نشریه علمی اختصاصی رایانه ایران زنجان
ISSN 2717-0233
RAYANIRAN
وزارت فرهنگ و ارشاد 7656
rayaniran.itcz.ir

هوش مصنوعی و تحلیل بقا برای پیش بینی طول عمر حرفه ای بازیکنان بسکتبال NBA با استفاده از داده های پزشکی و عملکردی

محمد حسین دانش پژوه,امیر حسین فقیه,حسین پرورش بوشهری,ابوالفضل حسن پور

محمد حسین دانش پژوه

کارشناسی مهندسی کامپیوتر دانشگاه خلیج فارس بوشهر دانشکده مهندسی جم

mohammadhdp@mehr.pgu.ac.ir

امیر حسین فقیه

کارشناسی مهندسی کامپیوتر دانشگاه خلیج فارس بوشهر دانشکده مهندسی جم

ah.faghih@mehr.pgu.ac.ir

ابوالفضل حسن پور

کارشناسی مهندسی کامپیوتر دانشگاه خلیج فارس بوشهر دانشکده مهندسی جم

a.hasanpour@mehr.pgu.ac.ir

حسین پرورش بوشهری

کارشناسی مهندسی کامپیوتر دانشگاه خلیج فارس بوشهر دانشکده مهندسی جم

h.parvaresh@mehr.pgu.ac.ir

چکیده

این مقاله به بررسی جامع ادبیات علمی موجود در زمینه به‌کارگیری هوش مصنوعی (AI) و تحلیل بقا (SA) برای پیش‌بینی طول عمر حرفه‌ای بازیکنان اتحادیه ملی بسکتبال (NBA) با ادغام داده‌های پزشکی و عملکردی می‌پردازد. در این پژوهش، روش‌شناسی‌های کلیدی، عوامل تأثیرگذار و چشم‌انداز رو به رشد تحلیل داده‌ها در ورزش مورد بررسی قرار گرفته است. یافته‌های اصلی نشان می‌دهند که ترکیبی از معیارهای عملکردی در زمین بازی، سوابق جامع آسیب‌دیدگی، و حتی عوامل بیوگرافیک، به طور قابل توجهی بر طول عمر بازیکنان تأثیر می‌گذارند. در حالی که مدل‌های هوش مصنوعی/یادگیری ماشین (AI/ML) مانند جنگل تصادفی (Random Forest) و تقویت گرادیان افراطی (XGBoost) دقت پیش‌بینی بالایی را از خود نشان می‌دهند، چالش‌هایی نظیر کیفیت داده‌ها، قابلیت تفسیر مدل و ملاحظات اخلاقی همچنان مطرح هستند. این مقاله بر پتانسیل تحول‌آفرین این رویکردهای تحلیلی پیشرفته برای توسعه بازیکنان، مدیریت تیم و تصمیم‌گیری‌های استراتژیک در بسکتبال حرفه‌ای تأکید می‌کند.

واژگان کلیدي: هوش مصنوعی، تحلیل بقا، بسکتبال، طول عمر بازیکنان NBA، داده‌های پزشکی، داده‌های عملکردی، یادگیری ماشین

-
- مقدمه

پیش‌بینی طول عمر حرفه‌ای ورزشکاران از مراحل اولیه، هم از دیدگاه ورزشکار و هم از منظر باشگاه، یک وظیفه جذاب است، زیرا تصمیم‌گیری در مورد استراتژی‌های آینده به نفع هر دو طرف است. این

وظیفه زمانی چالش‌برانگیزتر و در عین حال جذاب‌تر می‌شود که فرد بر آمارهای عملکرد حرفه‌ای در مراحل اولیه تکیه کند. اگرچه مقالات پژوهشی متعددی در این زمینه وجود دارد، اما در این مقاله، توجه خود را به ورزشکاران بسکتبال حرفه‌ای ایالات متحده که در لیگ NBA رقابت می‌کنند، محدود خواهیم کرد. به طور خاص، ما بر این موضوع تمرکز خواهیم کرد که آیا یک بازیکن NBA حداقل ۱۰ سال در لیگ دوام می‌آورد یا خیر، که حداقل تعداد سال‌های لازم برای دریافت کامل طرح بازنشستگی است.

