

Prediction the Number of Medals of Iranian Taekwondo Sport in the Olympics Through the Cultural and Social Components and Decision Tree Algorithm

ARTICLE INFO

Article Type

Research Article

Authors

Morteza Ghalibaf Tousi Nejad¹

Zinat Nikaeen ^{2*}

Faride Ashraf Ganjuee¹

Zahra Haji Anzahae¹

How to cite this article

Morteza Ghalibaf Tousi Nejad, Zinat Nikaeen, Faride Ashraf Ganjuee, Zahra Haji Anzahae, Prediction the Number of Medals of Iranian Taekwondo Sport in the Olympics Through the Cultural and Social Components and Decision Tree Algorithm, *Journal of Islamic Life Style Centeredon Health*, 2021:5(1): 405-412.

1. Department of Sports Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran.

2. Department of Sports Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran (Corresponding Author).

* Correspondence:

Address:

Phone:

Email: dr.zinatnikaeen@gmail.com

Article History

Received: 2021/02/23

Accepted: 2021/05/10

ABSTRACT

Purpose: Today, winning a medal in the Olympic event is very important for all countries. The purpose of the present study is to predict the number of medals in Iran's taekwondo sport in the Paris Olympics using the decision tree algorithm. Sampling of this research was done for Olympic experts using the snowball method.

Materials and methods: In this research, in order to determine the indicators that are effective on medal winning in the Olympics, preliminary studies and interviews to the point of saturation with 11 experts in the Olympic field were used. In the following, 78 indicators effective on winning medals in the Olympics were selected from the relief algorithm. These indicators were classified into five components: economic, political, cultural and social, technological and sports, and 54 indicators were extracted qualitatively and 24 indicators quantitatively. Data were collected for 24 quantitative indicators from library studies and sites and 54 qualitative indicators from Taekwondo Olympians of different Olympic eras. The statistical population includes all the Olympians as well as the number of medals won by Iran in Taekwondo for the Olympic periods. The statistical sample was 30 Olympians from the 2000 Sydney Olympics to the 2020 Tokyo Olympics. Data analysis was done using the decision tree data mining method.

Findings: The results of data analysis showed that the decision tree algorithm has 80.12% accuracy for predicting Olympic results.

Conclusion: As a result, by using the decision tree algorithm, it is predicted that Iran will win a bronze medal in Taekwondo at the Paris Olympics.

Keywords: Prediction, Olympics, Cultural and Social Component, Data Analysis, Taekwondo

پیش بینی تعداد مدال رشته ورزشی تکواندوی ایران در المپیک با استفاده از مولفه های فرهنگی و

اجتماعی و درخت تصمیم گیری

نتیجه گیری: در نتیجه با استفاده از الگوریتم درخت تصمیم پیش بینی می شود کشور ایران در المپیک پاریس در رشته ورزشی تکواندو موفق به کسب یک مدال برنز خواهد شد.

واژه های کلیدی: پیش بینی، المپیک، مولفه فرهنگی و اجتماعی، داده کاوی، تکواندو

تاریخ دریافت: ۱۳۹۹/۱۲/۰۵

تاریخ پذیرش: ۱۴۰۰/۰۲/۲۰

* نویسنده مسئول: dr.zinatnikaen@gmail.com

مرتضی قالیباف طوسی نژاد^۱

گروه مدیریت ورزشی، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران.

زینت نیک آیین^{۲*}

گروه مدیریت ورزشی، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران (نویسنده مسئول).

فریده اشرف گنجوی^۳

گروه مدیریت ورزشی، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران.

زهرا حاجی انزهایی^۴

گروه مدیریت ورزشی، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران.

چکیده

هدف: امروزه در رویداد المپیک کسب مدال برای تمامی کشورها از اهمیت زیادی برخوردار است. هدف از مطالعه حاضر پیش بینی تعداد مدال رشته ورزشی تکواندو ایران در المپیک پاریس با استفاده از الگوریتم درخت تصمیم است. نمونه گیری این پژوهش برای متخصصین حوزه المپیک با استفاده از روش گلوله برفی صورت گرفت.

