



Столыпинский

вестник

Научная статья
Original article
УДК 004.852

**КЛАССИФИКАЦИЯ РЕЗУЛЬТАТОВ СОРЕВНОВАНИЙ ПО
КОННОМУ СПОРТУ НА ОСНОВЕ МЕТОДА К-БЛИЖАЙШИХ
СОСЕДЕЙ**

**CLASSIFICATION OF THE RESULTS OF EQUESTRIAN COMPETITIONS
BASED ON THE K-NEAREST NEIGHBOR METHOD**

Катковская К.В., старший преподаватель кафедры «Информационные системы», ФГБОУ ВО СГУВТ, Россия, г. Новосибирск

Барц К.С., студент магистратуры, 1 курс, факультет «Электромеханический факультет», ФБГОУ ВО СГУВТ, Россия, г. Новосибирск

Katkovskaya K.V., senior lecturer of the Department "Information Systems", FGBOU VO SGUVT, Russia, Novosibirsk

K.S. Barts, Master's degree student, 1st year, Faculty of "Electromechanical Faculty", FBGOU IN SGUVT, Russia, Novosibirsk

Аннотация: В статье рассматривается классификация результатов соревнований по конному спорту в видах конкур и выездка на языке Python при помощи библиотеки Scikit-Learn. Предполагается использование метода классификации для обработки больших данных для дальнейшего прогнозирования и выбора более результативной лошади. Представлено описание метода k-ближайших соседей и алгоритма машинного обучения. Приведены графики распределения результатов соревнований по видам конкур и выездка. Показана важность классификации, как одного из

важнейших методов научного познания, осмысления и описания закономерностей окружающего мира, а также формой, так называемого контролируемого познания и одной из опор эмпирического знания, используется для упорядочения знаний.

Annotation: The article discusses the classification of the results of equestrian competitions in the types of show jumping and dressage in Python using the Scikit-Learn library. It is supposed to use the classification method for processing big data for further prediction and selection of a more efficient horse. A description of the k-nearest neighbors method and a machine learning algorithm is presented. The graphs of the distribution of the results of competitions by types of show jumping and dressage are given. The importance of classification is shown as one of the most important methods of scientific knowledge, understanding and describing the patterns of the surrounding world, as well as a form of so-called controlled cognition and one of the pillars of empirical knowledge, used to streamline knowledge.

Ключевые слова: метод k-ближайших соседей, классификация, машинное обучение, конный спорт, python, соревнования.

Keywords: k-nearest neighbors, classification, machine learning, equestrianism, python, competitions.

В настоящее время такой метод познания, как классификация, используется для достижения различных целей, как научно-теоретических, так и сугубо практических, например, классификация получаемых данных является важным этапом решения задач в любой деятельности человека [1]. В силу большого объема показателей и набора данных, необходимых для построения модели прогнозирования, целесообразно использовать информационные технологии, как механизм реализации вычислений и визуализации данных. В качестве таких информационных современных средств может выступать специальные среды разработки программ прогнозирования на языке программирования Python.

Метод классификации является одним из важнейших методов научного познания, осмысления и описания закономерностей окружающего мира, а также формой, так называемого контролируемого познания и одной из опор эмпирического знания, используется для упорядочения знаний. Без обобщения и упорядочения исследуемых объектов, их сортировки по определенным признакам невозможно обойтись ни в одной отрасли, ни в одном междисциплинарном исследовании [2].

Большая часть деятельности по классификации строится на представлении о том, что классифицируемые объекты существуют независимо от человека и классификатор открывает реально существующие объекты и описывает их виды и отдельные феномены такими, какие они есть на самом деле. Учение о принципах и практике классификации именуется таксономией.

Термин «классификация» является достаточно широко используемым, однако его конкретное наполнение зачастую может зависеть от контекста его применения. Продуктом классификация является логически взаимоувязанный сложный многокомпонентный конструкт.

Метод k-ближайших соседей (k-nearest neighbors algorithm, kNN) отличается своей простотой. Обучение алгоритма заключается в запоминании обучающей выборки [4]. В случае использования метода для классификации объект присваивается тому классу, который является наиболее распространенным среди k-соседей данного объекта, классы которых уже известны. В случае использования метода для регрессии, объекту присваивается среднее значение по k-ближайшим к нему объектам, значения которых уже известны.