استاو و هوانگ (۱۹۹۵) به بررسی عوامل مؤثر بر طول عمر بازیکنان NBA پرداختند و داده‌های درفت‌های ۱۹۸۰-۱۹۸۶ تا فصل ۱۹۹۰-۱۹۹۱ را تحلیل کردند. تحلیل تاریخچه رویداد آن‌ها نشان داد که شماره درفت اولیه بازیکن، متغیرهای عملکردی و مدت زمان حضور در لیگ، عوامل محوری مؤثر بر طول عمر حرفه‌ای هستند، به طوری که توانایی امتیازآوری به طور قابل توجهی بر زمان بازی و احتمال ماندن در لیگ تأثیر می‌گذارد. این مطالعه همچنین نشان داد که فرنچایزهای NBA تمایل داشتند انتخاب‌های درفت بالا را نسبت به انتخاب‌های درفت پایین حفظ کنند، در حالی که مهارت‌های دفاعی مانند ریباند و بلاک شات به طور مثبتی بر شانس ماندن بازیکن در لیگ تأثیر می‌گذاشت، به ویژه در تیم‌هایی که به مهارت‌های تیمی اهمیت می‌دادند. به موازات آن، گروتهویس و هیل (۲۰۰۴) و گروتهویس و هیل (۲۰۱۸) تحقیقات قابل توجهی در مورد تبعیض خروج در NBA، با تمرکز بر بازیکنان غیرآمریکایی، انجام دادند. گروتهویس و هیل (۲۰۰۴) عوامل مؤثر بر مدت زمان حرفه را بررسی کردند و بر تمایل مالکان تیم به حفظ بازیکنان مولد تأکید کردند، به طوری که پاس گل، بلاک و امتیاز در هر دقیقه بازی بر ادامه حضور در NBA تأثیر می‌گذاشت. قد، وزن و شماره درفت نیز به عنوان پیش‌بینی‌کننده‌های قابل توجه طول عمر حرفه‌ای شناسایی شدند، در حالی که نژاد به تبعیض خروج کمک نکرد. گروتهویس و هیل (۲۰۱۸) این موضوع را با آشکار ساختن این نکته گسترش دادند که بازیکنان خارجی بدون تجربه کالج در ایالات متحده تمایل به داشتن حرفه کوتاه‌تری داشتند، که احتمالاً نشان‌دهنده تبعیض خروج یا ترجیح برای پایان دادن به حرفه خود در کشورهای مبدأشان بود. برعکس، بازیکنان با تجربه کالج در ایالات متحده، طول عمر حرفه‌ای قابل مقایسه با بازیکنان بومی را نشان دادند، که بر تعامل پیچیده ترجیحات طرفداران، پویایی‌های فرهنگی و فرصت‌های سودآور در شکل‌دهی مسیرهای شغلی بازیکنان خارجی NBA تأکید می‌کند. این بینش‌ها به طور قابل توجهی به اقتصاد ورزش و گفتمان گسترده‌تر در مورد پویایی استعدادهای بین‌المللی در لیگ‌های ورزشی حرفه‌ای کمک می‌کند.

پترسن و همکاران (۲۰۱۱) اثر متیو ("ثروتمندترها ثروتمندتر می‌شوند") را نشان دادند که در آن طول عمر و موفقیت گذشته یک فرد به پیشرفت بیشتر شغلی کمک می‌کند. این مطالعه به طور مؤثری نشان داد که حتی یک نرخ پیشرفت متوسط در آغاز حرفه یک فرد، نقش حیاتی در شکل‌دهی مسیر طول عمر حرفه‌ای دارد. این مدل به طور پیچیده‌ای اثر متیو را در خود جای داد و بر اهمیت حیاتی توسعه شغلی اولیه تأکید کرد. این کار نابرابری‌های ذاتی بین حرفه‌های کوتاه و طولانی را روشن کرد و یک آمار قانع‌کننده را نشان داد که تقریباً ۳٪ از بازیکنان بسکتبال در صورتی که کمتر از ۱۲ دقیقه در هر بازی بازی کنند، حرفه NBA را تجربه می‌کنند. علاوه بر این، این تحقیق تأکید کرد که ورزشکارانی که از حرفه‌های طولانی‌مدت برخوردار بودند، به طور موفقیت‌آمیزی سطح بالایی از عملکرد را در طول یک بازه زمانی قابل توجه از زمان بازی حفظ کردند.

