Ниже представлены примеры оформления отчётов. Это не ГОТОВЫЕ 100% шаблоны по содержанию!!! Проявляйте осмысленный, критический подход при ознакомлении с материалом, учитесь принимать ответственность за свои решения, документировать, обосновывать, формулировать выводы и результаты!!!

ГУАП

КАФЕДРА № 42

ОТЧЕТ   
ЗАЩИЩЕН С ОЦЕНКОЙ

ПРЕПОДАВАТЕЛЬ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| профессор, д-р.т.н., профессор |  |  |  | В. В. Фомин |
| должность, уч. степень, звание |  | подпись, дата |  | инициалы, фамилия |

|  |
| --- |
| ОТЧЕТ О ЛАБОРАТОРНОЙ РАБОТЕ №1 |
| МЕТОД K-БЛИЖАЙШИХ СОСЕДЕЙ |
| по курсу: МЕТОДЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА |
|  |
|  |

РАБОТУ ВЫПОЛНИЛ

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| СТУДЕНТ гр. № | 4021 |  |  |  | О. А. Моисеенко |
|  |  |  | подпись, дата |  | инициалы, фамилия |

Санкт-Петербург 2024

1. **Цель работы:**

Изучение основ организация работы с технологической платформой для создания законченных аналитических решений KNIME, с использованием метода k-ближайших соседей.

1. **Задание:**

1. Для собственного набора данных реализовать кластеризацию методом k-ближайших соседей.

2. Выполнить оценку качества кластеризации.

3. Разработать визуальные отчеты для изучения кластеров.

1. **Ход выполнения работы:**

На рисунке 1 представлена схема рабочего процесса, в которую был добавлен блок Column Filter, для определения значимых атрибутов при анализе набора данных.

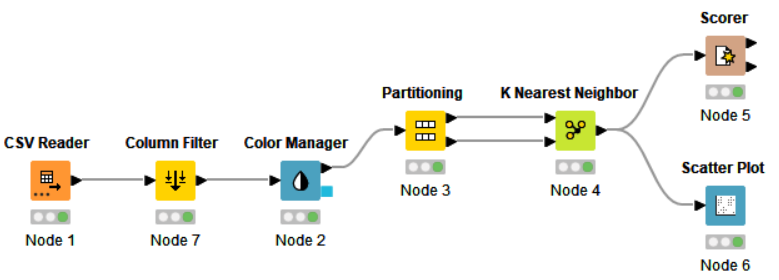


Рисунок 1 – Схема рабочего процесса

Для выполнения лабораторной работы был взят набор данных, содержащий прогнозы погоды. Набор данных состоит из 500 образцов со следующими видами погоды:

* sun;
* rain;
* drizzle;
* snow;
* fog.

Также были выбраны следующие значимые атрибуты:

* precipitation;
* temp\_max;
* temp\_min;
* wind.

На рисунке 2 представлен фрагмент результата классификации тестовой выборки. В столбце weather представлены начальные значения, а в столбце Class [kNN] метки класса. Проанализировав результат, можно сделать вывод, что больше всего отклонений в классах snow, drizzle и fog, т.к. в объем образцов с данным видом погоды в начальном наборе данных был значительно меньше, чем образцов с видом snow и rain.

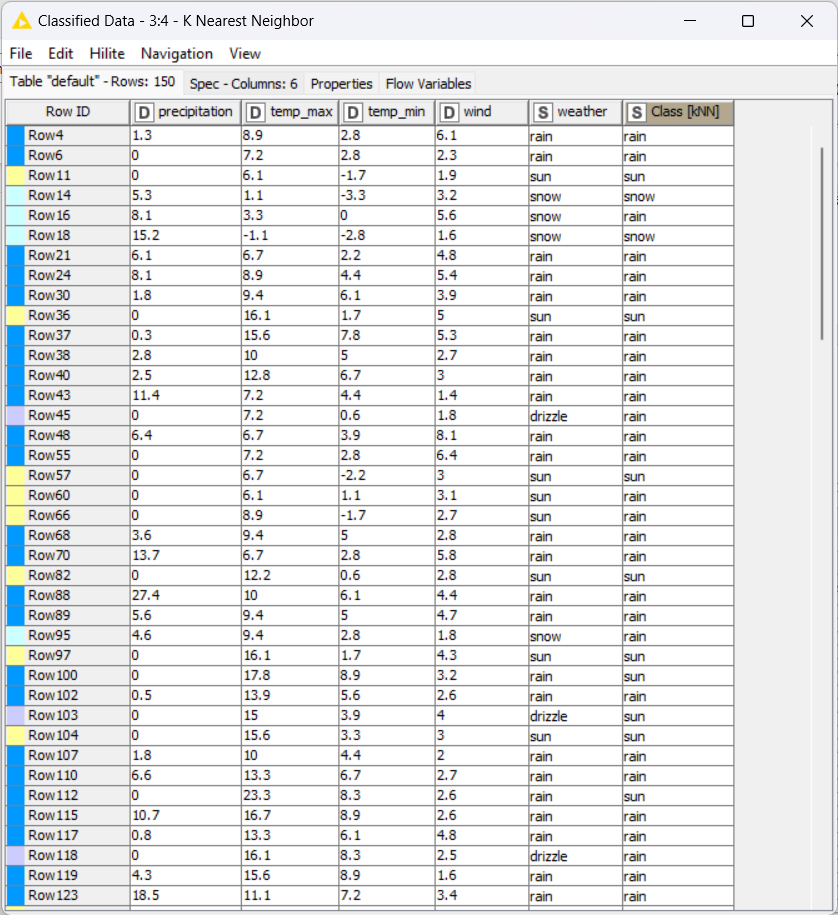


Рисунок 2 – Фрагмент классификации данных

Результат оценки классификатора в виде матрицы сопряженности представлен на рисунке 3. В результате получено 113 верных прогноза и 37 ложных срабатывания. Таким образом точность попадания составляет 75,333%.

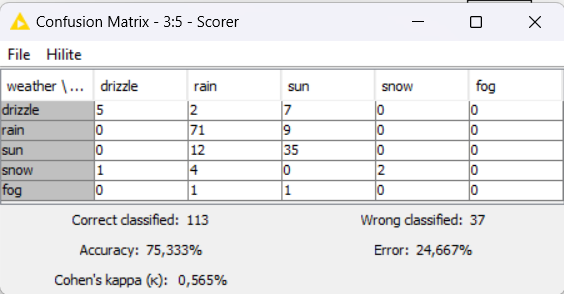


Рисунок 3 – Результат оценки классификатора

Для полной оценки качества классификации проанализируем основные метрики оценки качества представленные на рисунке 4:

1. Самые большие значения полноты содержат виды rain и snow, значительно превышающие значения полноты для остальных видов, как было сказано ранее, такой результат был получен в связи с малым процентом объема значений со значениями snow, fog и drizzle. Следовательно, узлу не хватило данных для выявления более точных закономерностей.
2. При том, что количество истинно-положительных решений для видов snow и drizzle мало, точность для данных видов приближенно равна 1, следовательно, те образцы, которые определены в данный класс, определены туда верно.