Алгоритм KNN для классификации может быть описан следующими шагами:

- 1 Выбрать число k-соседей и метрику расстояния;
- 2 Найти k-ближайших соседей объекта, который хотим классифицировать;
- 3 Присвоить метку класса мажоритарным голосованием.

Основываясь на выбранной метрике расстояния, алгоритм KNN находит в тренировочном наборе данных k объектов, которые являются самыми близкими к классифицируемому объекту (самыми похожими на него). Метка класса классифицируемого объекта определяется мажоритарным голосованием среди его k -ближайших соседей.

Ранее в статье упоминалось, что одним из современных средств реализации алгоритмов прогнозирования является язык программирования Python. Данный язык имеет ряд удобных библиотек машинного обучения и научных расчетов: Pandas, NumPy, Scikit-Learn, matplotlib.pyplot, которые позволяют построить рабочие модели. Библиотека Pandas отвечает за обработку и анализ данных, NumPy – за выполнение операций линейной алгебры и численных преобразований, matplotlib.pyplot – отвечает за визуализацию данных, Scikit-Learn – является непосредственно библиотекой машинного обучения [5, 6, 7, 8].

Машинное обучение – это методики анализа данных, которые позволяют аналитической системе обучаться в ходе решения множества сходных задач. Машинное обучение базируется на идее о том, что аналитические системы могут учиться выявлять закономерности и принимать решения с минимальным участием человека.

Машинное обучение обычно относится к изменениям в системах, выполняющих задачи, связанных с искусственным интеллектом (ИИ). Такие задачи включают в себя распознавание, диагностику, планирование, управление роботом, прогнозирование и т. д. Изменения могут быть либо усовершенствования уже работающих систем или синтез новых систем [3].

Алгоритмы машинного обучения помогают ответить на вопросы, ответ на которые сложно найти с помощью анализа, выполняемого вручную. С помощью алгоритмов отображается либо прогнозируется связь между двумя переменными.

Конный спорт – сборное понятие, под которым понимаются спортивные игры с использованием лошадей. При езде верхом всадник активно

взаимодействует с движением лошади. Лошадь при этом управляется перемещением веса всадника, сжатием бёдер и ослаблением или натяжением поводьев. Голос всадника тоже может влиять на движение лошади, но не всеми видами конного спорта это разрешено [9].

Конный спорт очень разнообразен, но к классическим видам его относятся только три:

- Конкур;
- Выездка;
- Троеборье.

Для начала необходимо обработать данные соревнований по выездке и конкуру. При обработке данных были выявлены данные:

- Тип зачета (общий зачет, зачёт для юношей, зачет для детей, зачёт для любителей, зачет для детей на пони);
- Место, занятое лошастью и всадником;
- Кличка лошади и код;
- Владелец лошади;
- Тренер всадника;
- Название конного клуба;
- Набранные за схему баллы;
- Набранные за схему проценты.

Для классификации по выездке были отобраны данные: кличка лошади и код; место, занятое лошастью и всадником; набранные за схему проценты. Для упрощения обработки данных кличкам лошадей были выданы id-коды.

На основании данных был построен график (см. рис. 1).