دو مطالعه منتشر شده در سال ۲۰۰۸ به بررسی اهمیت بسکتبال کالج در شکل‌دهی مسیر شغلی یک

بازیکن NBA پرداختند. کوتز و اوگان‌تایمین (۲۰۱۰) با تمرکز بر کلاس‌های درفت NBA از سال ۱۹۸۷ تا ۱۹۸۹، قابلیت پیش‌بینی حرفه‌های موفق را بر اساس عملکرد کالج با بررسی بازیکنان بازنشسته ارزیابی کردند. تحلیل، با در نظر گرفتن داده‌های جامع در مورد جزئیات درفت و معیارهای عملکرد، نشان داد که بازیکنان از کنفرانس‌های کوچک‌تر، کارایی بالاتری را نشان می‌دهند که ناشی از امتیازات و ریباندهای کالج برتر است. با وجود دستیابی به تولید NBA مشابه، بازیکنان از کنفرانس‌های کوچک‌تر، حرفه‌های کوتاه‌تری را نسبت به همتایان خود از کنفرانس‌های بزرگ‌تر تجربه کردند، که مفاهیم رایج در مورد تبعیض آماری و ارزش گزینه را به چالش می‌کشد. با بررسی همبستگی بین عملکرد کالج و NBA، کوتز و اوگان‌تایمین (۲۰۱۰) بینش‌های ارزشمندی را در مورد پیچیدگی‌های تصمیمات درفت و مسیرهای شغلی بازیکنان ارائه دادند. بارنز (۲۰۰۸) رابطه بین متغیرهای آماری حرفه قبل از NBA و طول عمر بازیکنان NBA را با توجه به پست‌های بازی بازیکنان، گارد، فوروارد و سنتر، بررسی کرد. با تحلیل داده‌های فصول کالج ۱۹۸۸-۲۰۰۲، آن‌ها از ۱۱ متغیر مستقل مانند امتیازات، پاس گل و توپ لو رفته استفاده کردند، در حالی که طول عمر حرفه‌ای متغیر وابسته بود. تحلیل آماری ارتباطات قابل توجهی را برای گاردها و فورواردها آشکار کرد و بر تأثیر پاس گل، توپ لو رفته، امتیازات، درصد پرتاب‌های میدانی و درصد پرتاب‌های آزاد بر طول عمر حرفه‌ای NBA تأکید کرد. نکته قابل توجه این است که این مطالعه برای سنترها اهمیت آماری را ناچیز یافت و آن را به ماهیت منحصر به فرد پست سنتر و اندازه نمونه کوچک‌تر نسبت داد. این یافته‌ها بر پتانسیل تحلیل آماری در کمک به مدیران کل و استعدادیاب‌های NBA در ارزیابی مؤثر بازیکنان و استراتژی‌های انتخاب، و ارائه بینش‌های ارزشمند در فرآیند پیچیده ساخت تیم‌های بسکتبال موفق، تأکید می‌کند. میگوئل و همکاران (۲۰۱۹) تحلیل گسترده‌ای از داده‌های درفت NBA از سال ۱۹۷۸ تا ۱۹۹۸ انجام دادند و بینش‌های قانع‌کننده‌ای را در مورد رابطه بین ترتیب انتخاب درفت و طول عمر حرفه‌ای بازیکنان آشکار کردند. بازیکنانی که در پنج انتخاب اول انتخاب شدند، به طور متوسط، حرفه‌ای طولانی‌تر در حدود ۱۴ سال داشتند، با یک روند غیرخطی قابل تشخیص که کاهش طول عمر را از انتخاب اول تا سی‌ام نشان می‌داد. با در نظر گرفتن سال‌های درفت، این مطالعه نوساناتی را در طول عمر حرفه‌ای شناسایی کرد، با افزایش تا سال ۱۹۸۵، تثبیت تا سال ۱۹۹۳ و افزایش بعدی.

فین و سوننشاین (۲۰۱۲) از معیارهای عملکردی مرسوم بازیکنان فاصله گرفتند و به جای آن، جوایز فردی را به عنوان معیاری برای موفقیت انتخاب کردند. آن‌ها تعداد جوایز فردی کسب شده را به همراه داده‌های بیولوژیکی بازیکن مانند قد و وزن به عنوان معیاری از عملکرد به کار گرفتند. آن‌ها اشاره کردند که قد بازیکن و تعداد جوایز کسب شده تأثیر مثبتی بر مدت زمان حرفه او دارد. ارتباط بین قد و طول عمر حرفه‌ای طولانی‌تر را می‌توان به کمیاب بودن بازیکنان در پست‌هایی مانند سنتر و فوروارد-سنتر در مقایسه با گاردها یا گارد-فورواردها نسبت داد، که باعث می‌شود اولی‌ها به دلیل توانایی‌هایشان در به ثمر رساندن شوت‌های نزدیک سبد، ریباند و بلاک شات، بدون در نظر گرفتن معیارهای عملکرد خاص، بیشتر مورد توجه قرار گیرند.

