

УТВЕРЖДАЮ

Директор САО РАН,

Г.Г. Валявин

«12» сентября 2024 года



ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

Федерального государственного бюджетного учреждения науки «Специальная астрофизическая обсерватория Российской академии наук (САО РАН)» на диссертацию Авдеевой Александры Сергеевны «Исследование параметров звёзд и определение межзвёздного поглощения по данным больших современных обзоров неба», представленную на соискание ученой степени кандидата физико-математических наук по специальности 1.3.1. – «Физика космоса, астрономия».

Диссертационная работа Авдеевой Александры Сергеевны посвящена актуальным проблемам современной астрономии. Астростатистика и астроинформатика стремительно развиваются как научные области, изучающие пересечение наблюдательной астрономии, статистики, разработки алгоритмов и науки о данных. За последнее десятилетие количество статей по статистическим методам увеличилось в 2.5 раза, а по методам машинного обучения — в 4 раза за пять лет. Вызовы астрономической науке, связанные с большим и постоянно нарастающим объемом информации, а также неизбежностью ее анализа, определяют необходимость применения передовых методологий и техник. Задачи, ранее считавшиеся нерешаемыми традиционными методами, теперь находятся на переднем крае исследований благодаря достижениям в астростатистике и астроинформатике. Таким образом, диссертационное исследование, несомненно, является актуальным.

Содержание работы

Диссертация А.С. Авдеевой состоит из введения, четырех глав и заключения. Полный объем работы составляет 122 страницы, включая 33 рисунка и 16 таблиц. Список литературы содержит 99 наименований.

Во Введении обосновывается актуальность работы, формулируются цели и задачи диссертационной работы, раскрывается научная новизна и практическая значимость, указывается личный вклад диссертанта.

В диссертационной работе современные статистические методы применены к нескольким задачам. Задача поиска коричневых карликов по показателям цвета актуальна, поскольку эти объекты являются малоизученными, что затрудняет создание полной картины их распространенности и характеристик. Увеличение выборки коричневых карликов поможет уточнить их физические свойства и проверить модели эволюции. Первые две главы диссертации Александры Сергеевны Авдеевой сосредоточены на совершенствовании методов поиска коричневых карликов в данных современных фотометрических обзоров, таких как WISE, 2MASS, DES и Pan-STARRS.

В первой главе диссертации описывается разработка фотометрических правил для поиска коричневых карликов в обзорах WISE, 2MASS и DES, причем с классификацией их на три подгруппы: яркие, транзитные и слабые коричневые карлики. Применение разработанных критериев в области пересечения трех обзоров позволило обнаружить 135 коричневых карликов. Дополнительное исследование 96 объектов из этой выборки подтвердило их принадлежность типу коричневых карликов, что свидетельствует о надежности разработанных правил. Показано, что глубина обзора Gaia может быть недостаточной для идентификации, по крайней мере, трети коричневых карликов. Обнаруженные в результате поиска 43 коричневых карлика, которые отсутствуют в базе данных Simbad, являются новыми открытиями, что подтверждает обоснованность разработанных в результате исследования фотометрических правил.

Вторая глава посвящена применению методов машинного обучения для распознавания коричневых карликов в фотометрических обзорах WISE, 2MASS и Pan-STARRS. С помощью четырех методов машинного обучения - случайный лес (Random Forest), опорных векторов (Support Vector Machine), XGBoost и TabNet Classifier - были реализованы модели для распознавания коричневых карликов по фотометрическим данным. Все модели показали стабильно высокие результаты классификации, которые превосходят результативность методов, известных из литературы. Установлено, что модели, основанные на деревьях решений (RF и XGBoost), используют в значительной степени блески объектов, что несколько снижает надежность моделей. Два других метода - SVM и TabNet — опираются на показатели цвета, что уменьшает возможность ошибочной классификации слабых источников. Обнаружено, что показатель цвета ($i-y$), ранее не использовавшийся в других работах для селекции коричневых карликов, показывает высокую результативность. Результаты проделанной работы демонстрируют, что подход, который опирается на методы машинного обучения, открывает новые возможности для автоматизации и совершенствования поиска коричневых карликов.

В третьей главе описывается реализация методики оценки надежности эффективных температур звезд в третьем выпуске данных Gaia (Gaia DR3). Отмечено, что у значительного числа звезд наблюдаются большие отклонения в оценках эффективных температур между данными Gaia и спектроскопическими обзорами высокого разрешения. Произведено сравнение эффективных температур обзора Gaia с эффективными температурами спектроскопических обзоров высокого разрешения APOGEE и GALAH. Были использованы три алгоритма бустинга - XGBoost, CatBoost и LightGBM. Модели машинного обучения, которые были применены к данным модуля Gaia GSP-Phot, позволили выявить объекты с надежными величинами температур, совпадающих с эталонными температурами в пределах 250K. Ограничение данного подхода заключается в том, что модели в основном извлекают объекты с эффективными температурами менее 7000 K, что связано с отсутствием существенного количества более горячих звезд в обучающих данных. Были определены флаги качества для эффективных температур модуля Gaia GSP-Phot. По результату работы 66% звезд Gaia DR3 имеют надежные оценки температуры.

