

ВСТУПИТЕЛЬНОЕ СЛОВО

DOI: 10.31857/S0005231022120017, EDN: KREMGС

В данном номере журнала представлены статьи, подготовленные на основе докладов, прочитанных на конференции «Математические методы распознавания образов» (ММРО-2021), посвященной памяти академика К.В. Рудакова. Первая часть статей была опубликована в № 10, 2022.

В статье Гришина В.А. «Задача формирования покрытия области неопределенности эталонами для систем оптической навигации» рассматривается процесс управления стыковкой некооперируемых космических аппаратов.

В работе Скачкова Н.А., Воронцова К.В. «Улучшение качества машинного перевода с использованием обратной модели» рассмотрены проблемы компьютерного перевода текстов. Машинный перевод — это задача обработки текстов естественного языка, ставящая своей целью перевод входного текста с одного языка на другой язык в автоматическом режиме. В данной работе рассматриваются подходы, использующие переводные модели обратных языковых направлений и улучшающие согласованность между переводами одного текста с помощью прямых и обратных им моделей перевода. В работе представлено общее теоретическое обоснование для таких методов с точки зрения решения задачи максимизации правдоподобия, а также предложен способ стабильного обучения современных моделей с использованием циклических переводов.

В статье Дюличевой Ю.Ю. «Выявление аффективных состояний на основе автоматического анализа текстов комментариев в социальных сетях» рассмотрена задача классификации 3553 англоязычных комментариев из социальной сети Reddit на основе различных подходов к векторизации текстов комментариев: мешок слов, TF-IDF, анализ биграмм на основе поточечной взаимной информации PMI и сентимента, глубокая модель представления языка BERT. Применение гибридного подхода с использованием векторизации текстов на основе BERT и анализа биграмм позволило повысить качество классификации до 91%.

В работе Королева Н.С., Сенько О.В. «Метод повышения эффективности обучения градиентного бустинга, основанный на модифицированных функциях потерь» рассматривается новый метод повышения качества обучения градиентного бустинга, а также увеличения его обобщающей способности, основанный на использовании модифицированных функций потерь.

¹ Статьи с 3 стр. по 140 стр. являются окончанием тематического выпуска № 10, 2022 г.

В статье Хрыльченко К.Я., Воронцова К.В. «Оптимизация весов модальностей в тематических моделях транзакционных данных» рассматривается актуальная задача обучения мультимодальных тематических моделей. Одной из основных проблем при обучении моделей такого вида является поиск оптимальных весов модальностей. Авторами предложен и теоретически обоснован новый вычислительно эффективный метод нахождения “сбалансированных” весов модальностей, который решает проблему “доминирующих модальностей”. Предложенный метод экспериментально исследован на задаче предсказания 90-дневной просрочки выплат по кредиту. Авторами показано, что данный метод превосходит по качеству как стандартные мультимодальные модели, так и лучшую по качеству унимодальную модель.

Статья Ерохина В.И., Кадочникова А.П., Сотникова С.В. «Достаточные условия значимости коэффициентов линейных моделей и полиномиальной сложности их определения по данным с интервальной неопределенностью» посвящена исследованию свойств интервальных систем линейных алгебраических уравнений, которые, в частности, возникают при построении линейной регрессии по данным с интервальной неопределенностью. Авторский вклад состоит в формулировке и доказательстве теоремы о достаточных условиях, при которых допустимая область интервальной системы линейных алгебраических уравнений является выпуклым ограниченным многогранником, целиком принадлежащим одному ортанту.

В работе Берикова В.Б. «Модель и метод построения разнородного кластерного ансамбля» рассматривается задача кластеризации данных с помощью разнородного ансамбля с использованием матрицы коассоциации. Формулируется вероятностная модель, учитывающая коррелированность оценочных функций, с помощью которой находятся соотношения между характеристиками ансамбля и показателями качества итогового решения. Найдено выражение для оптимальных весов базовых алгоритмов, для которых минимальна верхняя граница оценки вероятности ошибки кластеризации.

В работе Анциперова В.Е. «Генеративная модель автокодировщиков, самообучающихся на изображениях, представленных выборками отсчетов», обосновывается концепция автокодировщиков, ориентированных на автоматическую генерацию сжатых изображений. Предлагаемый подход к задаче синтеза подобных автокодировщиков опирается на методы и принципы машинного обучения, понимаемого здесь как обучение по выборке из самих же данных. Для этих целей разработано специальное представление изображений с помощью выборок отсчетов контролируемого размера (выборочных представлений).

Лазарев А.А.