طول عمر حرفه‌ای علاوه بر این، به عوامل مختلفی بستگی دارد، به طوری که آسیب‌دیدگی‌ها و بیماری‌های فصلی نقش قابل توجهی در تعیین طول عمر حرفه‌ای یک بازیکن NBA ایفا می‌کنند، همانطور که در مطالعات زیر نشان داده شده است. کستر و همکاران (۲۰۱۷) تحقیقات جامعی را در مورد تأثیر پارگی رباط صلیبی قدامی (ACL) بر بازیکنان NBA از سال ۱۹۸۴ تا ۲۰۱۴ انجام دادند. با وجود نرخ بازگشت ۸۶.۱٪ پس از بازسازی ACL، این مطالعه نشان داد که

میانگین زمان بازی پس از عمل به طور قابل توجهی کوتاه‌تر و ۱.۸۴ سال در مقایسه با گروه کنترل بود. تحلیل بقا بر نرخ بالای فرسایش زودهنگام برای بازیکنانی که تحت بازسازی ACL قرار می‌گیرند تأکید کرد و رابطه پیچیده بین این آسیب‌دیدگی‌ها، موفقیت توانبخشی و پیامدهای پایدار بر طول عمر حرفه‌ای بازیکنان بسکتبال حرفه‌ای را برجسته کرد. خلیل و همکاران (۲۰۲۰) با استفاده از گروه‌های کنترل همسان، پیامدهای پارگی تاندون آشیل (AT) بر حرفه بازیکنان NBA از سال ۱۹۷۰ تا ۲۰۱۹ را بررسی کردند و نشان دادند که از میان ۴۷ بازیکن با پارگی AT، ۷۲.۳٪ به طور موفقیت‌آمیزی پس از جراحی به مشارکت در NBA بازگشتند، هرچند با حرفه‌های بازی به طور قابل توجهی کوتاه‌تر در مقایسه با همتایان بدون آسیب‌دیدگی (به ترتیب به طور متوسط ۳.۱ در مقابل ۵.۸ فصل). جانز و همکاران (۲۰۲۱) مروری را انجام دادند که تأثیر پارگی تاندون آشیل (AT) را بر ۳۳۳ ورزشکار حرفه‌ای در لیگ‌های ورزشی بزرگ بررسی کرد. یافته‌ها نرخ بازگشت به بازی ۷۶.۴٪ را پس از ترمیم AT نشان می‌دهند، با میانگین زمان بهبودی ۱۱ ماه - دو برابر جمعیت عمومی. با این حال، ورزشکاران بازگشته کاهش قابل توجهی در عملکرد را تجربه کردند، به ویژه در NFL و NBA، که نشان‌دهنده یک پیامد بالقوه تغییردهنده حرفه است. این مطالعه چالش‌های این ورزشکاران را برجسته کرد و بینش‌های حیاتی را برای تعیین انتظارات مبتنی بر شواهد در بازگشت پس از عمل به ورزش حرفه‌ای ارائه داد.

مارتین و همکاران (۲۰۲۱) به طور خاص بر آسیب‌دیدگی‌ها در فصل تازه‌کار یک بازیکن NBA تمرکز کردند. با استفاده از داده‌های ۲۰۰۷ تا ۲۰۱۹، آن‌ها نرخ بالای آسیب‌دیدگی و بیماری را در بازیکنان تازه‌کار، به ویژه در مچ پا، آشکار کردند. آن‌ها ارتباط بین آسیب‌دیدگی‌های فصل تازه‌کار و طول عمر حرفه‌ای را بررسی کردند. نتایج نشان داد که کاهش قابل توجهی در مجموع فصول بازی شده برای تازه‌کاران با آسیب‌دیدگی وجود دارد، اما این اثر پس از در نظر گرفتن متغیرهای مخدوش‌کننده کاهش یافت. موقعیت‌های درفت پایین‌تر با حرفه‌های NBA کوتاه‌تر مرتبط بودند، که نشان می‌دهد عوامل عملکردی و سرمایه‌گذاری‌های سازمانی نقش دارند. الگوهای آسیب‌دیدگی خاص، به ویژه آسیب‌دیدگی‌های مچ پا و زانو، پیامدهای بلندمدت را برجسته کردند و از برنامه‌های کاهش هدفمند حمایت کردند. در حالی که تازه‌کاران خطر آسیب‌دیدگی بالاتری را نشان دادند، تحلیل‌های تعدیل شده نشان داد که طول عمر حرفه‌ای چندعاملی است، با بار آسیب‌دیدگی تجمعی که به عنوان یک عامل تعیین‌کننده بالقوه ظاهر می‌شود، که بر نیاز به تحقیقات مداوم و استراتژی‌های کاهش بهبود یافته تأکید می‌کند.

