*). Выбираем такой атрибут *X*, для которого Gain(*Xi*) максимально.

**Шаг 2.** Разбиваем таблицу по выбранному атрибуту на *N* подтаблиц. Проверяем каждую подтаблицу следующим образом.

**2.1.** Если подтаблица монотонна (все строки относятся к одному классу), то порождаемправило.

**2.2.** В противном случае рекурсивно применяем алгоритм C4.5 к полученной подтаблице

**Программа, реализующая алгоритм С4.5.**

import math

import pandas

from pandas import DataFrame

import numpy

columns = dict(Запах=DataFrame({"Запах": ["Сильный", "Слабый", "Нет"]}),

Движение=DataFrame({"Движение": ["Ровное", "Ведет", "Нет"]}),

Заводится=DataFrame({"Заводится": ["Хорошо", "Плохо", "Нет"]}),

Свет=DataFrame({"Свет": ["Работает", "Неполадки"]}),

Тормоза=DataFrame({"Тормоза": ["Работает", "Неполадки"]}))

def get\_table():

columns\_copy = columns.copy()

table = columns\_copy.pop(list(columns.keys())[0])

for column in columns\_copy:

table = table.merge(columns[column], how="cross")

table['Неисправность'] = numpy.where(

((table['Запах'] == 'Слабый') | (table['Запах'] == 'Нет')) &

((table['Движение'] == 'Ровное') | (table['Движение'] == 'Ведет')) &

((table['Заводится'] == 'Хорошо') | (table['Заводится'] == 'Плохо')) &

((table['Свет'] == 'Работает') & (table['Тормоза'] == 'Работает')), 'Нет',

'Да')

pandas.set\_option('display.max\_rows', None)

return table

def c45(table: pandas.DataFrame, condition):

# info\_table(table)

column, splits = next\_column(table)

for split in splits:

if is\_monotonous(split):

path = condition\_path(split, column, condition)

print\_condition\_path(\*path)

else:

condition[column] = split.iloc[0][column]

next\_table = delete\_column(split, column)

c45(next\_table, condition.copy())

def next\_column(table: pandas.DataFrame):

without\_defect = delete\_defect\_column(table)

max\_gain = 0

max\_column = None

max\_split = None

for column in without\_defect:

split = magic\_split(table, column)

column\_gain = gain(split, rows\_count(table), info\_table(table))

if max\_gain <= column\_gain:

max\_gain = column\_gain

max\_column = column

max\_split = split

return max\_column, max\_split

def delete\_defect\_column(table: pandas.DataFrame):

return delete\_column(table, 'Неисправность')

def delete\_column(table, column):

result = table.copy()

del result[column]

return result

def magic\_split(table: pandas.DataFrame, column: str):

tables = []

all\_values = columns[column]

for value in all\_values.values:

value = value[0]

tables.append(table.query(f"{column} == '{value}'"))

return tables

def rows\_count(table):

return len(table.index)

def info\_table(table: pandas.DataFrame) -> int:

number\_of\_yes = table.query("Неисправность == 'Да'").shape[0]

number\_of\_no = table.query("Неисправность == 'Нет'").shape[0]

total = number\_of\_yes + number\_of\_no

p\_yes = number\_of\_yes / total

p\_no = number\_of\_no / total

p\_yes = 1 if p\_yes == 0 else p\_yes

p\_no = 1 if p\_no == 0 else p\_no

return -p\_yes \* math.log2(p\_yes) - p\_no \* math.log2(p\_no)

def gain(tables: list[pandas.DataFrame], total\_row, main\_table\_info) -> int:

info\_t = 0

for table in tables:

info\_t = info\_t + info\_table(table) \* rows\_count(table) / total\_row

return main\_table\_info - info\_t

def is\_monotonous(table):

yes\_row\_count = rows\_count(table.query("Неисправность == 'Да'"))

return rows\_count(table) == yes\_row\_count or yes\_row\_count == 0

def condition\_path(table, final\_column, conditions):

final\_conditions = conditions.copy()

last\_column\_value = table.iloc[0][final\_column]

final\_conditions[final\_column] = last\_column\_value

result = table.iloc[0]["Неисправность"]

return final\_conditions, result

# print\_condition\_path(final\_conditions, result)

def print\_condition\_path(conditions, result):

rule = 'IF'

separator = ''

for column\_name, value in conditions.items():

rule = rule + f"{separator} {column\_name} = {value}"

separator = ','

rule = rule + f" THEN Неисправность = {result}"

print(rule)

def main():

