

что позволяет улучшить качество учебного процесса и достичь лучших результатов.

В целом, опыт внедрения веб-приложения в учебном процессе для двух учебных дисциплин можно считать положительным. После начала использования веб-приложения в масштабах всего факультета появится возможность сравнивать успеваемость студентов по различным учебным дисциплинам и находить корреляции между ними. Это позволит своевременно обнаруживать и даже предупреждать появление академических задолженностей в масштабах всего факультета.

ЛИТЕРАТУРА

1. О применении системы дистанционного обучения для компьютерного тестирования знаний студентов по дисциплине «Физическая химия» / А.К. Болвако, Г.П. Дудчик // Труды БГТУ. Сер. VIII, Учеб.-метод. работа. – 2015. – № 8 (181). – С. 124–127.

2. Урбанович, П.П. Дистанционное обучение: тенденция, естественный процесс или вынужденная мера? / П.П. Урбанович, Е.А. Блинова, Н.В. Ржеутская // VIII Международная научно-техническая интернет-конференция «Информационные технологии в образовании, науке и производстве», 21-22 ноября 2020 года [Электронный ресурс] / Белорусский национальный технический университет; сост. Е.В. Кондратёнок. – Минск: БНТУ, 2020. – С. 116–122.

3. Пилинога, М.М. Веб-приложение для оценки текущей успеваемости студентов / М.М. Пилинога, Е.А. Блинова //Информационные технологии. Физика и математика: материалы 88-й науч.-техн. конф. профессорско-преподавательского состава, научных сотрудников и аспирантов (с международным участием), Минск, 29 января – 16 февраля 2024 года [Электронный ресурс] / отв. за издание И.В. Войтов; УО БГТУ. – Минск: БГТУ, 2024. – С. 46–49.

УДК 004.89

Л.В. Серебряная., зав. каф. (БГЭУ, БГУИР, г. Минск)

МОДЕЛИ И МЕТОДЫ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА И ОБРАБОТКИ ДАННЫХ

Термин «интеллектуальный» до недавнего времени относился к тем видам деятельности, которые были под силу только человеку. В настоящее время ситуация изменилась: активно создаются аппаратные и программные средства, способные имитировать действия человека в области обработки информации. Интеллектуальный анализ данных

особенно эффективен в применении к задачам, в которых отсутствует строгая формализация. Разработка интеллектуальных моделей и методов относится к общетеоретическим исследованиям технологии искусственного интеллекта. Сюда входят процедура машинного обучения и алгоритмы на ее основе, классификация данных, семантические и искусственные нейронные сети, которым посвящена данная работа [1].

Для решения поставленных задач необходимо уметь выявлять закономерности предметной области, что позволяют выполнить знания, которые не ограничиваются рамками данных. Информация, находящаяся в знаниях, должна содержать все необходимые сведения о предметной области, в которой решается задача.

Объемы информации, необходимой для решения поставленных задач, быстро растут. Для использования данных необходимы инструменты, способные извлекать скрытые закономерности и прогнозировать будущие события. Полезным для этого оказывается машинное обучение. Его суть заключается в создании структур данных, алгоритмов и моделей, которые способны автоматически извлекать знания из данных, решать задачи или предказывать результаты.

Машинное обучение можно представить в виде модели, похожей на «черный ящик». Он принимает на вход условие задачи, а на выходе формирует ответ. При этом у ящика есть много дополнительных параметров, влияющих на то, как именно будут обрабатываться данные [1].

Задачам, решаемым на основе машинного обучения, характерна неполнота исходных данных, которую должно дополнить машинное обучение. Такие задачи часто ассоциируют с методами принятия решений или распознаванием образов. Образ – это модель объектов, явлений или процессов реальной жизни. Образы задаются наборами признаков, которые являются самой важной информацией для машинного обучения и решения задачи распознавания образов.

Большинство задач интеллектуального анализа и обработки данных можно связать с двумя видами машинного обучения: с учителем, его еще называют контролируемым, и без учителя (самообучение). В контексте машинного обучения «учителем» может быть любое вмешательство человека в процесс обработки информации. В ходе процедуры обучения выполняется алгоритм работы с данными, часто это связано с классификацией или кластеризацией данных. В классы (кластеры) попадают объекты со схожими признаками. Массив исходных данных состоит из двух частей: обучающей и тестовой выборок. Если совокупность классов известна заранее, то применяется контролируемое обучение. Если, кроме признаков образов, нет никакой дополнительной информации, то используется самообучение [1].

В задачах интеллектуального анализа данных часто образы задаются векторами, чьи координаты являются признаками объектов. Многие методы анализа данных основаны на понятии пространства признаков и их обработке, которая связана с нахождением расстояний между объектами. Чем меньше расстояние между ними, тем вероятнее, что объекты относятся к одному классу (кластеру).

В качестве примера метода, использующего процедуру контролируемого обучения, можно привести алгоритм k -средних. Его цель – определить центрами классов k типичных образов и максимально компактно распределить вокруг них остальные объекты выборки [2].

Исходные данные алгоритма: число образов (векторов) и число классов (k), на которое нужно разделить все образы. Количества образов и классов выбираются произвольно. Признаки объектов задаются случайно, это координаты векторов. Обычно k элементов из набора векторов случайно назначают первоначальными центрами классов.

Примером метода, использующего процедуру самообучения, является алгоритм *максимина*. Его цель – исходя из произвольного выбора, максимально компактно разделить объекты на классы, определив центр каждого из них [2].

Исходные данные: число образов, которые нужно разделить на классы. Количество образов выбирается произвольно. Признаки объектов задаются случайным образом, это координаты векторов.