هدف این مقاله با کارهای تحقیقاتی قبلی متفاوت است، زیرا به جای تلاش برای پیش‌بینی مدت زمان حضور بازیکنان NBA، تلاش می‌کند احتمال ماندن در لیگ برای حداقل ۱۰ سال را پیش‌بینی کند. بازیکنانی که حداقل ۳ سال در لیگ خدمت کرده‌اند، واجد شرایط دریافت حداقل بسته بازنشستگی NBA هستند، اما کسانی که ۱۰ سال خدمت کرده‌اند، مستحق طرح بازنشستگی کامل هستند که شامل تمام مزایای ممکن است. بازیکنان NBA می‌توانند پرداخت‌های ماهانه کوچک‌تر را، در طول یک دوره زمانی طولانی‌تر، از سن ۴۵ سالگی تحت طرح بازنشستگی زودهنگام NBA دریافت کنند. بازیکنان تشویق می‌شوند تا دریافت پرداخت‌ها را تا روز بازنشستگی عادی در سن ۶۲ سالگی به تعویق بیندازند تا بالاترین پرداخت‌های ممکن را دریافت کنند. به عنوان مثال، یک بازیکن با تنها سه سال خدمت که در سن ۶۲ سالگی بازنشسته می‌شود، حداقل مبلغ ۵۶,۹۹۸ دلار در سال دریافت خواهد کرد و یک بازیکن با حداقل ۱۰ سال خدمت می‌تواند تا ۲۱۵,۰۰۰ دلار سالانه در سن ۶۲

سالگی دریافت کند. مبلغ بازنشستگی بر اساس ترکیبی از عوامل شامل سال‌های خدمت، سن و سابقه حقوق است (hoopshype.com). لازم به ذکر است که از جولای ۲۰۲۳ به بعد، مبلغ ماهانه برای هر سال خدمت معتبر قابل پرداخت به عنوان حقوق بازنشستگی عادی ۱,۰۰۱.۴۷ دلار است.

الگوریتم‌های آماری و یادگیری ماشین (ML) برای این منظور به کار گرفته شدند، با استفاده از داده‌های جمع‌آوری شده از بازیکنان درفت شده بین سال‌های ۱۹۹۹ و ۲۰۰۶. آمارهای عملکرد سال دوم آن‌ها به الگوریتم‌ها وارد شد و عملکرد پیش‌بینی‌کننده بالایی را به دست آورد. همین تحلیل دوباره انجام شد، اما این بار با حفظ مهم‌ترین آمارها، و عملکرد پیش‌بینی‌کننده در همان سطوح دقت باقی ماند. به عنوان یک طرح اعتبارسنجی اضافی برای مدل بهینه انتخاب شده، داده‌های اخیرتر جمع‌آوری شد که عملکرد بسیار رضایت‌بخشی از مدل را نشان داد. این مدل ساده برای استفاده و تفسیر است و می‌تواند به عنوان اولین پیش‌بینی طول عمر حرفه‌ای یک بازیکن عمل کند، در حالی که تأثیر هر عملکرد را بر احتمال بقا در لیگ برای حداقل ۱۰ سال نشان می‌دهد.

مزایای بازنشستگی

شایان ذکر است که تغییر قابل توجهی در واجد شرایط بودن بازیکنان برای مزایای بازنشستگی بین توافقنامه چانه‌زنی جمعی (CBA) قدیمی ۲۰۱۸-۲۰۲۳ و توافقنامه به‌روز شده که در جولای ۲۰۲۳ به اجرا درآمد، وجود دارد. طبق CBA قبلی، یک بازیکن صرفاً با شرکت یا فعال بودن در یک بازی، یک فصل کامل را به پایان رسانده تلقی می‌شد. این قانون حتی برای بازیکنان دوطرفه نیز اعمال می‌شد، که حقوق آن‌ها برای روزهای کاری NBA در مجموع حقوق و مزایا گنجانده می‌شد و در نتیجه به سهم بازیکنان از درآمد مرتبط با بسکتبال (BRI) کمک می‌کرد. با این حال، با اجرای CBA جدید، که از جولای ۲۰۲۳ به اجرا درآمد، معیارهای یک بازیکن برای قرار گرفتن در لیست تیم دستخوش بازنگری قابل توجهی شد. اکنون، یک بازیکن در صورتی در لیست تیم تلقی می‌شود که در تاریخ ۲ فوریه فصل عادی جاری، به عنوان فعال، غیرفعال، یا در لیست دوطرفه هر تیمی قرار گرفته باشد، یا اگر برای حداقل پنجاه درصد (۵۰٪) از کل بازی‌های فصل عادی انجام شده توسط تیم، در لیست فعال باشد. این تغییر در معیارهای واجد شرایط بودن به طور خاص بر بازیکنان در سال دوم حرفه حرفه‌ای آن‌ها تأثیر می‌گذارد، زیرا طبق قانون CBA قدیمی تا فصل ۲۰۲۲-۲۰۲۳، آن‌ها می‌توانستند صرفاً با داشتن قرارداد برای حداقل یک بازی در سال سوم حضور خود در لیگ، واجد شرایط بازنشستگی شوند. به طور مشابه، بازیکنان با قرارداد دوطرفه می‌توانستند با قرار گرفتن در لیست تیم NBA برای تنها یک بازی، واجد شرایط مزایای بازنشستگی شوند.