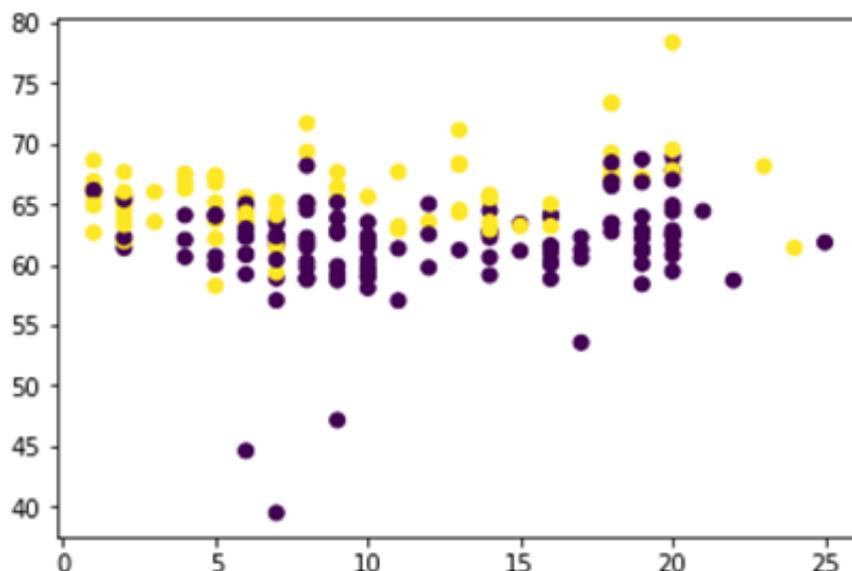


Рисунок 1. Распределение данных по выезде

По результатам обработанных данных можно пронаблюдать результативность лошадей за большой период времени по выезде. По оси ординат расположен процент выполнения норматива лошадью и всадником.

Для классификации по конкуру были отобраны данные: кличка лошади и код; место, занятое лошадью и всадником, набранные за схему проценты.

На основании данных был построен график (см. рис. 2).

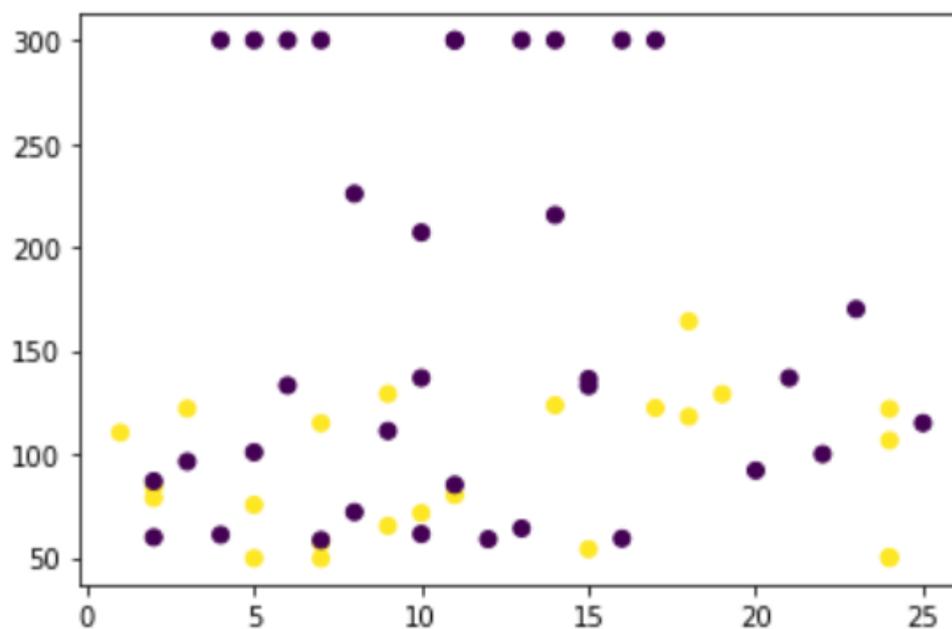


Рисунок 2. Распределение данных по конкуру

По результатам обработанных данных можно пронаблюдать результативность лошадей за большой период времени по конкуру. На графике по оси абсцисс расположены id-коды лошадей, что упрощает возможность отследить результативность лошади на соревнованиях по конкуру и выездке. По оси ординат расположено время прохождения схемы лошастью и всадником.

Благодаря данному методу можно быстро обрабатывать и классифицировать большие данные. На графиках наглядно видно соотношение побед и поражений, и неповиновений лошадей. Всадник и тренер, после ознакомления с результатами, может выбрать под всадника наиболее подходящую лошадь для участия в последующих соревнованиях по выездке и конкуру. Метод поможет отобрать наиболее способных лошадей под всадников для повышения шанса выигрыша на областных, всероссийских соревнованиях и повышения рейтинга конного клуба.