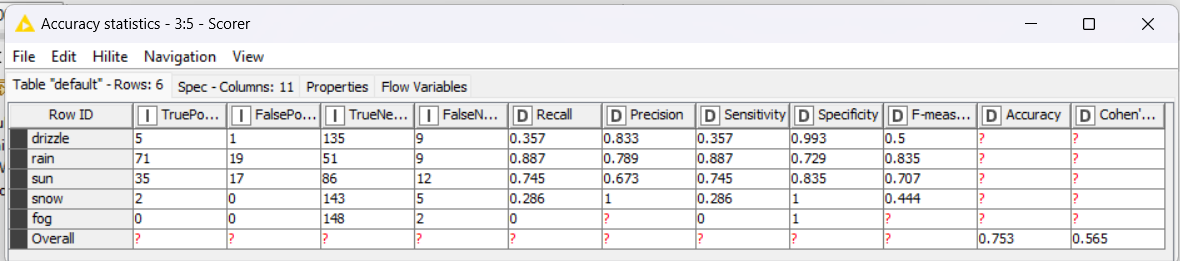


Рисунок 4 – Метрики оценки качества

На рисунке 5 представлена визуализация распределения данных с помощью узла Scatter Plot в виде диаграммы рассеивания. На горизонтальной оси отсчитываются значения параметра temp\_min, а на вертикальной – temp\_max.

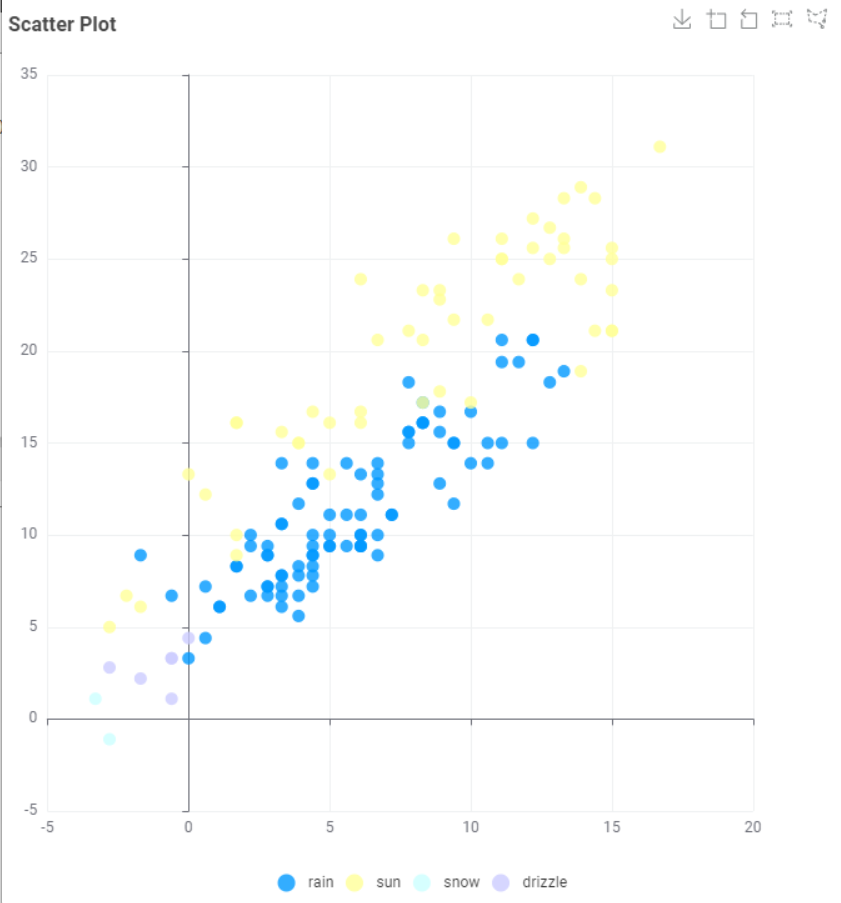


Рисунок 5 – Точечный график распределения данных

Анализируя полученный график, можно увидеть, что виды погоды располагаются достаточно близко друг к другу, в отрицательных температурах находятся виды snow, sun и drizzle, rain существует в дни температур с малой разницей, а sun – с большой.

Также существуют выбросы rain в дни с отрицательной температурой.

1. **Выводы:**

В ходе лабораторной работы был проведено изучение основ организация работы с технологической платформой для создания законченных аналитических решений KNIME, с использованием метода k-ближайших соседей.

При изучении метода k-ближайших соседей и анализе результатов был использован набор данных, имеющий 5 видов погоды, с преобладанием 2х типов: sun и rain. В результате анализа высокая полнота решений была именно у этих видов, однако, у видов snow и drizzle была повышенная точность решений.

Из вышесказанного можно сделать вывод, что наборы данных лучше брать с равномерным распределением данных в наборе.

ГУАП

КАФЕДРА № 42

ОТЧЕТ   
ЗАЩИЩЕН С ОЦЕНКОЙ

ПРЕПОДАВАТЕЛЬ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| профессор, д-р.т.н., профессор |  |  |  | В. В. Фомин |
| должность, уч. степень, звание |  | подпись, дата |  | инициалы, фамилия |

|  |
| --- |
| ОТЧЕТ О ЛАБОРАТОРНОЙ РАБОТЕ №2 |
| МЕТОД ЛИНЕЙНОЙ РЕГРЕССИИ |
| по курсу: МЕТОДЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА |
|  |
|  |

РАБОТУ ВЫПОЛНИЛ

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| СТУДЕНТ гр. № | 4021 |  |  |  | О. А. Моисеенко |
|  |  |  | подпись, дата |  | инициалы, фамилия |

Санкт-Петербург 2024

1. **Цель работы:**

Изучение основ организация работы с технологической платформой для создания законченных аналитических решений KNIME, с использованием метода линейной регрессии.

1. **Задание:**

1. Для собственного набора данных реализовать кластеризацию методом k-ближайших соседей.

2. Выполнить оценку качества кластеризации.

3. Разработать визуальные отчеты для изучения кластеров.

1. **Ход выполнения работы:**

На рисунке 1 представлена схема рабочего процесса.