مواد و روش ها: در این پژوهش به منظور تعیین شاخص های موثر بر مدال آوری در المپیک از مطالعات اولیه و مصاحبه تا حد اشباع با ۱۱ نفر از خبرگان حوزه المپیک استفاده شد. در ادامه از الگوریتم رلیف ۷۸ شاخص مؤثر بر مدال آوری در المپیک انتخاب شد. این شاخص ها در پنج مؤلفه اقتصادی، سیاسی، فرهنگی و اجتماعی، فناوری و ورزشی طبقه بندی شد و ۵۴ مورد از شاخص ها بصورت کیفی و ۲۴ شاخص دیگر بصورت کمی استخراج گردید. برای ۲۴ شاخص کمی از مطالعات کتابخانه ای و سایت ها و ۵۴ شاخص کیفی از المپین های تکواندوی ادوار مختلف المپیک داده ها جمع آوری گردید. جامعه آماری شامل همه المپین ها و همچنین تعداد مدال کسب شده ایران در رشته تکواندو برای ادوار المپیک می باشد. نمونه آماری تعداد ۳۰ نفر المپین از المپیک ۲۰۰۰ سیدنی تا المپیک ۲۰۲۰ توکیو بود. تجزیه و تحلیل داده ها با استفاده از روش داده کاوی درخت تصمیم صورت گرفت.

یافته ها: نتایج تجزیه و تحلیل داده ها نشان داد که الگوریتم درخت تصمیم از دقت ۸۰/۱۲٪ برای پیش بینی نتایج المپیک برخوردار است.

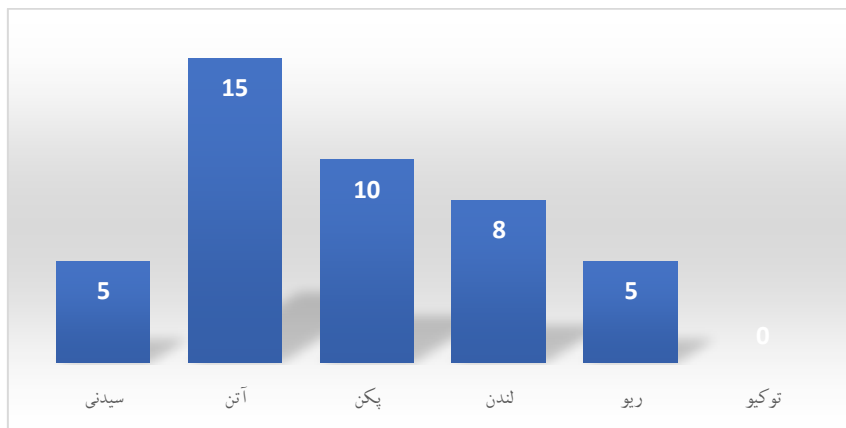
مقدمه

پیش بینی دقیق عملکرد المپیک در دهه های گذشته مورد توجه محققین قرار گرفته است، عمدتاً به این دلیل که پیش بینی مدال، برای ارائه یک دولت به شهروندان ضروری است (۱)، اما در مسیر پیش بینی همواره مشکلاتی وجود داشته که انجام پیش بینی های دقیق و قابل اعتماد را تقریباً غیر ممکن کرده است. پیش بینی مدال های المپیک جریان فکری جدیدی نیست، این نوع تحقیقات از اوایل دهه ۱۹۷۰ شروع شده است (۲). ورزش سیستمی است که در آن چندین عامل با هم کار می کنند، وجود پارامترهای زیاد و در بسیاری موارد پنهان، پیش بینی را به مساله ای سخت تبدیل نموده است که الگوریتم های پیچیده ریاضی را نیز از ارائه راهکارهای مناسب برای ساخت یک الگوی پیش بینی کارآمد عاجز کرده است (۳). مشاهده تکرار در وقوع رویدادها این پیش فرض را به وجود آورد که می توان نتایج و تبعات آن را به کلیه رویدادهایی از آن جنس تعمیم داد. بنابراین استفاده از فناوری یادگیری ماشین به محققان اجازه می دهد تا الگوریتم هایی را طراحی و سیستم هایی را شبیه سازی کنند که بتواند با قدرت بالاتری نتایج را پیش بینی کند (۴). در این میان تعداد معیارها، پیچیدگی داده ها و پویایی محیط از جمله عواملی است که مساله تصمیم گیری در ورزش را با چالشی جدی در دهه اخیر مواجه کرده است که هر چه این عوامل بیشتر باشند یادگیری ماشین دشوارتر خواهد شد (۵). مقالات متعددی نیز بیان کردند که پیش بینی های ورزشی که مبتنی بر داده و اطلاعات است با آنچه که به صورت تصادفی مانند پیش بینی های شانس انجام می پذیرد، کاملاً متفاوت است (۶).

