

ОЦЕНКА ЭФФЕКТИВНОСТИ ДЕРЕВА РЕШЕНИЙ В ЗАДАЧЕ РАСПОЗНАВАНИЯ ВОЗДУШНЫХ ОБЪЕКТОВ ПО ИХ ТРАЕКТОРНЫМ ПРИЗНАКАМ

Коломиец В. В., Свинарский М. В., Чжо Ту

Военная академия Республики Беларусь

Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: mechislav1993@gmail.com

Рассматривается задача радиолокационного распознавания воздушных объектов, актуальность которой обусловлена повышением точности и надежности принятия решений о классе наблюдаемого объекта. В качестве перспективного инструмента для решения этой задачи предлагается алгоритм машинного обучения с учителем – дерево решений С4.5. Приведены результаты оценки эффективности функционирования построенного дерева решений по тестовой выборке с искусственно введенным шумом.

ВВЕДЕНИЕ

В условиях стремительного развития технологий автоматизация процессов приобретает все большее значение. Неотъемлемой частью автоматизированных систем являются системы классификации, играющие ключевую роль в принятии решений, особенно в сферах, где последствия ошибок могут иметь критические значения. Военная сфера, где решения как правило принимаются в условиях ограниченного времени и высокой степени неопределенности, является ярким примером такой сферы. История насчитывает множество трагических случаев, когда гражданские самолеты сбивали как во время военных действий, так и в мирное время. Это подчеркивает актуальность задачи классификации воздушных объектов, особенно в контексте военного применения.

Основным источником информации о воздушной обстановке выступает радиолокационная станция, в которой решается три основные задачи [1]: обнаружение, измерение и распознавание. Под задачей распознавания принято понимать принятие решения о классе (типе) наблюдаемого объекта по выделенным классификационным признакам. В качестве основных классификационных признаков, которые могут выделять большинство радиолокационных станций, выступают траекторные признаки воздушных объектов [1]. Поскольку различные классы (типы) воздушных объектов обладают различными значениями высоты и скорости полета, то данные признаки являются информативными для решения задачи распознавания [1].

Важно отметить, что решение задачи распознавания необходимо осуществлять на дальних границах действия систем противовоздушной обороны с целью эффективного поражения наблюдаемого объекта. В связи с этим задача распознавания рассматривается применительно к радиолокационным станциям обзора, дальность действия которых составляет несколько сотен километров [1]. При этом особенностью движения воздушных объектов на дальних рубежах является

их полет на крейсерских значениях высоты и скорости полета, что обеспечивает максимальную дальность (продолжительность) полета объекта.

ОСНОВНАЯ ЧАСТЬ

В контексте решаемой задачи рассматривается применение дерева принятия решений. Деревья решений представляют собой эффективный класс моделей машинного обучения, отличающихся высокой точностью и интерпретируемостью [2]. Их иерархическая структура, состоящая из узлов и листьев (рисунок 1), обеспечивает наглядное представление информации, что делает их предпочтительными по сравнению с другими моделями. Узлы содержат решающие правила, которые применяются для классификации данных, а листья представляют собой результат принятия решения о классе наблюдаемого объекта [2].

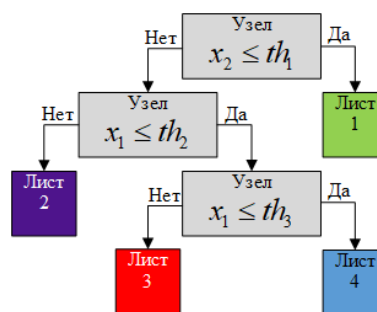


Рис. 1 – Пример дерева принятия решений

Процесс построения дерева решений основан на рекурсивном разбиении данных на подмножества, удовлетворяющие и не удовлетворяющие определенным правилам. Каждое подмножество затем снова делится на два, и так далее (рисунок 1). Процесс продолжается до тех пор, пока не будут выполнены условия останковки алгоритма. Следует отметить, что листьям соответствуют определенные области пространства признаков, где все объекты относятся к одному классу (рисунок 2).