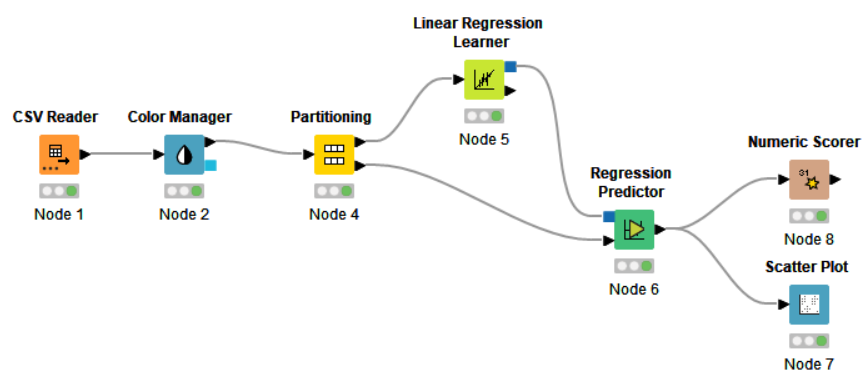


Рисунок 1 – Схема рабочего процесса

Для выполнения лабораторной работы был взят набор данных, содержащий информацию успеваемости обучающихся. Набор данных состоит из 50 образцов со следующими атрибутами:

* Порядковый номер студента;
* Среднее время, затраченное на обучение;
* Оценка студента;

На рисунке 2 представлены фрагмент прогнозируемые ответы тестовой выборки с использованием регрессионной модели. В столбце Marks представлены начальные значения, а в столбце Predication – прогнозируемые значения. Проанализировав результат, можно выявить следующую закономерность: чем ниже значение оценки, тем больше его расхождение в значении с спрогнозированной оценкой.

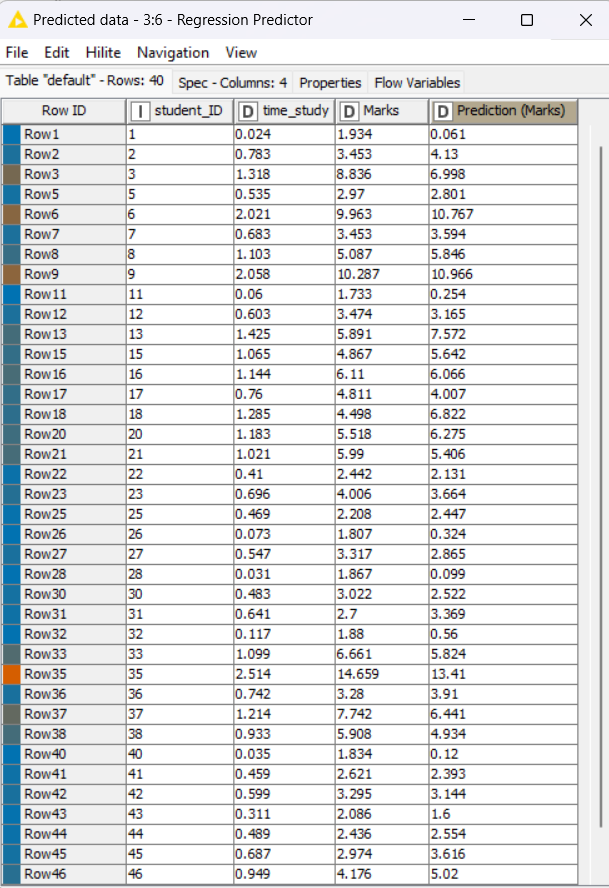


Рисунок 2 – Прогнозируемые ответы

Результат прогнозирования оценивается по параметрам модели, которые представлены на рисунке 3. Из значений данных параметров можно сделать следующие выводы:

Одним из самых важных параметров является коэффициент детерминации (R2), в нашем случае он приближенно равен 1, следовательно, модель работает хорошо, между значениями существует линейная зависимость и 86,9% дисперсии оценок студентов можно объяснить количеством потраченного на обучение времени.

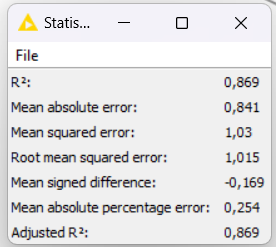


Рисунок 3 – Параметры модели линейной регрессии

На рисунке 5 представлена визуализация распределения данных с помощью узла Scatter Plot в виде диаграммы рассеивания. На горизонтальной оси отсчитывается время, потраченное студентом на обучение, а на вертикальной – спрогнозированное значение оценок.

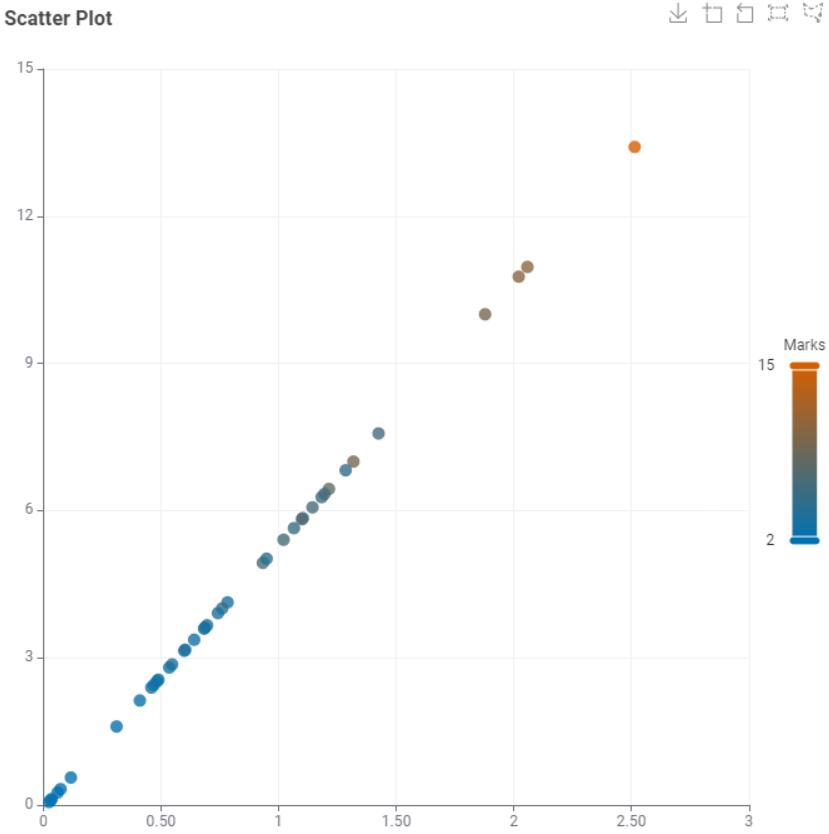


Рисунок 5 – Точечный график распределения данных

Анализируя полученный график, можно явно отследить линейную зависимость полученной студентом оценки от времени, потраченного им на обучение.

1. **Выводы:**

В ходе лабораторной работы был проведено изучение основ организация работы с технологической платформой для создания законченных аналитических решений KNIME, с использованием метода линейной регрессии.

При изучении метода линейной регрессии и анализе результатов был использован набор данных, содержащий 50 образцов с информацией об успеваемости обучающихся. В результате анализа было доказано существование линейной зависимости между оценкой обучающегося и временем, потраченным им на обучение. Полученная оценка прямо пропорционально зависит от времени обучения студента.