افراد خبره با قضاوت های ذهنی، ممکن است در شرایط کمتر معلوم و بیشتر نامطمئن به علت استفاده از معیار و ملاک های کیفی در مقایسه با مدل های آماری پیش بینی بهتری را انجام دهند (۷). بررسی های زیادی برای مقایسه این دو روش در زمینه های مختلف مانند پزشکی، موفقیت های دانشگاهی، تصمیم گیری در تجارت، پیش بینی وضعیت آب و هوا، پیش بینی های اقتصاد کلان، نرخ تورم، نتایج انتخابات سیاسی و ... صورت پذیرفته است که در اغلب آن ها الگوهای آماری پیش بینی بهتری کرده اند (۸). تحقیقات متفاوتی صورت گرفته است که در پیش بینی از روش های داده کاوی استفاده کرده اند و نتیجه آن پیش بینی ها در مقایسه با الگوهای آماری بهتر

به غیرخطی بودن مقوله‌های مورد بررسی در این پژوهش می‌بایست از روش‌های غیرخطی استفاده نمود. بسیاری از محققان از الگوریتم‌های داده‌کاوی مختلف برای بدست آوردن بالاترین دقت پیش‌بینی نتایج استفاده کردند (۱۵). لذا در این پژوهش روش داده‌کاوی درخت تصمیم استفاده گردید. کشور ایران تا اکنون تنها در ۶ رشته ورزشی موفق به کسب مدال المپیک شده است که این رشته‌ها عبارتند از کشتی، وزنه‌برداری، تکواندو، دو و میدانی، کاراته و تیراندازی. در میان این رشته‌های ورزشی روند نزولی در دریافت مدال را برای رشته ورزشی تکواندو شاهد هستیم بر اساس تعداد مدال‌های کسب شده در ادوار المپیک در صورتی که برای مدال طلا ده امتیاز، نقره هشت امتیاز و برنز پنج امتیاز در نظر بگیریم می‌توان این روند را در نمودار ۱ مشاهده کرد. به همین دلیل رشته ورزشی مورد هدف در این پژوهش تکواندو انتخاب گردید تا براساس میزان دقت خروجی پیش‌بینی، بتوان پیش‌بینی کننده‌ی مناسب را جهت اخذ تصمیمات برطبق شاخص‌های مؤثر بر مدال‌آوری رشته ورزشی تکواندو در المپیک با حداکثر میزان کارایی ارائه نمود و روند پیش رو را در مسیر درست قرار داد تا مانع سیر نزولی کسب مدال در این رشته ورزشی گردد.

بوده است. هنگ لی^۱ آنالیز مدل پیش‌بینی مسابقات ورزشی با استفاده از روش داده‌کاوی شبکه عصبی و رگرسیون را انجام داد که مدل پیش‌بینی ارائه شده توسط الگوریتم شبکه عصبی دارای خطای پیش‌بینی کمتری نسبت به مدل رگرسیون خطی داشت (۹)، گراهام^۲ از روش داده‌کاوی و رگرسیون برای پیش‌بینی پذیرش و میزان ازدهام در بخش اورژانس بیمارستان استفاده کرد که روش داده‌کاوی درخت تصمیم از دقت بالاتری برخوردار بود (۱۰). پژوهش پیش‌بینی نتایج مسابقات جهانی کریکت با استفاده از روش‌های داده‌کاوی توسط جاسکانی^۳ صورت گرفت که نتایج روش جنگل تصادفی بهتر بود نسبت به رگرسیون (۱۱) همچنین مزیدی در پژوهشی با عنوان پیش‌بینی نتایج مسابقات ورزشی با استفاده از دسته‌بندی کننده‌ها بدست آورد که پیش پردازش داده‌ها نتایج دقیق‌تری برای پیش‌بینی نسبت به بدون پیش پردازش دارد (۱۲) و مقایسه روش‌های رگرسیون، جنگل تصادفی و نایویز برای شیردهی و بقای گاوهای هولشتاین توسط ون دیر هیده^۴ صورت گرفت که نتایج نشان داد جنگل تصادفی و نایویز به نسبت قدرت پیش‌بینی بیشتری نسبت به رگرسیون را دارا هستند (۱۳). آگاروال^۵ در پژوهشی به پیش‌بینی مسابقات هنرهای رزمی پرداخت که در نتیجه روش‌های یادگیری ماشین توانست پیش‌بینی با دقت بالایی را داشته باشد (۱۴). با توجه



نمودار (۱): مدال‌های کشور ایران در المپیک برای رشته ورزشی تکواندو

مراحل پژوهش حاضر شامل (الف) جمع‌آوری داده‌ها، (ب) انتخاب ویژگی‌ها، (ج) اجرای الگوریتم پیشنهادی، (د) اندازه‌گیری دقت عملکرد الگوریتم پیش‌بینی و (ه) پیش‌بینی تعداد مدال المپیک برای رشته ورزشی تکواندو است.