علاوه بر این، به عنوان بخشی از توافقنامه CBA 2017، بهبودهای قابل توجهی در مزایای مراقبت‌های بهداشتی و آموزشی در دسترس بازیکنان NBA انجام شد. به طور خاص، بازنشستگان با حداقل سه سال خدمت در NBA اکنون پوشش مراقبت‌های بهداشتی مادام‌العمر دریافت می‌کنند، که این یک بند بی‌سابقه در سایر انجمن‌های بازنشستگان است. علاوه بر این، کسانی که یک دهه یا بیشتر در لیگ خدمت کرده‌اند، از پوشش مراقبت‌های بهداشتی جامع نه تنها برای خود بلکه برای همسر و فرزندانشان نیز برخوردار هستند. این برنامه جامع مراقبت‌های بهداشتی استاندارد جدیدی را در ورزش حرفه‌ای تعیین می‌کند و اطمینان می‌دهد که بازیکنان بازنشسته NBA و خانواده‌هایشان برای زندگی به خوبی مراقبت می‌شوند. علاوه بر مزایای مراقبت‌های بهداشتی، CBA همچنین مقرراتی را برای حمایت آموزشی معرفی کرد. بازیکنان بازنشسته‌ای که مایل به ادامه تحصیل هستند، می‌توانند شهریه خود را تا سقف ۳۳,۰۰۰ دلار سالانه بازپرداخت کنند، که در CBA 2023 به ۶۲,۵۰۰ دلار

برای هر سال تقویمی تغییر یافت. این کمک با هدف تسهیل انتقال به حرفه‌های پس از بسکتبال و تشویق یادگیری مادام‌العمر است. مزایای آموزشی حتی فراتر می‌رود، زیرا هر بازیکن واجد شرایط با سه (۳) سال یا بیشتر خدمت در NBA از تاریخ توافقنامه CBA 2023، یک افزایش یک‌باره در مزایای بازپرداخت شهریه معادل ۲۴,۰۰۰ دلار دریافت خواهد کرد تا یا مدرک ناتمام خود را تکمیل کند یا تحصیلات بیشتر را دنبال کند. طبق آخرین آمارهای موجود تا سپتامبر ۲۰۱۹، ۲۸ بازیکن قبلاً برای بازپرداخت شهریه تأیید شده‌اند، و بیش از ۵۰ نفر دیگر در انتظار تأیید هستند. این ابتکارات بر تعهد NBA به حمایت از بازیکنان خود فراتر از حرفه بازی آن‌ها، ترویج رفاه کلی و توسعه شخصی مستمر آن‌ها تأکید می‌کند (NBA-NBPA Collective Bargaining Agreement 2017).

همچنین مهم است که توجه داشته باشیم که NBA یک طرح مزایای قوی ۴۰۱(k) را در CBA 2011 معرفی کرد که متعاقباً در CBA 2017 و 2023 بازنویسی شد و به طور خاص برای بازیکنان خود طراحی شده است. این ابتکار یک مسیر ساختاریافته برای امنیت مالی در طول و پس از حرفه بازی آن‌ها فراهم می‌کند. تحت این طرح، بازیکنان می‌توانند بخشی از درآمد خود را به حساب ۴۰۱(k) اختصاص دهند، با مشارکت‌هایی که به صورت پیش از کسر مالیات انجام می‌شود، که به طور مؤثری درآمد مشمول مالیات آن‌ها را کاهش می‌دهد و به سرمایه‌گذاری‌ها اجازه می‌دهد بدون مالیات رشد کنند. علاوه بر این، NBA به عنوان یک کارفرما، این مشارکت‌ها را تطبیق می‌دهد و تا ۱۴۰ درصد از مشارکت‌های خود بازیکن را ارائه می‌دهد. این طرح تطبیق سخاوتمندانه به عنوان یک تشویق قانع‌کننده برای بازیکنان عمل می‌کند تا برنامه‌ریزی مالی بلندمدت را در اولویت قرار دهند. در طرح ۴۰۱(k)، بازیکنان با طیف متنوعی از گزینه‌های سرمایه‌گذاری مواجه می‌شوند که آن‌ها را قادر می‌سازد تا استراتژی‌های سرمایه‌گذاری خود را با اهداف مالی فردی و تحمل ریسک خود تنظیم کنند. علاوه بر این، ساختار طرح NBA (k)۴۰۱ عادات پس‌انداز منظم را تضمین می‌کند، زیرا بازیکنان عموماً نمی‌توانند تا رسیدن به سن ۵۹-۶۰ سالگی بدون متحمل شدن جریمه به وجوه دسترسی پیدا کنند. این محافظت برای تقویت بازیکنان طراحی شده است تا با پول قابل توجهی بیشتر در پس‌انداز از بازی کناره‌گیری کنند.