ГУАП

КАФЕДРА № 42

ОТЧЕТ   
ЗАЩИЩЕН С ОЦЕНКОЙ

ПРЕПОДАВАТЕЛЬ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| профессор, д.т.н., профессор |  |  |  | В. В. Фомин |
| должность, уч. степень, звание |  | подпись, дата |  | инициалы, фамилия |

|  |
| --- |
| ОТЧЕТ О ЛАБОРАТОРНОЙ РАБОТЕ №3 |
| ПРОГНОЗИРОВАНИЕ ТИПОВ ПОГОДЫ  С ИСПОЛЬЗОВАНИЕМ МЕТОДА ДЕРЕВА РЕШЕНИЙ |
| по курсу: МЕТОДЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА |
|  |
|  |

РАБОТУ ВЫПОЛНИЛ

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| СТУДЕНТ гр. № | 4021 |  |  |  | О. А. Моисеенко |
|  |  |  | подпись, дата |  | инициалы, фамилия |

Санкт-Петербург 2024

1. **Цель работы:**

Изучение основ организация работы с технологической платформой для создания законченных аналитических решений KNIME, с использованием метода деревьев решений.

1. **Задание:**

1. Для набора данных, включающих информацию о погоде реализовать классификацию методом дерева решений.

2. Выполнить оценку качества классификации.

3. Построить дерево решений и выявить набор логических правил.

1. **Описание предметной области:**

Для выполнения лабораторной работы был взят набор данных, содержащий прогнозы погоды из 150 образцов со следующими видами погоды:

* Sun – солнечно;
* Rain – дождливо;
* Snow – снежно.

Также были выбраны следующие значимые атрибуты:

* Precipitation – осадки, мм;
* temp\_max – максимальная суточная температура;
* temp\_min – минимальная суточная температура;
* wind – ветер, м/с.

На рисунке 1 представлена схема рабочего процесса.

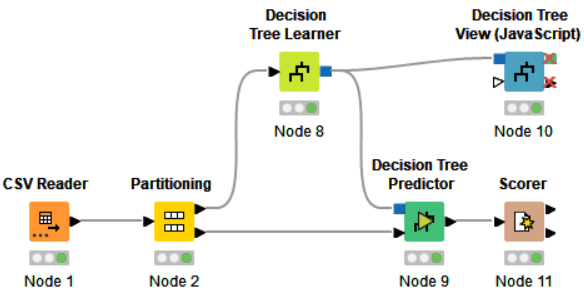


Рисунок 1 – Схема рабочего процесса

Результат оценки классификатора в виде матрицы сопряженности представлен на рисунке 2. В результате получено 40 верных прогнозов и 5 ложных срабатывания. Таким образом точность попадания составляет 88,889%, что значительно превышает точность 1 лабораторной работы (75,333%).

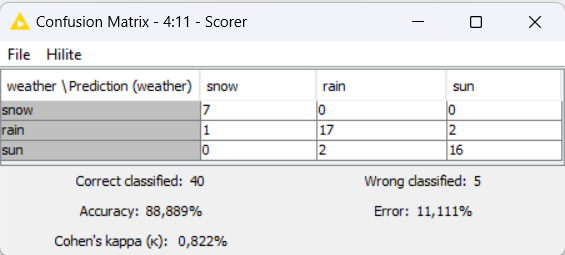


Рисунок 2 – Результат оценки классификатора

Для полной оценки качества классификации проанализируем основные метрики оценки качества представленные на рисунке 3:

Метрики полноты и точности для всех видов погоды приближенно равны 1, что свидетельствуем о высоком качестве и корректности результатов анализа.

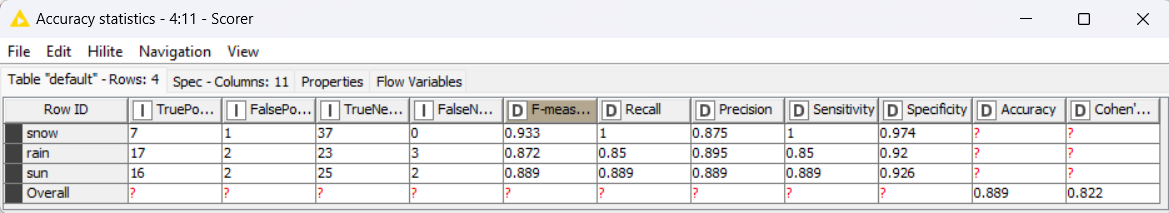


Рисунок 3 – Метрики оценки качества

На рисунке 4 представлена дерево данных, на основании которого было вычислено 3 логических правила:

1. (Если precipitation > 0.15 и temp\_min > 3.60) ИЛИ (если precipitation ≤ 0.15 и temp\_max ≤ 18.05 и wind > 2.80 и temp\_min > 5.30), то вид погоды *rain*;
2. Если precipitation > 0.15 и temp\_min ≤ 3.60 или , то вид погоды *snow*;
3. (Если precipitation ≤ 0.15 и temp\_max > 18.05) ИЛИ (если precipitation ≤ 0.15 и temp\_max ≤ 18.05 и wind ≤ 2.80) ИЛИ (если precipitation ≤ 0.15 и temp\_max ≤ 18.05 и wind > 2.80 и temp\_min ≤ 5.30), то вид погоды *sun*;

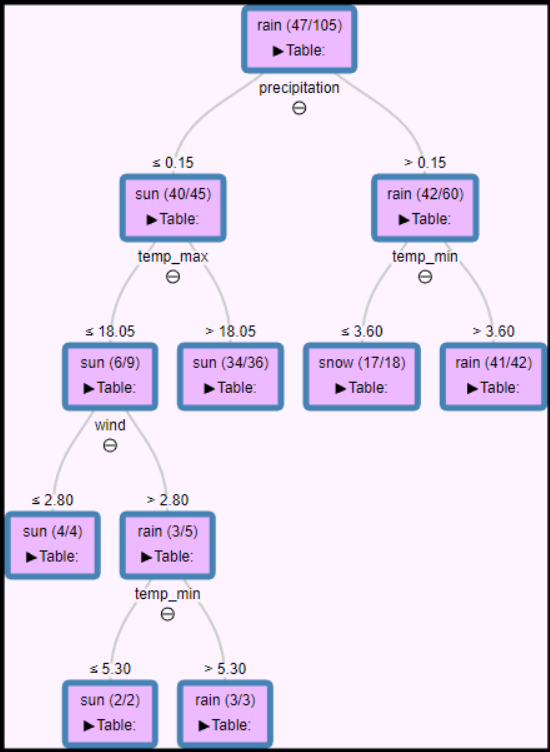


Рисунок 4 – Дерево решений

1. **Выводы:**

При изучении метода деревьев решений и анализе результатов был использован набор данных, имеющий 3 вида погоды. При сравнении данных, полученных с помощью метода k-средних и деревьев решений, можно сделать выводы, о том, что удаление видов погоды, имеющих недостаточное количество значений в наборе данных, значительно повлияло на качество анализа в лучшую сторону (точность попадания увеличилась с 75,333% до 88,889%).