مواد و روش‌ها

در این پژوهش از روش داده‌کاوی درخت تصمیم برای پیش‌بینی تعداد مدال رشته ورزشی تکواندو ایران در المپیک ۲۰۲۴ پاریس استفاده شد. شکل (۱) مراحل اجرای روش پژوهش را نشان می‌دهد.

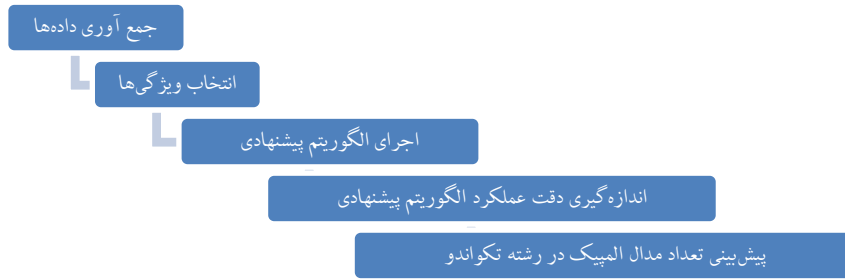
4- Van der Heide

5- Aggarwal

1- Hang Li

2- Graham

3- Jaskani



شکل ۱. مراحل اجرای روش پژوهش

خطی هستند و استخراج داده‌ها که بصورت چندین لایه بودند و سپس انتقال داده‌ها به یک دسته‌بندی کننده^۲ صورت گرفت. که دسته‌بندی کننده تمامی این شاخص‌ها را با هم ترکیب می‌کند تا بتواند پیش‌بینی را انجام دهد. ویژگی‌های پیچیده را نمی‌توان بصورت مستقیم از داده‌های ورودی بدست آورد. می‌توان از لحاظ ریاضی نشان داد که بهترین ویژگی‌هایی که می‌شود از یک داده با استفاده از یک لایه (بدون سلسله مراتب) بدست آورد فقط لایه‌ها^۳ و توده-ها^۴ باشد. دلیل این امر هم را می‌توان داده‌های محدود از یک تبدیل غیرخطی ورودی دانست. برای بدست آوردن و با تولید ویژگی‌هایی که شامل اطلاعات بیشتری هستند نمی‌توانستیم بصورت مستقیم بر روی داده‌های ورودی کار کنیم و لازمه این کار این است که شاخص‌های اولیه ورودی (مثل لایه‌ها و توده‌ها) مجدد تبدیل شود تا شاخص‌های پیچیده‌تری که شامل داده‌های بیشتری برای تمایز بین کلاس‌ها مورد نیاز هست بدست بیاید. ۷۸ شاخص جمع‌آوری شده بصورت اعداد کمی به عنوان "داده‌های ورودی" به سیستم داده شد. تعداد مدال کسب شده در هر المپیک نیز به عنوان متغیر ملاک به سیستم وارد شد که برای هر دوره، الگوریتم اجرا گردید و سپس با تعداد مدال واقعی کسب شده آموزش صورت گرفت. این مسئله علاوه بر خیره‌سازی الگوریتم درخت تصمیم به کاهش خطای پیش-بینی موجود کمک کرد.

الگوریتم پیش‌بینی درخت تصمیم شاخه‌ای از بحث یادگیری ماشین و هوش مصنوعی و مجموعه‌ای از الگوریتم‌هایی است، که تلاش می‌کنند، مفاهیم انتزاعی سطح بالا را با استفاده از یادگیری در سطوح و لایه‌های مختلف مدل کنند. در ساختار درخت تصمیم، پیش‌بینی به دست آمده از درخت در قالب یک سری قواعد توضیح داده می‌شود. هر مسیر از ریشه تا یک برگ درخت تصمیم، یک قانون را بیان می‌کند و در نهایت برگ با کلاسی که بیشترین مقدار رکورد در آن تعلق گرفته برچسب می‌خورد. این الگوریتم روشی افزایشی از هرس کردن درخت را بکار می‌گیرد تا خطای طبقه‌بندی ناشی از نویز با جزئیات خیلی زیاد را در داده‌های آموزشی کاهش دهد. هرس کردن با جایگزینی گره داخلی با گره برگ رخ می‌دهد که بدان وسیله درصد یا میزان خطا کاهش می‌یابد.

