



АЛГОРИТМ ПРОГНОЗИРОВАНИЯ ПОДХОДЯЩЕГО УПРАВЛЕНИЯ ДЛЯ СЛОЖНЫХ СИСТЕМ НА ОСНОВЕ ТЕОРИИ РАСПОЗНАВАНИЯ ОБРАЗОВ

¹Т.Р.Нурмухаммедов

Ташкентский Государственный Транспортный Университет,
ntolaniddin@mail.ru,

²Ш.И.Юлдашев

Ташкентский университет прикладных наук,
yuldashev-1405@mail.ru,
<https://www.doi.org/10.5281/zenodo.7929954>

ARTICLE INFO

Received: 01st April 2023

Accepted: 14th April 2023

Online: 25th April 2023

KEY WORDS

Совокупность признаков, базовый набор, эмпирические данные, синтез, объекты эталоны, критерий качества.

ABSTRACT

В данной статье рассмотрены вопросы разработки метода, обобщающего функциональные возможности алгоритмов распознавания образов для повышения их качества. При решении практических задач распознавания широко используются метрические алгоритмы классификации и комбинаторная теория восстановления связи с прецедентной информацией. Также приведена задача исследования комбинаторных оценочных алгоритмов и программных средств, обладающих способностью обобщения.

Введение. В нашей стране, как и во всех сферах, большое внимание уделяется развитию информационно-коммуникационных и компьютерных технологий. В настоящее время все большее распространение информации, связанной с анализом возникающих объектов, увеличение применения новых методов порождает потребность в получении эффективных решений задач интеллектуального анализа данных. Методы интеллектуального анализа данных решают такие вопросы, как обработка данных, повышение качества и выявление скрытых закономерностей. Одним из подходов к повышению качества распознавания образов и снижению вычислительных затрат является предварительный анализ обучающих данных. Основной целью данного анализа является оценка основных информационных свойств образовательного комплекса, в частности, определение веса (информативности) признаков. Распознавание – это процесс обработки информации, осуществляемый некоторым модификатором информации (интеллектуальным информационным каналом, системой распознавания), имеющим вход и выход. На вход системы поступают сведения о свойствах представленных объектов. На выходе системы отображается информация о том, к каким классам (обобщенным символам) отнесены определенные объекты. При создании и эксплуатации автоматизированной системы распознавания образов решается ряд задач. Формулы этих задач и сам набор не совместимы у разных авторов, так как в той или иной мере зависят от конкретной



математической модели, основанной на той или иной системе распознавания. Также в определенных моделях распознавания некоторые задачи не имеют решения и не ставятся соответствующим образом. В таких случаях использование различных интенсивных методов в системе распознавания образов дает положительные результаты.

Отличительной особенностью интенсивных методов является то, что они используют различные описания свойств и их отношений как элементов операций при построении и применении алгоритмов распознавания образов. Такими элементами могут быть интервалы, средние и дисперсии отдельных значений или значений свойств, матрицы отношений свойств и т. п., в которых действия осуществляются аналитически или конструктивно. Однако объекты в этих методах не рассматриваются как неотъемлемые единицы информации, а служат индикаторами для оценки взаимодействий и поведения их атрибутов. Эти методы распознавания образов вытекают из классической теории статистических решений, в которой объекты исследования рассматриваются как реализация многомерной случайной величины, распределенной в пространстве свойств по некоторому закону. Они основают на байесовской схеме принятия решений, учитывающей априорные вероятности объектов, принадлежащих определенному известному классу, и условную плотность распределения значений векторов свойств. Эти методы сокращают для определения отношения вероятностей в различных областях многомерной характерной области.

Группа методов, основанных на оценке плотности распределения значений свойств, непосредственно связана с методами дискриминантного анализа. Байесовский подход к принятию решений является одним из самых передовых в современной статистике и представляет собой один из так называемых параметрических методов, для которых аналитическое выражение закона распределения (в данном случае нормального закона) известно и лишь незначительно. Необходимо оценить количество параметров (усредненные векторы и ковариационные матрицы) [1].

Методы исследования. Набор обобщенных алгоритмов вычисления оценок задается в виде $\{A\}$, а его элементы записываются в виде $A_q \in \{A\}$. Основная задача состоит в том, чтобы определить экстремальный алгоритм из множества $\{A\}$. Решение этой задачи зависит от переклассификации объектов обучающей выборки T_{nml} и определении качества алгоритма A в каждом процессе. Нахождение экстремального алгоритма, правильно определяющего класс распознаваемого объекта, приводит к снижению затрат на классификацию контролируемых объектов. В процессе обучения алгоритму A объекты обучающей выборки T_{nml} могут обучаться индивидуально или часть обучающей выборки (например, 10%) может быть взята в качестве контрольной выборки и реализована путем их обучения.