Модель характеризуется точностью результатов предсказаний в 88.9%, а, следовательно, является работоспособной.

ГУАП

КАФЕДРА № 42

ОТЧЕТ   
ЗАЩИЩЕН С ОЦЕНКОЙ

ПРЕПОДАВАТЕЛЬ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| профессор, д.т.н., профессор |  |  |  | В. В. Фомин |
| должность, уч. степень, звание |  | подпись, дата |  | инициалы, фамилия |

|  |
| --- |
| ОТЧЕТ О ЛАБОРАТОРНОЙ РАБОТЕ №4 |
| ПРОГНОЗИРОВАНИЕ ТИПОВ ПОГОДЫ  С ИСПОЛЬЗОВАНИЕМ МЕТОДА НЕЙРОННЫХ СЕТЕЙ |
| по курсу: МЕТОДЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА |
|  |
|  |

РАБОТУ ВЫПОЛНИЛ

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| СТУДЕНТ гр. № | 4021 |  |  |  | О. А. Моисеенко |
|  |  |  | подпись, дата |  | инициалы, фамилия |

Санкт-Петербург 2024

1. **Цель работы:**

Изучение основ организация работы с технологической платформой для создания законченных аналитических решений использованием метода нейронных сетей.

1. **Задание:**

1. Для набора данных, включающих информацию о погоде реализовать классификацию методом нейронных сетей.

2. Выполнить оценку качества классификации.

1. **Описание предметной области:**

Для выполнения лабораторной работы был взят набор данных, содержащий прогнозы погоды из 150 образцов со следующими видами погоды:

* Sun – солнечно;
* Rain – дождливо;
* Snow – снежно.

Также были выбраны следующие значимые атрибуты:

* precipitation – осадки, мм;
* temp\_max – максимальная суточная температура;
* temp\_min – минимальная суточная температура;
* wind – ветер, м/с.

На рисунке 1 представлена модель рабочего процесса для классификации данных.

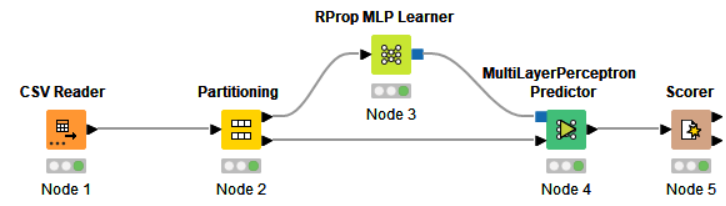


Рисунок 1 – Модель рабочего процесса

В результате обучения нейронной сети был построен график распространения ошибки на итерациях, на котором можно увидеть, что величина ошибки с течением итераций стремиться к 0, следовательно, обученную нейронную сеть можно считать работоспособной.

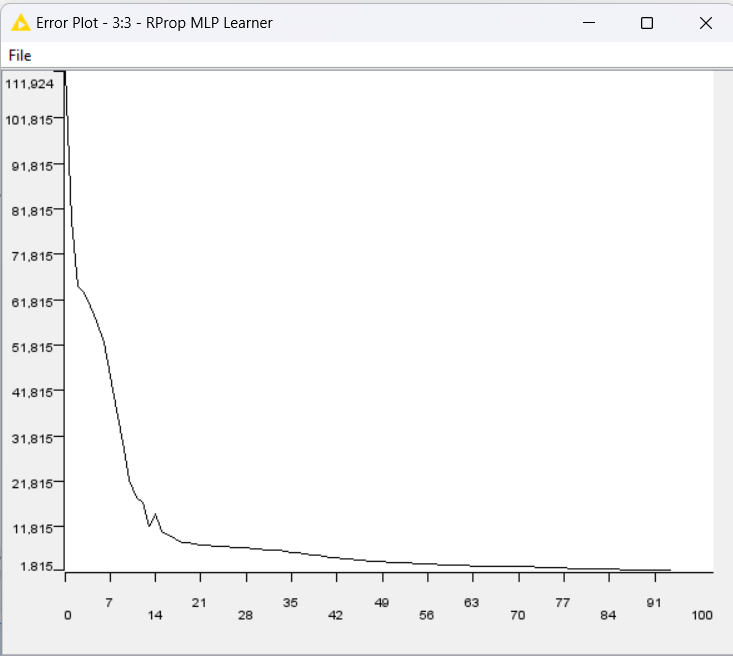


Рисунок 2 – График распространения ошибки

На рисунке 3 представлен фрагмент результата классификации данных на основе обученной модели многослойного перцептрона.