Четвертая глава диссертации А.С. Авдеевой посвящена определению межзвездного поглощения в выделенных направлениях на небе. Надежные величины межзвездного поглощения критически важны правильной интерпретации данных о звездах и галактиках. Поглощение света межзвездной средой и пылью искажает наблюдаемые спектры, изменяя их интенсивность и цвет, что может привести к ошибочным выводам о физических

свойствах небесных объектов. Методика определения межзвездного поглощения по эффективной температуре отдельных объектов применена к данным спектроскопического обзора RAVE для южного полушария неба. Для большинства исследуемых направлений были определены параметры закона косеканса. В работе был сделан вывод о том, что качество данных обзора RAVE недостаточно для получения более точных величин, чем известно на данный момент.

В Заключение сформулированы положения, выносимые на защиту, перспективы дальнейшей разработки темы исследования, сведения об апробации работы и приведен список публикаций.

Достоверность результатов, публикации, апробация

Достоверность полученных результатов и выводов в диссертационной работе обеспечивается обоснованным выбором объектов исследования, использованием современных данных наблюдений и апробированных методов обработки, а также согласованностью с опубликованными результатами других авторов. Достоверность результатов подтверждается пятью публикациями в рецензируемых журналах, рекомендованных ВАК.

Результаты проведенных исследований были представлены на различных всероссийских и международных конференциях и семинарах, в том числе на Всероссийской астрономической конференции ВАК-2021, международных конференциях «Аналитика и управление данными в областях с интенсивным использованием данных» (2022 и 2023 гг), международной конференции Astronomical Data Analysis Software and Systems (ADASS) в 2022 г., Всероссийской конференции «Физика звезд: теория и наблюдения» (2023 г.) и других.

Все результаты, выносимые на защиту, получены автором в результате совместных исследований, опубликованных с соавторами в научных статьях. При этом личный вклад автора в эти работы является определяющим. Соискатель принимала активное участие на всех стадиях исследования: в постановке задачи, подборе и анализе наблюдательных данных, подборе и обучении моделей, а также в обсуждении полученных результатов и их подготовке к публикации.

Вопросы и замечания

1. Почему при поиске коричневых карликов не использовались обзоры ближнего инфракрасного диапазона, проводящиеся на телескопе UKIRT, - LAS, GPS, UHS, которые глубже обзора 2MASS?
2. Не лишним было бы привести более подробное описание программного обеспечения OPTUNA.
3. На стр. 27 приводится ссылка на рис. 1.1, в то время как, исходя из контекста, подразумевается рис. 1.4.
4. В гл. 2 таблицы отнесены в конец главы, что неудобно при чтении и понимании текста.
5. На стр. 42 в тексте ссылка на рис. 2.4, хотя описание относится к рис. 2.3. К рисунку 2.4. в тексте главы нет объяснений.

6. В гл. 2 при рассмотрении задачи классификации ничего не сказано о том, какие объекты можно “спутать” с коричневыми карликами и достаточно ли они представлены в тренировочном наборе данных.
7. В гл. 3 рассматривается одинаковое абсолютное значение ошибки эффективной температуры для всего диапазона температур, в то время как разумнее было бы рассмотреть относительную ошибку.

Указанные недостатки не умаляют достоинств диссертации и не влияют на положения, выносимые на защиту. Обоснованность и достоверность выводов диссертации обеспечивается высоким авторитетом изданий, в которых опубликованы результаты. Полученные данные могут быть использованы в ИНАСАН, ИКИ РАН, САО РАН, ГАИШ МГУ и других как отечественных, так и зарубежных научных организациях. Хочется отметить применение нескольких алгоритмов машинного обучения к решению астрофизических задач, а также высокий уровень владения языком в донесении своих исследований до читателя.

Заключение

Диссертация Авдеевой Александры Сергеевны «Исследование параметров звёзд и определение межзвёздного поглощения по данным больших современных обзоров неба» является законченным научным исследованием, выполненным на высоком уровне, и отвечает требованиям, установленным к работам подобного рода в ВАК. Содержание диссертации соответствует специальности 1.3.1. Физика космоса, астрономия (по физико-математическим наукам). Автореферат диссертации в полном объеме отражает основные результаты, полученные в работе. Таким образом, соискатель Авдеева Александра Сергеевна достойна присуждения ученой степени кандидата физико-математических наук по специальности 1.3.1. Физика космоса, астрономия.

Текст отзыва рассмотрен и одобрен астрофизическим семинаром САО РАН 12 сентября 2024 г.

Заведующая лабораторией информатики САО РАН,
кандидат физико-математических наук
тел. +7 (928) 3889832,
email: zhe@sao.ru



О.П. Желенкова

Сведения о ведущей организации:

Федеральное государственное бюджетное учреждение науки «Специальная астрофизическая обсерватория Российской академии наук (САО РАН)», 369167, п. Нижний Архыз, Зеленчукский район, Карачаево-Черкесская республика, САО РАН.

тел. +7 (87878) 46336, e-mail: admsao@sao.ru