Каждый алгоритм задаётся распознаваемому объекту S в следующей формальной форме:

$$A(S) = (\alpha_1^A(S), \alpha_2^A(S), \dots, \alpha_l^A(S)), \alpha_i^A(S) \in \{0, 1, \Delta\}, i = 1, 2, \dots, l. \quad (1)$$



Здесь, если $\alpha_u^A(S) = 1$, алгоритм заносит объект S в класс K_u , если $\alpha_u^A(S) = \Delta$, то решение алгоритма S объект не принадлежит классу K_u , когда $\alpha_u^A(S) = 0$, этот алгоритм отменяет классификацию объекта S относительно класса K_u .

Ниже приведены сведения о функции качества, определяющей качество алгоритмов и составляющих их параметрах, а также схема и шаги, основанные на определении параметров обобщенных алгоритмов. Элементы $A_q \in \{A\}$ множества $\{A\}$ являются алгоритмами, Z ориентирован на решение задачи. Для каждого алгоритма A множества $\{A\}$ существует функция определения качества решения задачи Z , и она обозначается $\varphi_A(Z)$, $\varphi_A(Z) \geq 0$.

Находя количественные $\varphi(A)$ качественные функциональные значения параметрического алгоритма $A \in \{A\}$, необходимо определить алгоритм $A^* \in \{A\}$ такой, что $\varphi(A^*) = \underset{A \in \{A\}}{\text{extr}} \varphi(A)$ пусть значение функционала достигает минимума[1].

Если $\varphi(A)$ получить по отношению к количеству ошибок, то $\varphi(A) \rightarrow \min$, если выражать через качество обучения (в процентах), то $\varphi(A) \rightarrow \max$.

Стандартное функциональное качество распознавания образов пишется в виде

$$\varphi(A) = \frac{1}{nm} \sum_{i=1}^n \sum_{j=1}^m |\alpha_{ij} - \alpha_{ij}^A| \quad (2)$$

В некоторых публикациях приведены и другие способы его расчета.

Также при решении задачи Z A считается некорректным алгоритмом, но это не обязательно должен быть $\alpha_j^A(S'_j) = P_j(S'_j)$. Сейчас рассматриваются алгоритмы, отрицающие вычисление значения $P_j(S'_j)$, $1 \leq j \leq l$ для S'_j . Этот аргумент определяется как: $\alpha_j^A(S'_j) = \Delta$.

При решении задачи Z содержит в алгоритме A ряд дополнительных требований по классу некорректных алгоритмов. Пусть заданный множеств алгоритмов $\{A\}$ выглядит следующим образом:

$$A(I_0(K_1, \dots, K_l), I(S)) = (\beta_1^A(S), \dots, \beta_l^A(S)), \beta_1^A(S) \in \{0, 1, \Delta\}. \quad (3)$$

В таком наборе $\{A\}$ алгоритмов $\varphi(A)$ называется количественным функционалом - функциональным качеством или функцией качества алгоритма A . Основная задача, подлежащая выяснению (задача Z): из среды алгоритмов, определенных в (3), необходимо найти алгоритм A^* следующим образом:

$$\varphi(A^*) = \sup_{A \in \{A\}} \varphi(A) \quad (4)$$

Принцип алгоритма вычисления оценок основан на вычисление оценок, которые распознаются набором признаков и характеризуют «близость» эталонных объектов, отделенных от заданного набора признаков. Идею разработки алгоритмов вычисления оценок выдвинул академик РАН Ю.И. Журавлев в конце 60-х годов прошлого века. Научная школа, основанная Журавлевым на этой идее и направлении, до сих пор обновляет и развивает научные взгляды. Это направление широко распространено во многих республиках, и ученые научных школ, работающих в них, сотрудничают в развитии этой области (например, алгоритмы вычисления значений, алгебраическая теория признакового представления, определение значимости признаков и др.). Эта деятельность была отмечена во многих научных публикациях [2].



Результаты и обсуждение. В отличие от методов и алгоритмов, рассмотренных выше, алгоритм вычисления оценок представляет собой новый подход к характеристикам, характеризующим объекты. Для этого алгоритма объект существует одновременно в разных частях признакового пространства. Алгоритмы вычисления оценок дополняют идею использования признаков: так как не всегда понятно, какая комбинация признаков является информативной. В алгоритме вычисления оценок отклонение объектов рассчитывается с помощью всех комбинаций или предопределенных комбинаций описаний признаков, описывающих объект.