Анализ данных не всегда можно выполнить, оценивая только расстояния между объектами. Тогда рекомендуется использовать математический аппарат в виде решающих правил и разделяющих функций. Они применяются совместно, при этом разделяющие функции определяют границы между классами, а решающие правила отвечают на вопрос, к какому классу относить тот или иной объект. Разделяющую функцию часто представляют в виде линейной суммы:

$$f(\bar{x}) = \omega_0 + \omega_1 x_1 + \omega_2 x_2 + \dots + \omega_n x_n,$$

где ω_i – весовые коэффициенты, каждый из которых относится к определенной составляющей разделяющей функции. В процессе построения разделяющей функции основная задача заключается в том, чтобы найти весовые коэффициенты вида $\bar{\omega}_i = \{\omega_{0i}, \omega_{1i} \dots\}$ для каждого конкретного применения [2].

Для того чтобы ЭВМ имела возможность манипулирования знаниями о проблемной области, они должны быть представлены в виде модели. Модель представления знаний (МПЗ) – это способ и результат

формального описания знаний в базе знаний. Она должна быть понятной пользователю и обеспечивать однородность представления знаний.

В основе использования МПЗ лежит аксиоматический метод. В любой науке он состоит в том, что выделяется некоторое небольшое множество истинных утверждений, опираясь на которые можно вывести все истинные утверждения данной науки.

Рассмотрим одну из популярных МПЗ – семантическую сеть. Это – информационная модель предметной области, имеющая вид ориентированного графа. Его вершины соответствуют объектам предметной области, а дуги (ребра) задают отношения между ними. Следовательно, семантическая сеть отражает семантику предметной области в виде понятий и отношений. Семантические сети позволяют выделять смысл текста в виде понятий и связей между ними, образующих граф. Количество типов отношений в семантической сети определяется её разработчиком исходя из конкретных целей.

К достоинствам семантических сетей можно отнести следующее: универсальность, наглядность системы знаний, соответствие современным представлениям об организации долговременной памяти человека.

Еще один способ интеллектуального анализа и обработки данных – на основе искусственной нейронной сети (ИНС). ИНС – это математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей, т.е. сетей нервных клеток живого организма. Современные ИНС по сложности и «интеллекту» постоянно растут и развиваются, демонстрируя ряд ценных свойств [3].

Обучение. ИНС могут менять свое поведение в зависимости от внешней среды. После предъявления входных сигналов, возможно вместе с требуемыми выходами, они самонастраиваютя, чтобы обеспечить ожидаемую реакцию.

Обобщение. Отклик сети после обучения может быть до некоторой степени нечувствителен к небольшим изменениям входных сигналов. ИНС делает обобщение автоматически благодаря своей структуре.

Абстрагирование. Если, например, предъявить сети несколько искаженных вариантов входного образа, то она сама сможет создать на выходе «идеальный» образ, с которым никогда раньше не встречалась.

Рассмотрим одну из наиболее популярных моделей ИНС на основе контролируемого обучения – многослойный персепtron. Для его работы необходимо задать структуру сети: количество слоев и нейронов в каждом слое. Кроме того, задаются X – вектор входных сигналов сети, Y – вектор выходных сигналов, W – вектор весовых коэффициентов. Выбираются все пороговые значения сети и функция ее активации.

В результате этапа обучения на основе эталонных наборов данных вход-выход ИНС настраивается таким образом, чтобы в дальнейшем для произвольного входного сигнала выдать достаточно точный результат. Перед началом обучения весовые коэффициенты устанавливаются равными некоторым случайнym значениям. В процессе обучения сеть должна корректировать весовые коэффициенты так, чтобы максимально уменьшить значение общей ошибки. По завершении успешного обучения сети можно переходить к работе с тестовыми объектами [3].

Интеллектуальный анализ и обработка данных представляют собой мощный инструмент, способствующий преобразованию больших объемов данных в ценные знания и информацию. Применение этого подхода позволяет раскрывать скрытые закономерности и тенденции, что важно для принятия стратегических решений. В целом, развитие методов и технологий интеллектуального обработки данных обещает значительное усиление аналитических возможностей и эффективность принимаемых на их основе решений.

ЛИТЕРАТУРА

1. Гифт Н. Прагматичный ИИ. Машинное обучение и облачные технологии : науч. изд./ Н. Гифт ; пер. с англ. И. Пальти. – СПб : Питер, 2019. – 300 с.
2. Серебряная Л.В., Третьяков Ф.И. Методы и алгоритмы принятия решений: учеб.-метод. пособие для студ. спец. «Программное обеспечение информационных технологий» всех форм обуч. / Л. В. Серебряная, Ф. И. Третьяков. – Минск: БГУИР, 2014. – 50 с.
3. Серебряная, Л.В. Методы построения искусственных нейронных сетей для классификации данных / Л.В. Серебряная // Цифровая трансформация. – 2022. – Т. 28, № 1. – С. 20–26.

УДК 004.03.26+004.56

Д.В. Сазонова, асп.,
П.П. Урбанович, проф. (БГТУ, г. Минск)

АНАЛИЗ МОДЕЛЕЙ И СРЕДСТВ СТЕГАНОГРАФИЧЕСКОГО ПРЕОБРАЗОВАНИЯ ИНФОРМАЦИИ НА ОСНОВЕ НЕЙРОСЕТЕВЫХ ТЕХНОЛОГИЙ

Стеганография – наука о способах передачи или хранения информации при сохранении в тайне самого факта такой передачи (хранения) [1]. С развитием искусственного интеллекта (ИИ) интеграция сте-