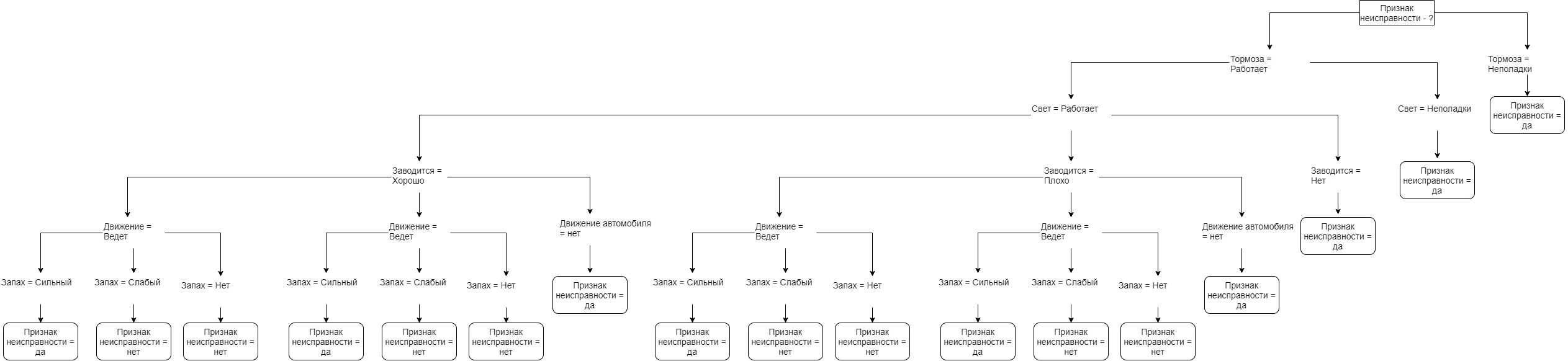
table = get\_table()

c45(table, dict())

if \_\_name\_\_ == '\_\_main\_\_':

main()

**Процесс получения дерева решений из таблицы**



**Новые продукционные правила**

1. IF Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ровное, Запах = Сильный   
   THEN Неисправность = Да
2. IF Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ровное, Запах = Слабый   
   THEN Неисправность = Нет
3. IF Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ровное, Запах = Нет   
   THEN Неисправность = Нет
4. IF Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ведет, Запах = Сильный   
   THEN Неисправность = Да
5. IF Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ведет, Запах = Слабый   
   THEN Неисправность = Нет
6. IF Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Ведет, Запах = Нет   
   THEN Неисправность = Нет
7. IF Тормоза = Работает, Свет = Работает, Заводится = Хорошо, Движение = Нет   
   THEN Неисправность = Да
8. IF Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ровное, Запах = Сильный   
   THEN Неисправность = Да
9. IF Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ровное, Запах = Слабый   
   THEN Неисправность = Нет
10. IF Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ровное, Запах = Нет   
    THEN Неисправность = Нет
11. IF Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ведет, Запах = Сильный   
    THEN Неисправность = Да
12. IF Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ведет, Запах = Слабый   
    THEN Неисправность = Нет
13. IF Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Ведет, Запах = Нет   
    THEN Неисправность = Нет
14. IF Тормоза = Работает, Свет = Работает, Заводится = Плохо, Движение = Нет   
    THEN Неисправность = Да
15. IF Тормоза = Работает, Свет = Работает, Заводится = Нет   
    THEN Неисправность = Да
16. IF Тормоза = Работает, Свет = Неполадки   
    THEN Неисправность = Да
17. IF Тормоза = Неполадки   
    THEN Неисправность = Да