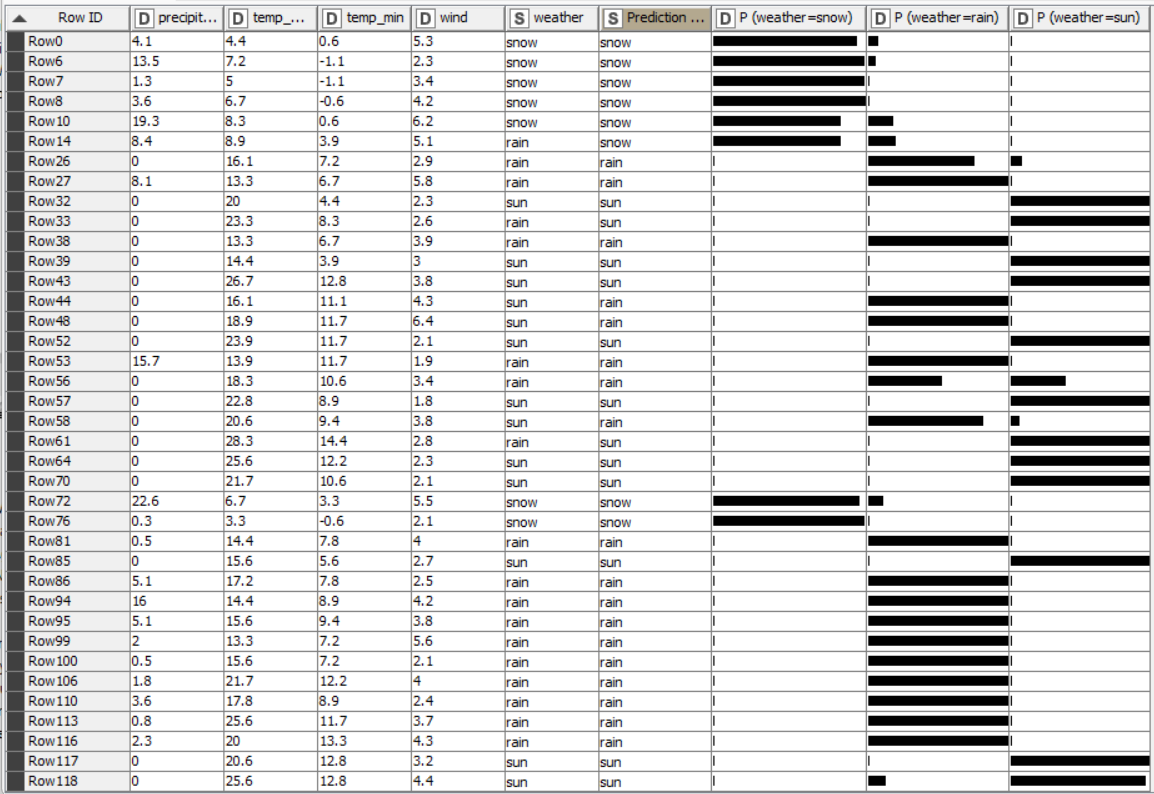


Рисунок 3 – Вероятность появления классов

Результат оценки классификатора в виде матрицы сопряженности представлен на рисунке 4. В результате получено 38 верных прогнозов и 7 ложных срабатывания. Таким образом точность попадания составляет 84,444%.

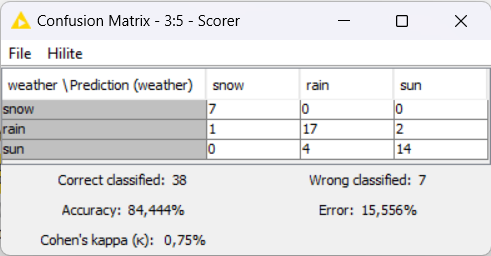


Рисунок 4 – Результат оценки классификатора

Для полной оценки качества классификации проанализируем основные метрики оценки качества представленные на рисунке 5:

Метрики полноты и точности для всех видов погоды приближенно равны 1, что свидетельствуем о высоком качестве и корректности результатов анализа.

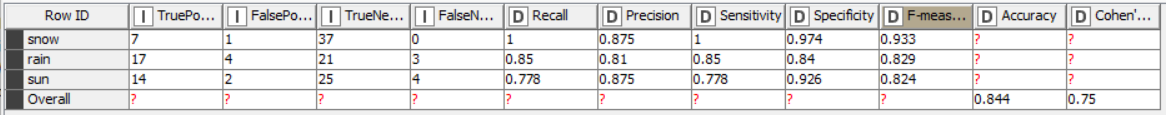


Рисунок 5 – Метрики оценки качества

1. **Выводы:**

Точность попадания составляет 84,444%, что говорит о работоспособности модели. Однако, в связи с тем, что у объектов разных классов могут быть схожие показатели атрибутов, в обучении наблюдаются незначительные ошибки.

ГУАП

КАФЕДРА № 42

ОТЧЕТ   
ЗАЩИЩЕН С ОЦЕНКОЙ

ПРЕПОДАВАТЕЛЬ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| профессор, д.т.н., профессор |  |  |  | В. В. Фомин |
| должность, уч. степень, звание |  | подпись, дата |  | инициалы, фамилия |

|  |
| --- |
| ОТЧЕТ О ЛАБОРАТОРНОЙ РАБОТЕ №5 |
| ПРОГНОЗИРОВАНИЕ ТИПОВ ПОГОДЫ  С ИСПОЛЬЗОВАНИЕМ БАЙЕСОВСКОГО ПОДХОДА |
| по курсу: МЕТОДЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА |
|  |
|  |

РАБОТУ ВЫПОЛНИЛ

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| СТУДЕНТ гр. № | 4021 |  |  |  | О. А. Моисеенко |
|  |  |  | подпись, дата |  | инициалы, фамилия |

Санкт-Петербург 2024

1. **Цель работы:**

Изучение основ организация работы с технологической платформой для создания законченных аналитических решений использованием наивного Байесовского подхода.

1. **Задание:**

1. Для набора данных, включающих информацию о погоде реализовать классификацию с помощью Байесовского подхода.

2. Выполнить оценку качества классификации.

1. **Описание предметной области:**

Для выполнения лабораторной работы был взят набор данных, содержащий прогнозы погоды из 150 образцов со следующими видами погоды:

* Sun – солнечно;
* Rain – дождливо;
* Snow – снежно.

Также были выбраны следующие значимые атрибуты:

* precipitation – осадки, мм;
* temp\_max – максимальная суточная температура;
* temp\_min – минимальная суточная температура;
* wind – ветер, м/с.

На рисунке 1 представлена модель рабочего процесса для классификации данных.

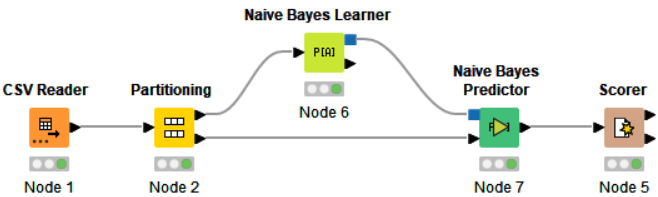


Рисунок 1 – Модель рабочего процесса

На рисунке 2 представлена таблица статистики, состоящая из средних значений и стандартного отклонения, для всех атрибутов по каждому типу погоды.

Можно заметить, что значения атрибута «ветер» для всех классов приближенно равно, следовательно, данный атрибут не имеет значимости при классификации.

Ключевым для класса «солнечно» является атрибут «осадки», а для класса «снежно» – минимальная и максимальная температура.

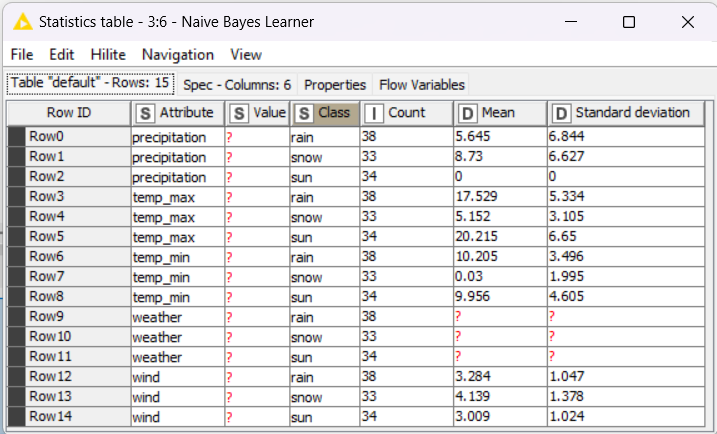


Рисунок 2 – Таблица статистики

На рисунке 3 представлен фрагмент таблицы классификации, построенную в результате выполнения модели, обученной на методах наивного байесовского подхода. В четырех последних столбцах содержатся вероятности классов для трех видов (классов) погоды, а также прогнозы предсказания.