Литература:

1. О классификации спортивных соревнований в федеральных стандартах спортивной подготовки: Научная электронная библиотека «КиберЛенинка». [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/o-klassifikatsii-sportivnyh-sorevnovaniy-v-federalnyh-standartah-sportivnoy-podgotovki/viewer> (дата обращения: 05.06.2022).
2. Классификация как метод научного исследования, в частности в юридической науке: Научная электронная библиотека «КиберЛенинка». [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/klassifikatsiya-kak-metod-nauchnogo-issledovaniya-v-chastnosti-v-yuridicheskoy-nauke/viewer> (дата обращения: 05.06.2022).
3. Nils J. Nilsson. Introduction to machine learning: textbook / Nils J. Nilsson. – Robotics Laboratory Department of Computer Science Stanford University. – 2005.– 188 с.

4. Метод k-ближайших соседей: Википедия свободная энциклопедия [Электронный ресурс]. URL: https://ru.wikipedia.org/wiki/Метод_k-ближайших_соседей (дата обращения: 05.06.2022).
5. NumPy: Официальная документация [Электронный ресурс]. URL: <https://numpy.org/> (дата обращения: 02.06.2022).
6. Pandas: Официальная документация [Электронный ресурс]. URL: <https://pandas.pydata.org/> (дата обращения: 02.06.2022).
7. Scikit-Learn: Официальная документация [Электронный ресурс]. URL: <https://scikit-learn.org/stable/> (дата обращения: 02.06.2022).
8. Matplotlib: Официальная документация [Электронный ресурс]. URL: <https://matplotlib.org/stable/index.html> (дата обращения: 02.06.2022).
9. Виды конного спорта: Конный спорт [Электронный ресурс]. URL: vetkzn.ru/konnyi_sport/horsesports/ (дата обращения: 02.06.2022).

Literature:

1. On the classification of sports competitions in the federal standards of sports training: Scientific electronic library "CyberLeninka". [electronic resource]. URL: <https://cyberleninka.ru/article/n/o-klassifikatsii-sportivnyh-sorevnovaniy-v-federalnyh-standartah-sportivnoy-podgotovki/viewer> (accessed: 06/05/2022).
2. Classification as a method of scientific research, in particular in legal science: Scientific electronic library "CyberLeninka". [electronic resource]. URL: <https://cyberleninka.ru/article/n/klassifikatsiya-kak-metod-nauchnogo-issledovaniya-v-chastnosti-v-yuridicheskoy-nauke/viewer> (accessed: 06/05/2022).
3. Nils J. Nilsson. Introduction to machine learning: textbook / Nils J. Nilsson. – Robotics Laboratory Department of Computer Science Stanford University. – 2005. – 188 p.
4. The k-nearest neighbors method: Wikipedia free encyclopedia [Electronic resource]. URL: https://ru.wikipedia.org/wiki/Method_k-nearest_neighbors (accessed: 06/05/2022).

- 5 NumPy: Official documentation [Electronic resource]. URL: <https://numpy.org/> (accessed: 06/02/2022).
- 6 Pandas: Official Documentation [Electronic resource]. URL: <https://pandas.pydata.org/> (accessed: 06/02/2022).
- 7 Scikit-Learn: Official Documentation [Electronic resource]. URL: <https://scikit-learn.org/stable/> (accessed: 02.06.2022).
- 8 Matplotlib: Official documentation [Electronic resource]. URL: <https://matplotlib.org/stable/index.html> (accessed: 02.06.2022).
- 9 Types of equestrian sports: Equestrian sports [Electronic resource]. URL: vetkzn.ru/konnyi_sport/horsesports/ (accessed: 02.06.2022).

© Катковская К.В., Барц К.С. 2022 Научный сетевой журнал
«Столыпинский вестник» №6/2022

Для цитирования: Катковская К.В., Барц К.С. КЛАССИФИКАЦИЯ РЕЗУЛЬТАТОВ СОРЕВНОВАНИЙ ПО КОННОМУ СПОРТУ НА ОСНОВЕ МЕТОДА К-БЛИЖАЙШИХ СОСЕДЕЙ // Научный сетевой журнал «Столыпинский вестник» №6/2022