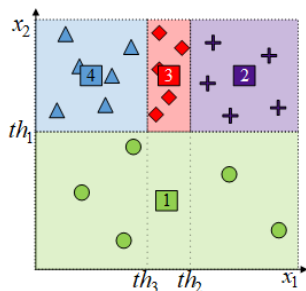


Рис. 2 – Пример областей принятия решений о классе объекта

Классификация новых объектов осуществляется путем прохождения по дереву от корневого узла до соответствующего листа. Каждый объект последовательно проходит по всем узлам, проверяя соответствие правилам, пока не достигнет листа, определяющего его класс. Такая иерархическая структура обеспечивает единственность решения, поскольку к каждому листу ведет только один путь, и объект может быть отнесен только к одному классу. Деревья решений представляют собой кусочно-постоянную аппроксимацию целевой зависимости, где аппроксимируемая функция является постоянной в каждой области пространства признаков, определяемой листом дерева. Отсутствие использования градиентных методов в процессе обучения обусловлено тем, что производная полученной функции равна нулю. При этом кусочно-постоянная аппроксимация может привести к переобучению, если дерево слишком сложное [2].

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

Для оценки эффективности алгоритма радиолокационного распознавания по крейсерским значениям высоты и скорости полета наблюдаемого объекта проводилось математическое моделирование. В качестве распознаваемых объектов рассматривалось четыре класса объектов: крылатая ракета, вертолет, винтовой самолет, реактивный самолет. По представленной выборке осуществлялось построение дерева принятия решений по алгоритму C4.5 со следующими гиперпараметрами: максимальная глубина дерева составляла 10, минимальное количество наблюдений в узле 2, минимальное количество наблюдений в листе 1. Результат построенного дерева принятия решений в виде графов представлен на рисунке 3.

Для проверки эффективности функционирования построенного дерева решений формировалась тестовая выборка, представляющая собой исходную выборку объектов с учетом ошибок измерения радиолокационной станции. При этом ошибки измерения распределены по нормальному закону распределения с нулевым математическим ожиданием и заданным среднеквадратическим отклонением. Так для оценки высоты полета ошибка составляла 200 м, а для скорости 5 м/с.

В качестве основных показателей эффективности функционирования алгоритма дерева решений были выбраны: точность (precision), полнота (recall), F1-мера (F1-score) [2]. Результаты математического моделирования сведены в таблицу 1.

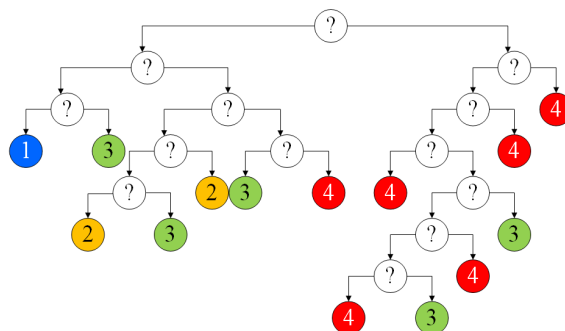


Рис. 3 – Вид построенного дерева решений

Таблица 1 – Результаты оценки эффективности алгоритма распознавания

Название	Precision	Recall	F1-score
Крылатая ракета	0.94	0.75	0.83
Вертолет	0.86	0.91	0.88
Винтовой самолет	0.62	0.47	0.53
Реактивный самолет	0.72	1	0.84
Среднее значение	0.78	0.78	0.77

Анализ представленных результатов позволяет утверждать, что анализируемый алгоритм в среднем характеризуется точностью 0.78, что является высоким результатом. При этом класс характеризующейся низким качеством распознавания является винтовой самолет. Данное обстоятельство обусловлено пересечением значений классификационных признаков с остальными классами, особенно с учетом ошибок измерения. Дальнейшие исследования направлены на анализ эффективности функционирования дерева решений построенного с учетом ошибок измерения высоты и скорости наблюдаемого объекта.

ЗАКЛЮЧЕНИЕ

Таким образом, полученные результаты демонстрируют перспективность использования алгоритмов дерева решений для решения задачи радиолокационного распознавания, особенно в условиях ограниченных вычислительных ресурсов и необходимости высокой интерпретируемости результатов.

СПИСОК ЛИТЕРАТУРЫ

1. Радиоэлектронные системы. Основы построения и теория: справочник / Я. Д. Ширман [и др.] ; под ред. Я. Д. Ширмана. – 2-е изд., перераб. и доп. – М. : Радиотехника, 2007. – 512 с.
2. Рашка, С. Python и машинное обучение: машинное и глубокое обучение с использованием Python, scikit-learn и TensorFlow 2 / С. Рашка, В. Мирджалили // 3-е изд.; Пер. с англ. – СПб. : ООО «Диалектика», 2020. – 848 с.