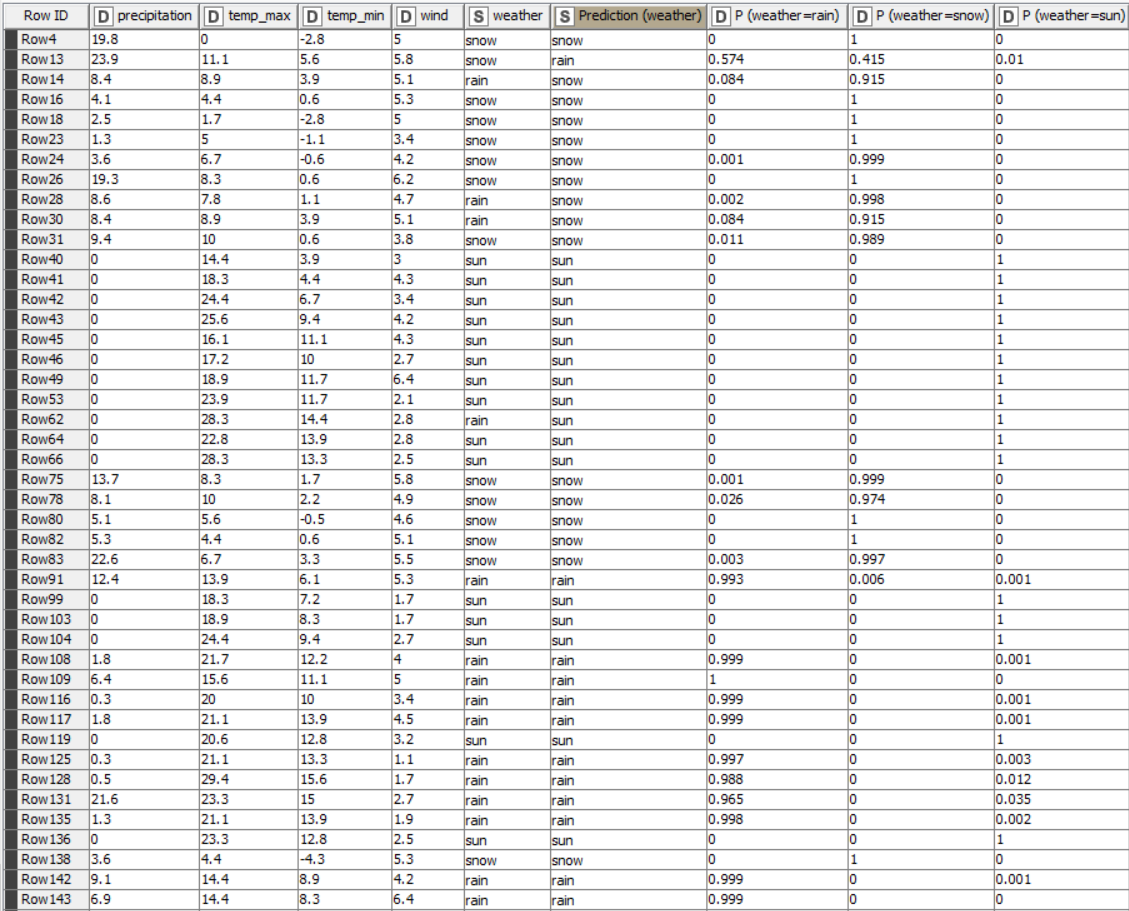


Рисунок 3 – Фрагмент таблицы классификации

Результат оценки классификатора в виде матрицы сопряженности представлен на рисунке 4. В результате получено 40 верных прогнозов и 5 ложных срабатывания. Таким образом точность попадания составляет 88,889%.

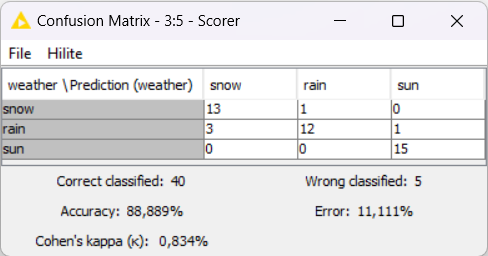


Рисунок 4 – Результат оценки классификатора

Для полной оценки качества классификации проанализируем основные метрики оценки качества представленные на рисунке 5:

Метрики полноты и точности для всех видов погоды приближенно равны 1, что свидетельствуем о высоком качестве и корректности результатов анализа.

Для класса «дождливо» значения метрик самые низкие, это произошло так как в статистических данных (рис. 2) у данного класса нет атрибутов, которые отличали бы его от остальных классов, следовательно, верно предсказать принадлежность образца к данному классу сложнее.

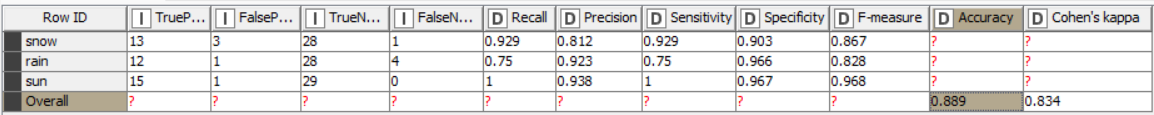


Рисунок 5 – Метрики оценки качества

1. **Выводы:**

Точность попадания составляет 88,889%, что говорит о работоспособности модели.

Каждый класс, кроме класса «дождливо», имеет ключевые атрибуты, по которым его можно опознать. Следовательно, прогнозирование дождливой погоды можно считать неточным, тогда как для солнечной и снежной предсказания можно считать достоверными.

ГУАП

КАФЕДРА № 42

ОТЧЕТ   
ЗАЩИЩЕН С ОЦЕНКОЙ

ПРЕПОДАВАТЕЛЬ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| профессор, д.т.н., профессор |  |  |  | В. В. Фомин |
| должность, уч. степень, звание |  | подпись, дата |  | инициалы, фамилия |

|  |
| --- |
| ОТЧЕТ О ЛАБОРАТОРНОЙ РАБОТЕ №6 |
| ПРОГНОЗИРОВАНИЕ ТИПОВ ПОГОДЫ  С ИСПОЛЬЗОВАНИЕМ МЕТОДА  МАШИННЫХОПОРНЫХ ВЕКТОРОВ (SVM) |
| по курсу: МЕТОДЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА |
|  |
|  |

РАБОТУ ВЫПОЛНИЛ

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| СТУДЕНТ гр. № | 4021 |  |  |  | О. А. Моисеенко |
|  |  |  | подпись, дата |  | инициалы, фамилия |

Санкт-Петербург 2024

1. **Цель работы:**

Изучение основ организация работы с технологической платформой для создания законченных аналитических решений использованием метода машинных опорных векторов.

1. **Задание:**

1. Для набора данных, включающих информацию о погоде реализовать классификацию с помощью метода машинных опорных векторов.

2. Выполнить оценку качества классификации.

1. **Описание предметной области:**

Для выполнения лабораторной работы был взят набор данных, содержащий прогнозы погоды из 150 образцов со следующими видами погоды:

* Sun – солнечно;
* Rain – дождливо;
* Snow – снежно.