$$E(S) = \sum_{i=1}^c -p_i \log_2 p_i \quad (۱) \text{ فرمول}$$

$E(S)$: آنتروپی مجموعه‌ای از داده‌ها

الف) جمع‌آوری داده‌ها: در پژوهش حاضر تعیین شاخص‌هایی موثر بر مدال آوری در المپیک از طریق مصاحبه با خبرگان حوزه المپیک صورت گرفت که شاخص‌های جمع‌آوری شده به دو صورت کمی و کیفی بودند. در بخش کمی (جمع‌آوری اطلاعات از طریق اینترنت و منابع معتبر صورت گرفت) و کیفی (سؤال‌های کیفی از المپین‌ها جهت تعیین پاسخ به شاخص‌های تعیین شده) انجام شد. در بخش اول از طریق مصاحبه با ۱۱ خبره در حوزه المپیک تا حد اشباع در پنج مؤلفه‌ی اقتصادی، سیاسی، فرهنگی و اجتماعی، فناوری و ورزشی با استفاده از روش گلوله برفی صورت گرفت. جامعه پژوهش در این بخش شامل خبرگان حوزه المپیک می‌باشند. در بخش المپین‌ها جامعه پژوهش شامل تمامی تکواندوکاران حاضر در المپیک هستند که از میان آنها تعداد ۳۰ نفر بطور تصادفی انتخاب شدند.

ب) انتخاب ویژگی‌ها: پس از مصاحبه با خبرگان حوزه المپیک از میان شاخص‌های مختلف به منظور کاهش بار محاسباتی اضافی و تعیین شاخص‌های موثر بر مدال آوری در المپیک از الگوریتم رلیف^۱ که یک روش مبتنی بر وزن و الهام گیر از نمونه می‌باشد، استفاده شد. الگوریتم رلیف از بین داده‌های آموزشی یک نمونه را به صورت تصادفی انتخاب می‌کند و سپس فاصله اقلیدسی آن نمونه تا نزدیک‌ترین نمونه از طبقه مشابه و نزدیک‌ترین نمونه از طبقه متفاوت را به دست می‌آورد و سپس این فاصله‌ها را برای به‌روز کردن وزن هر متغیر به کار می‌برد. در نهایت الگوریتم آن دسته از متغیرهایی را انتخاب می‌کند که وزن آن‌ها از یک حد آستانه‌ی از پیش تعریف شده به‌وسیله کاربر، بیشتر است. از جمله نکات قوت این الگوریتم می‌توان به ساده بودن اصول و عدم پیچیدگی آن، قابل حل بودن با توابع چند جمله‌ای مرتبه پایین و نیاز به تعداد کم داده‌های آموزشی اشاره کرد.

ج) اجرای الگوریتم پیشنهادی: تکواندو از المپیک ۲۰۰۰ سیدنی به عنوان یکی از رشته‌های ورزشی المپیک شناخته شد به همین دلیل بازه زمانی المپیک ۲۰۰۰ تا ۲۰۲۰ جهت گردآوری داده‌ها تعیین شد. ۵۴ شاخص کیفی از ۳۰ المپین تکواندو به صورت مقیاس ۵ ارزشی لیکرت سوال شد و از ۱ تا ۵ تبدیل به مقیاس کمی گردید. همچنین ۲۴ شاخص کمی از طریق کتابخانه و سایت‌های معتبر (بانک جهانی، سازمان بهداشت جهانی و مرکز اطلاعات و آمار ایران) جمع‌آوری گردید. پس از تعیین شاخص‌ها که بصورت غیر

³ - Edge

⁴ - Blob

¹ - Relief

² - Classifier

C: تعداد کلاس‌ها در یک سیستم

P_i: تعداد نمونه‌های موجود در کلاس i

د) اندازه‌گیری دقت عملکرد الگوریتم پیش‌بینی: نحوه کاری که در درخت تصمیم مشاهده می‌شود به این صورت است که یک درخت تصمیم یک ساختار نموداری درخت مانند است که در آن هر گره داخلی یک ویژگی را نشان می‌دهد، هر شاخه یک نتیجه را نشان می‌دهد، برچسب کلاس توسط هر گره برگ (گره پایانی) نشان داده می‌شود. بنابراین، درخت تصمیم سه مولفه اصلی دارد: گره ریشه: گره اصلی در بالای درخت که گره ریشه نامیده می‌شود. گره برگ: شامل سوالات و معیارهایی است که باید به آن‌ها پاسخ داده شود.

شاخه‌ها: همان ساختارهای فلش ماندی هستند که گره‌ها را به یکدیگر وصل می‌کنند.

ه) پیش‌بینی تعداد مدال المپیک برای رشته ورزشی تکواندو: در پایان با استفاده از الگوریتم پیش‌بینی کننده درخت تصمیم تعداد مدال ایران در المپیک پاریس در رشته ورزشی تکواندو به عنوان داده‌های خروجی مشخص گردید.