Используемая комбинация признаков (подмножество) называется авторами базовым набором или подмножеством объектов. Вводится общее понятие близости между распознаваемым, контрольным объектом и объектами так называемого эталонного образца (с определенной классификацией). Это сходство обнаруживается путем сравнения объекта управления с эталонными объектами на основе комбинации наборов деталей. Таким образом, традиционный алгоритм вычисления оценок представляет собой обобщенную, расширенную версию алгоритма k близких соседей. В этом алгоритме близость объектов может быть вычислена только в том случае, если они заданы одним пространством признаков.

Еще одним важным аспектом алгоритма расчета оценок является то, что адаптивная модель может быть построена на основе корректировки значений ряда параметров алгоритма на основе заданной учебной выборки. В качестве критерия качества получена ошибка распознавания. Параметры алгоритма вычисления оценок являются основными составляющими и задаются в виде лимитов, порогов. Принцип алгоритма вычисления оценок заключается в определении значений контрольных (распознаваемых) объектов и эталонных (ранее классифицированных) объектов, отражающих степень сходства. В системе набор для голосования представляет собой набор признаков, которые представляют заданный набор признаков. Опыт решения многих задач узнавания показывает, что во многих случаях основная «различимая» информация сопоставляется не только с одним признаком, но и с различными сочетаниями сочетаний признаков. Класс алгоритмов вычисления значений в настоящее время находится на стадии разработки идеи использования именно комбинации символов до логического конца, т.е. степень близости объектов определяется всеми возможными сравнениями комбинаций символов в Описание предметов требует многотомных циклических процессов, происходящих в обучении. Это приводит к увеличению технических и временных затрат. Поэтому необходима разработка подходов, направленных на снижение накладных расходов, возникающих в семействе алгоритмов. Базовый подход к вычислению значений параметров ориентирован на процесс обучения или оптимизации, а полученные результаты должны служить для достижения высокоточного качества распознавания. Для выбора параметров, повышающих качество распознавания, с использованием алгоритмов расчета значений и методов определения их значений необходимо решить следующие вопросы: выбор соответствующей метрической шкалы для признаков; определение необходимых параметров; найти оптимальные значения имеющихся параметров.

Определение всех параметров алгоритмов вычисления оценок осуществляется последовательно в 6 этапов. Модель IRIS использовалась для экспериментальных исследований, основанных на этом алгоритме. В ней 150 объектов разбиты на 3 класса по 50 и описываются 4 признаками. Таким образом, набор эталонных объектов отделяется от исходной обучающей выборки и обеспечивает корректность классификации. Сложность выбора базовых объектов связана с тем, что структура прецедентов каждого класса зависит от выбора других классов в качестве базовых. В более общем случае альтернативное решение проблемы потребовало бы рассмотрения всех вариантов посредством полного перебора [3].

С помощью приведенных выше алгоритмов на основе статистических данных, зафиксированных в республике в 2014-2021 годах – количество пожаров, число погибших и пострадавших в результате пожара, материальный ущерб, причиненный в результате пожара, - результаты прогнозирования данных показателей на 2022-2025 годы приведены в следующих таблицах и графиках:

Таблица 1

Прогноз возможных пожаров к 2025 г.