توصیف داده‌ها

نمونه شامل بازیکنانی است که در حین تحصیل در کالج ایالات متحده یا پس از اتمام دوره کارشناسی درفت شده‌اند. بازیکنانی که خارج از کالج ایالات متحده درفت شده‌اند، به عنوان مثال بازیکنان حرفه‌ای از اروپا یا سایر قاره‌ها، از تحلیل حذف شدند. دلیل این امر جلوگیری از سوگیری احتمالی ناشی از این واقعیت است که بازیکنان قبلاً در بسکتبال حرفه‌ای تجربه داشتند یا سن بالاتری داشتند.

بازیکنانی که از سال ۱۹۹۹ تا ۲۰۰۶ درفت شده بودند، مورد بررسی قرار گرفتند و آمارهای عملکردی درون زمین سال دوم آن‌ها، مانند امتیازات کسب شده (PTS)، مجموع ریباندها (TRB)، ریباندهای تهاجمی (ORB)، پاس گل (AST)، بلاک (BLK)، توپ ربایی (STL)، توپ لو رفته (TO) و دقایق بازی (MP) جمع‌آوری شد. علاوه بر این، کارایی یک بازیکن در حمله و دفاع با بررسی درصدهای پرتاب‌های میدانی موفق (FG٪)، پرتاب‌های سه امتیازی (3P٪) و پرتاب‌های آزاد (FT٪) در نظر گرفته شد. سایر متغیرهایی که در مدل گنجانده شده‌اند عبارتند از سن (AGE) در سال دوم، بازی‌های انجام شده (GP) در آن فصل و انتخاب درفت (DP). علاوه بر این، پست بازیکن نیز در نظر گرفته شد، به طوری که فورواردها و سنترها به دلیل قد و وزن خود، احتمال بیشتری برای ماندن طولانی‌تر در NBA دارند، همانطور که از مقالات تحقیقاتی قبلی نشان داده شده است.

برای ارائه بینش‌های عمیق‌تر در مورد کارایی بازیکن،

دو متغیر جدید معرفی شد: نسبت پاس گل به توپ لو رفته (AST/TO) و نسبت پاس گل به علاوه امتیازات به توپ لو رفته ((TO/(AST + PTS). این نسبت‌ها توانایی بازیکن را در کمک مثبت به عملکرد تیم خود با ایجاد فرصت‌های امتیازآوری از طریق پاس گل و امتیازات و در عین حال به حداقل رساندن توپ لو رفته، اندازه‌گیری می‌کنند. مقادیر بالاتر این نسبت‌ها نشان‌دهنده مشارکت کلی و کارایی بیشتر در تیم است که به طور بالقوه می‌تواند حرفه یک بازیکن را در لیگ طولانی‌تر کند. با گنجاندن طیف گسترده‌ای از شاخص‌های عملکردی و معیارهای کارایی، مدل ما قصد دارد درک جامعی از طول عمر بالقوه یک بازیکن در NBA را با در نظر گرفتن عملکرد فردی ارائه دهد. این رویکرد چندوجهی، ارزیابی دقیق‌تری از تأثیر و دوام یک بازیکن در طولانی مدت در لیگ را تضمین می‌کند.

ترکیب تمام سال‌های داده در یک پایگاه داده واحد برای تحلیل ما به دلیل اطلاعات گمشده از منابع مختلف، یک چالش بود. این ناهماهنگی عمدتاً به این دلیل ایجاد شد که برخی از بازیکنان در حرفه خود وقفه داشتند، یا به دلیل آسیب‌دیدگی یا ترک موقت لیگ در سال دوم خود، و سپس بازگشت. برای رفع این مشکل، یک رویکرد دقیق برای تعریف سال دوم یک بازیکن اتخاذ شد. به جای صرفاً تکیه بر سال تقویمی متوالی، سال دوم واقعی در لیگ با در نظر گرفتن تنها آن فصولی که بازیکن به طور فعال در حداقل یک بازی با زمان بازی شرکت کرده بود، شناسایی شد. این امر نمایش دقیق‌تری از پیشرفت و تداوم هر بازیکن در لیگ را تضمین کرد. تمام محاسبات با استفاده از نرم‌افزار آماری R Team (2023) انجام شد.