یافته‌ها

برای دریافت خروجی از الگوریتم پیشنهادی اطلاعات ۷۸ شاخص تعیین شده و متغیر ملاک که همان تعداد مدال‌های طلا، نقره و برنز می‌باشد را به سیستم اعمال کرده و برای هر کدام از المپیک‌ها طی بازه ۲۰۰۰ تا ۲۰۲۰ پیاده سازی شد.

جدول ۱. تعداد مدال ایران کسب شده کشور ایران در المپیک به تفکیک رنگ مدال

۲۰۰۰ سیدنی	۲۰۰۴ آتن	۲۰۰۸ پکن	۲۰۱۲ لندن	۲۰۱۶ ریو	۲۰۲۰ توکیو
۰	۱	۱	۰	۰	۰
۰	۰	۰	۱	۰	۰
۱	۱	۰	۰	۱	۰

از بین ۱۲۰ شاخص، ۷۸ شاخص مؤثر بر مدال‌آوری در المپیک با استفاده از الگوریتم رلیف تعیین گردید، که از میان آنها ۲۴ شاخص کمی و ۵۴ شاخص کیفی تعیین شد. مؤلفه اقتصادی ۱۴ شاخص، سیاسی ۱۲ شاخص، فرهنگی اجتماعی ۱۲ شاخص، فناوری ۱۰ شاخص و ورزشی ۳۰ شاخص را به خود اختصاص دادند.

براساس داده‌های جدول (۱) از میان ۶ مدال المپیک کسب شده توسط تیم ملی ایران در رشته ورزشی تکواندو از سال ۲۰۰۰ سیدنی تا ۲۰۲۰ توکیو (رشته ورزشی تکواندو از سال ۲۰۰۰ سیدنی جزء رشته‌های المپیکی قرار گرفته است)، بهترین نتیجه، در المپیک آتن با کسب ۱ مدال طلا و ۱ برنز و ضعیف‌ترین نتیجه در المپیک توکیو بدون کسب مدال ثبت شده است.

جدول ۳. ۷۸ تعداد و میزان درصد متغیرهای پنج شاخص فناوری، ورزشی، اقتصادی، فرهنگی اجتماعی و سیاسی مؤثر بر مدال‌آوری تکواندو ایران در المپیک

جمع شاخص‌ها	فناوری	ورزشی	اقتصادی	فرهنگی اجتماعی	سیاسی
۷۸	۱۰	۳۰	۱۴	۱۲	۱۲
٪۱۰۰	٪۱۳	٪۳۹	٪۱۸	٪۱۵	٪۱۵

عنوان ورودی وارد سیستم گردید و در خروجی تعداد مدال رشته ورزشی تکواندو در المپیک پاریس مورد پیش‌بینی قرار گرفت. خروجی سیستم میانگین دقت پیش‌بینی برای روش درخت تصمیم در ادواری که تکواندو در المپیک حضور داشته (۲۰۰۰ تا ۲۰۲۰) بر اساس شاخص‌های ورودی با استفاده از تابع $Accuracy = \frac{TP+TN}{TP+TN+FP+FN+TP}$ محاسبه گردید. دقت پیش‌بینی، نسبت تعداد پیش‌بینی صحیح در الگوریتم یادگیری ماشین به تعداد کل ورودی نمونه‌ها است. جدول (۴) ماتریس درهم‌ریختگی را نشان می‌دهد که اغلب برای تعریف خروجی

بر طبق داده‌های جدول (۳) از میان ۷۸ متغیر استخراج شده در ۵ شاخص فناوری، ورزشی، اقتصادی، فرهنگی اجتماعی و سیاسی، شاخص ورزشی با ۳۰ متغیر و ۳۹٪ درصد بیشترین اثر را در کسب مدال المپیک برای رشته ورزشی تکواندو دارد و فناوری با ۱۰ متغیر و ۱۳٪ کمترین اثر برای کسب مدال در المپیک را به خود اختصاص داده است.

الگوریتم درخت تصمیم طراحی شده برای هر دوره از المپیک‌ها پیاده‌سازی شد و از طریق متغیر ملاک آموزش دید و با داده‌های المپیک پیشین خود جمع گردید، این مسئله علاوه بر خبره‌سازی سیستم، اشتباهات الگوریتم را نمایان می‌ساخت. در پایان داده‌ها به

الگوریتم بروی مجموعه‌ای از داده‌های آزمایشی استفاده می‌شود (۱۶).

جدول ۴. ماتریس درهم‌ریختگی

مقدار واقعی	مقدار پیش‌بینی شده	
	مثبت	منفی
	TP	FP
	FN	TN


الگوریتم درخت تصمیم با میانگین خطای ۱۹/۸۸ در ۵ شاخص ورودی دارای دقت ۸۰/۱۲٪ برای پیش‌بینی تعداد مدال تکواندو ایران در المپیک پاریس اجرا گردید.