Также были выбраны следующие значимые атрибуты:

* precipitation – осадки, мм;
* temp\_max – максимальная суточная температура;
* temp\_min – минимальная суточная температура;
* wind – ветер, м/с.

1. **Основные результаты и сценарий работы:**

На рисунке 1 представлена модель рабочего процесса для классификации данных.

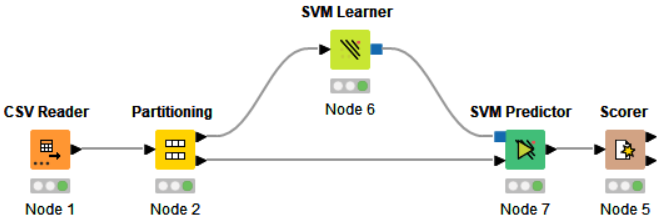


Рисунок 1 – Модель рабочего процесса

На рисунке 2 представлен фрагмент опорных векторов, где каждая строка представлена вектором, который описывается четырьмя атрибутами, охватывающими пространство, а последнее значение указывает класс.

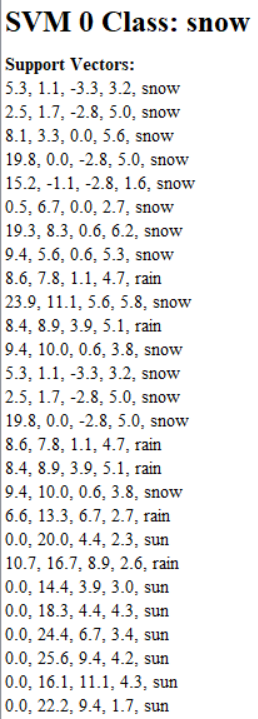


Рисунок 2 – Фрагмент представления опорных векторов

На рисунке 3 представлен фрагмент таблицы классификации, построенную в результате выполнения модели, обученной на методах опорных машинных векторов. В четырех последних столбцах содержатся вероятности для трех видов (классов) погоды, а также прогнозы предсказания.

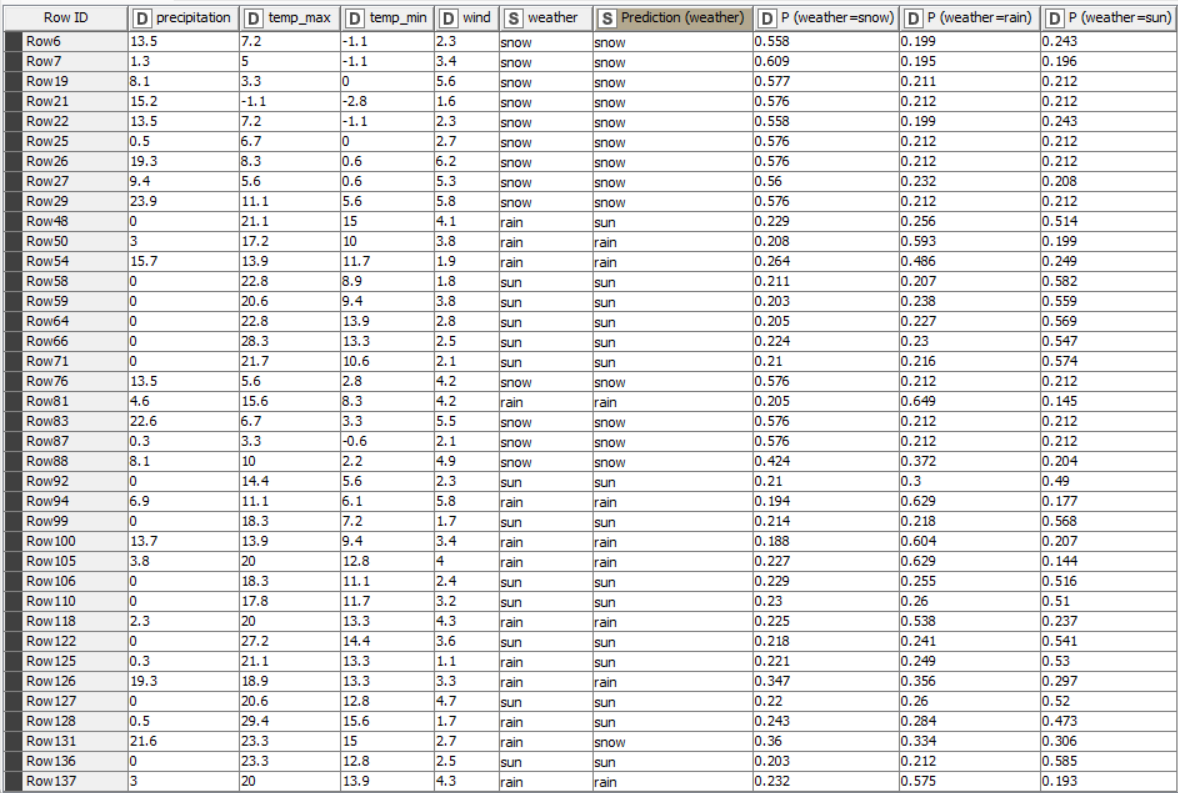


Рисунок 3 – Фрагмент таблицы классификации

Результат оценки классификатора в виде матрицы сопряженности представлен на рисунке 4. В результате получено 39 верных прогнозов и 6 ложных срабатывания. Таким образом точность попадания составляет 86,667%.

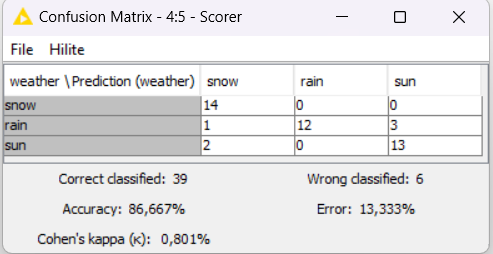


Рисунок 4 – Результат оценки классификатора

Для полной оценки качества классификации проанализируем основные метрики оценки качества представленные на рисунке 5:

Метрики полноты и точности для всех видов погоды приближенно равны 1, что свидетельствуем о высоком качестве и корректности результатов анализа.

Для класса «дождливо» значения метрик самые низкие, это произошло так как в статистических данных (рис. 2) у данного класса нет атрибутов, которые отличали бы его от остальных классов, следовательно, верно предсказать принадлежность образца к данному классу сложнее.

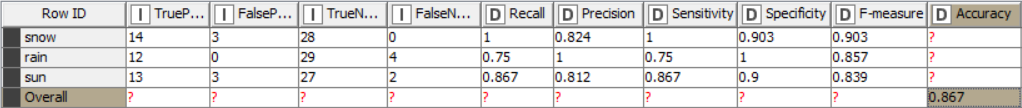


Рисунок 5 – Метрики оценки качества

1. **Выводы:**

Точность попадания составляет 86,667%, что говорит о работоспособности модели.