Год	2022	2023	2024	2025
Количество пожаров	733	697	661	624

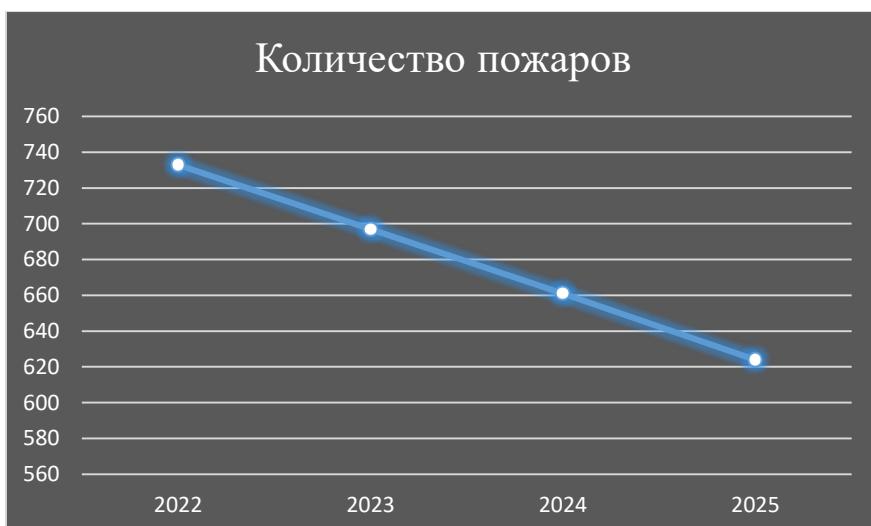


Рис.1. Информация о пожарах представлена в виде диаграммы

Таблица 2

Прогноз числа людей, которые могут погибнуть к 2025 году

Год	2022	2023	2024	2025
Количество погибших	10	10	9	9

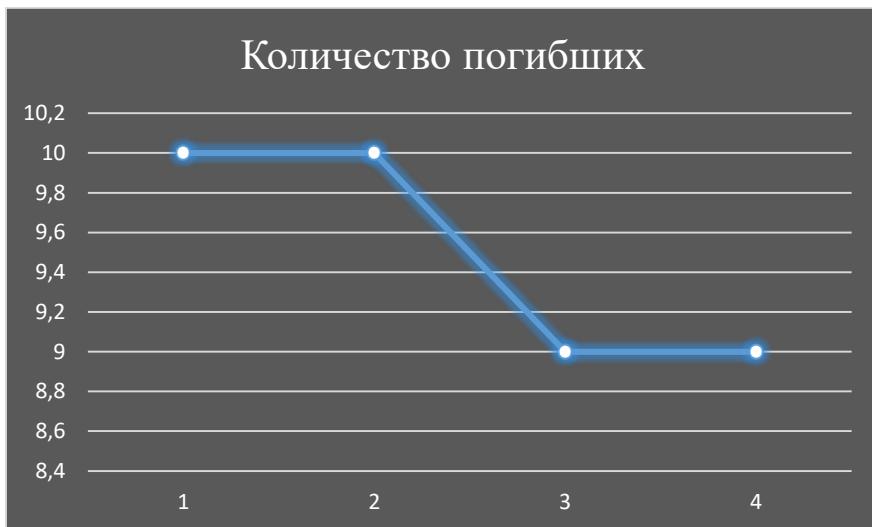


Рис.2. Число погибших в виде диаграммы

Таблица 3

Прогнозные данные о количестве людей, которые могут получить травмы до 2025 года

Год	2022	2023	2024	2025
Количество травм	6	5	4	2

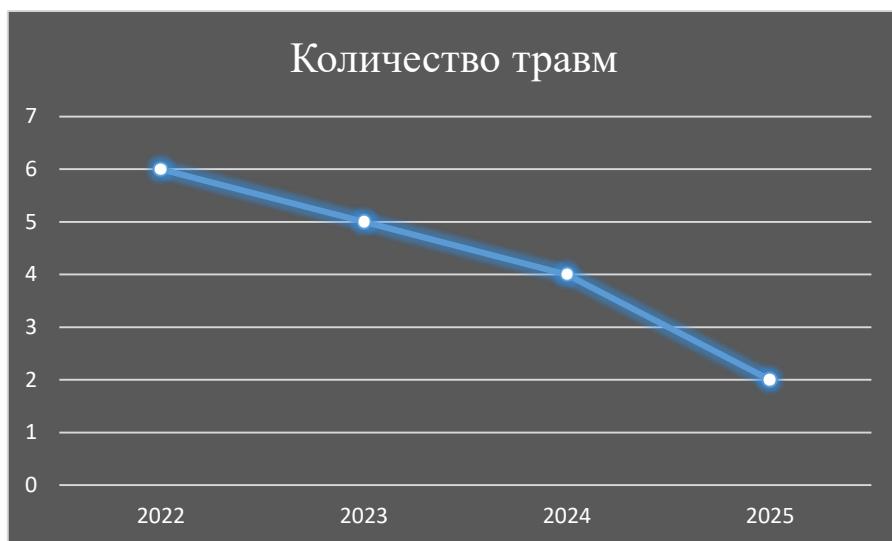


Рис.3. Количество пострадавших в виде диаграммы

Заключение. Это один из частных случаев в разработке алгоритмов распознавания образов, который позволяет оценить способность алгоритмов распознавать процесс синтеза эмпирических данных и решать такие задачи, как управление процессами и статистический анализ. Эти алгоритмы также разрабатываются как способ обобщения функциональных возможностей для повышения качества алгоритмов распознавания образов.



References:

1. Воронцов К.В. комбинаторные оценки качества обучения по прецедентам // Докл. РАН.-2004. -Т.394, №2. –С.175-178.
2. Худайбердиев М.Х., Хамроев А.Х., Мамиева Д.З. Алгоритм выставления оценок при формировании обучающего и контрольного выбора объектов. Республиканская научно-техническая конференция: «Проблемы информационных и телекоммуникационных технологий». - Ташкент, 10-11 марта 2016 г. –С. 185-187.
3. Худайбердиев М.Х., Алимов И.А. О способах повышения качества метрических алгоритмов классификации. Республиканская научно-техническая конференция: «Роль информационно-коммуникационных технологий в инновационном развитии реальных секторов экономики». Ташкент, 2017, 6-7 апреля. –С.39-41 .
4. Юлдашев Ш.И. О метрических алгоритмах определения символов для уточнения групп значимых объектов. // Научный вестник НамГУ, 2021, №1, С. 22-26.