جدول ۵. میانگین دقت و خطای هر شاخص در روش داده کاوی نایو بیز

میانگین دقت	میزان دقت	میانگین خطا	میزان خطای هر شاخص	شاخص ارزیابی شده	روش داده کاوی
۸۰/۱۲	۷۹/۴۷	۱۹/۸۸	۲۰/۵۳	فرهنگی اجتماعی	درخت تصمیم
	۷۹/۴۰		۲۰/۶۰	سیاسی	
	۸۰/۸۵		۱۹/۱۵	اقتصادی	
	۸۰/۵۴		۱۹/۴۶	ورزشی	
	۸۰/۳۵		۱۹/۶۵	فناوری	

در خروجی الگوریتم درخت تصمیم که بصورت دهدهی بدست آمد پیش‌بینی شد کشور ایران در المپیک پاریس ۱ مدال برنز بدست خواهد آورد.

جدول ۳. تبدیل خروجی دودویی به دهدهی تعداد مدال تکواندو ایران در المپیک ۲۰۲۰ پاریس

تبدیل به خروجی دهدهی	در نتیجه	تعداد مدال	رنگ مدال
$2^3 \times 0 + 2^2 \times 0 + 2^1 \times 0 + 2^0 \times 1$		۰	برنز
$2^3 \times 0 + 2^2 \times 0 + 2^1 \times 0 + 2^0 \times 0$		۰	نقره
$2^3 \times 0 + 2^2 \times 0 + 2^1 \times 0 + 2^0 \times 0$		۱	طلا

نتیجه‌گیری

یکی از کاربردهای حیاتی پیش‌بینی در ورزش که نیاز به دقت زیادی دارد، پیش‌بینی تعداد مدال در المپیک است. با پیشرفت فناوری در دهه‌های گذشته، یادگیری ماشین به عنوان یک روش هوشمند نو ظهور نتایج مطلوبی در حوزه‌های پیش‌بینی و طبقه‌بندی ارائه می‌دهد. در پژوهش حاضر ۷۸ شاخص مدال‌آوری در المپیک پس از مصاحبه با خبرگان حوزه المپیک از طریق الگوریتم رلیف انتخاب شد. با استفاده از داده‌های جمع‌آوری شده طی سال‌های ۲۰۰۰ تا ۲۰۲۰ تعداد مدال کسب شده به عنوان متغیر ملاک به الگوریتم درخت تصمیم جهت پیش‌بینی تعداد مدال رشته ورزشی تکواندو در المپیک ۲۰۲۴ پاریس آموزش داده شد. برای هر دوره از المپیک الگوریتم اجرا و با توجه به نتایج بدست آمده واقعی خبره‌سازی در پیش‌بینی صورت گرفت. دقت پیش‌بینی ۸۰/۱۲٪ بطور میانگین برای شاخص‌های موثر بر مدال‌آوری تکواندوی ایران در المپیک پاریس با روش درخت تصمیم بدست آمد. نتایج پژوهش حاضر با تحقیق‌های آگاروال که پیش‌بینی مسابقات رزمی را مورد مطالعه

قرار داده، دقت ۳۵٪ را در روش نایو بیز نشان داد و شلمبیک (۱) در پژوهش پیش‌بینی تعداد مدال آمریکا در المپیک با دقت ۶۰٪ و همچنین پژوهش دوکی (۱۶) با موضوع طبقه بندی ضربات تکواندو با استفاده از یادگیری ماشین که دقت ۷۳٪ را نشان داد، همسو بوده و در این پژوهش‌ها نیز روش داده کاوی از دقت پیش‌بینی بالایی برخوردار نمی‌باشد، در مقابل با پژوهش‌های شارما (۳) پیش‌بینی مسابقات بدمینتون از طریق الگوریتم نایو بیز با دقت ۹۳٪ و پژوهش گراهام (۱۰) با موضوع استفاده از داده کاوی برای پیش‌بینی میزان پذیرش در بیمارستان که دقت ۸۰٪ را نشان داد، نا همسو بوده و داده کاوی و یادگیری ماشین در این پژوهش‌ها از دقت بالایی برخوردار است.

در پژوهش صورت گرفته بالاترین دقت پیش‌بینی مربوط به شاخص اقتصادی با درصد ۸۰/۸۵ و کمترین آن مربوط به شاخص سیاسی با درصد ۷۹/۴۰ می‌باشد. براساس خروجی الگوریتم پیاده‌سازی شده تعداد ۱ مدال برنز برای المپیک ۲۰۲۴ پاریس پیش‌بینی شد. نتیجه حاصل در مقایسه با تحقیقات مشابه از دقت پایینی برخوردار است.

tournament classification using machine learning technique based on decision tree, Naïve Bayes, and random forest algorithm. IOP Conference Series: Materials Science and Engineering,

9. Li, H. (2020). Analysis on the construction of sports match prediction model using neural network. *Soft Computing*, 24(11), 8343-8353. <https://doi.org/10.1007/s00500-020-04823-w>

10. Graham, B., Bond, R., Quinn, M., & Mulvenna, M. (2018). Using Data Mining to Predict Hospital Admissions From the Emergency Department. *IEEE Access*, 6, 10458-10469.

<https://doi.org/10.1109/ACCESS.2018.2808843>

11. Basit, A., Alvi, M. B., Jaskani, F. H., Alvi, M., Memon, K. H., & Shah, R. A. (2020, 5-7 Nov. 2020). ICC T20 Cricket World Cup 2020 Winner Prediction Using Machine Learning Techniques. 2020 IEEE 23rd International Multitopic Conference (INMIC),

12. Mazidi, A., Golsorkhtabar, M., & Etminan, N. (2020). Sport Result Prediction Using Classification Methods. 3, 39-48.

13. van der Heide, E. M. M., Veerkamp, R. F., van Pelt, M. L., Kamphuis, C., Athanasiadis, I., & Ducro, B. J. (2019). Comparing regression, naive Bayes, and random forest methods in the prediction of individual survival to second lactation in Holstein cattle. *Journal of Dairy Science*, 102(10), 9409-9421.

<https://doi.org/https://doi.org/10.3168/jds.2019-16295>

14. Aggarwal, K., Yadav, N., & Dwivedy, M. (2019). A Comparative Study of Machine Learning Algorithms for Prior Prediction of UFC Fights. In *Harmony Search and Nature Inspired Optimization Algorithms* (pp. 67-76). Springer.

15. Fazlollahi, P., Afarineshkaki, A., & Nikbakhsh, R. (2020). Predicting the

از دلایل بالا نبودن قدرت پیش بینی درخت تصمیم در پژوهش حاضر می توان کم بودن داده های ورودی به دلیل تعداد دوره های محدود حضور تکواندو در المپیک نام برد. برای پژوهش های آتی پیشنهاد می شود از دیگر روش های یادگیری ماشین استفاده شود و همچنین برای افزایش اطلاعات ورودی می توان نتایج مسابقات جهانی را نیز لحاظ کرد.

References

1. Schlembach, C., Schmidt, S. L., Schreyer, D., & Wunderlich, L. (2020). Forecasting the Olympic medal distribution during a pandemic: a socio-economic machine learning model. *arXiv preprint arXiv:2012.04378*.

2. Scelles, N., Andreff, W., Bonnal, L., Andreff, M., & Favard, P. (2020). Forecasting national medal totals at the Summer Olympic Games reconsidered. *Social Science Quarterly*, 101(2), 697-711.

3. Sharma, M., Kumar, N., & Kumar, P. (2021). Naive bayes-correlation based feature weighting technique for sports match result prediction. *Evolutionary Intelligence*, 1-16.

4. Lotfi, S. (2021). Machine Learning for sport results prediction using algorithms. *International Journal of Information Technology and Applied Sciences (IJITAS)*, 3(3), 148-155.

5. Osawa, K., Swaroop, S., Khan, M. E. E., Jain, A., Eschenhagen, R., Turner, R. E., & Yokota, R. (2019). Practical deep learning with Bayesian principles. *Advances in neural information processing systems*, 32.

6. Carloni, L., Angelis, A. D., Sansonetti, G., & Micarelli, A. (2021). A Machine Learning Approach to Football Match Result Prediction. *International Conference on Human-Computer Interaction*,

7. Fernández, J., Bornn, L., & Cervone, D. (2019). Decomposing the immeasurable sport: A deep learning expected possession value framework for soccer. *13th MIT Sloan Sports Analytics Conference*,

8. Dikananda, A. R., Ali, I., & Rinaldi, R. A. (2021). Genre e-sport gaming

Medals of the Countries Participating in the Tokyo 2020 Olympic Games Using the Test of Networks of Multilayer Perceptron (MLP). *Annals of Applied Sport Science*, 8(4), 0-0.

16. Duki, M., Shapiee, M., Abdullah, M. A., Khairuddin, I., Razman, A., & P P Abdul Majeed, A. (2021). The Classification of Taekwondo Kicks Via Machine Learning: A feature selection investigation. *MEKATRONIKA*, 3, 61-67. <https://doi.org/10.15282/mekatronika.v3i1.7153>.