آمار توصیفی داده‌ها

جدول ۱ آمار توصیفی داده‌ها را ارائه می‌دهد. از ۳۲۲ بازیکن NBA در نمونه ما، ۱۵۶ نفر (۴۸.۴۵٪) به طرح بازنشستگی کامل دست یافتند. علاوه بر این، تنها ۴ بازیکن در فصل ۲۰۲۳-۲۰۲۴ هنوز فعال هستند، یعنی لبرون جیمز، کریس پاول، رودی گی و کایل لوری. شایان ذکر است که جیمز از معدود بازیکنانی است که ۲۱ سال در لیگ بازی کرده است و لوری به عنوان انتخاب ۲۴ درفت شد. از آنجایی که معیارهای عملکرد پس از سال دوم بازیکنان جمع‌آوری شده است، واضح است که آن‌ها حداقل دو سال در لیگ مانده‌اند و حداقل یک بازی در سال دوم خود انجام داده‌اند. اکثر بازیکنان صرفاً در پست‌های گارد یا فوروارد بازی می‌کنند، در حالی که بازیکنانی نیز بودند که یک سال یا حتی دو سال پس از درفت شدن شروع به بازی کردند.

تمام متغیرهای پیش‌بینی‌کننده به طور آماری به طور معنی‌داری با سال‌های حضور در لیگ مرتبط هستند، همانطور که در جدول ۱ ارائه شده است ، و همانطور که انتظار می‌رود، طول عمر حرفه‌ای بازیکنان با سن و انتخاب درفت آن‌ها رابطه منفی دارد. تمام متغیرهای پیش‌بینی‌کننده از نظر آماری معنی‌دار تلقی شدند (اکثر مقادیر p برابر با ۰.۰۰۱ یا کمتر بودند به جز نسبت پاس گل به توپ لو رفته که برابر با ۰.۰۱ بود) هنگامی که رگرسیون لجستیک برای متغیر پاسخ (دستیابی به بازنشستگی کامل) استفاده شد. آزمون استقلال χ^2 به طور جزئی مفروضات استقلال بین متغیر پاسخ و سال درفت (مقدار p=۰.۰۴۵)، بین متغیر پاسخ و راند انتخاب (مقدار p=۰.۰۲۶) را رد کرد، اما آن را بین متغیر پاسخ و پستی که بازی می‌کنند (مقدار p=۰.۲۲۷) رد نکرد. در نهایت، آزمون F ولش فرض برابری میانگین سال‌های حضور در لیگ را در ۸ فصل مورد مطالعه رد نکرد (مقدار p=۰.۱۱۱).

جدول ۱: آمار توصیفی داده‌ها

متغیر	همبستگی	حداقل	حداکثر	میانه	میانگین	انحراف معیار	پست	سال درفت	سال شروع
YRS		2	22	9	8.991	4.932	G: ۱۰۵	۱۹۹۰: ۷۵	همان سال: ۲۷۴
AGE	0.-328	19	26	23	22.484	1.565	G-F: ۳۸	۲۰۰۰: ۳۴	۱ سال بعد: ۴۷
GP	0.448	1	82	63	56.304	23.925	F: ۱۰۴	۲۰۰۱ ۴۰:	۲ سال بعد: ۱
FG%	0.253	0	0.75	0.437	0.433	0.084	F-C: ۳۵	۲۰۰۲: ۳۰	
۳P%	0.204	0	1	0.268	0.217	0.172	C: ۴۰	۲۰۰۳: ۳۱	
FT%	0.236	0	1	0.732	0.7014	0.154		۲۰۰۴: ۳۵	
MP	0.576	1	42.4	18	19.525	9.812		۲۰۰۵: ۴۲	
PTS	0.579	0	27.2	6	7.492	5.131		۲۰۰۶: ۳۵	
TRB	0.444	0	12.5	2.95	3.413	2.241			
ORB	0.322	0	4	0.8	1.032	0.773			
AST	0.4	0	9.3	1	1.564	1.695			

ستون اول به همبستگی‌های پیرسون بین سال‌های حضور در لیگ و متغیرهای پیش‌بینی‌کننده اشاره دارد، با همبستگی‌های آماری معنی‌دار در سطح معنی‌داری ۵٪ که به صورت پررنگ ظاهر می‌شوند.

دستیابی به بازنشستگی و شناسایی عوامل کلیدی عملکرد

پنج الگوریتم آماری و یادگیری ماشین مختلف برای پیش‌بینی احتمال دستیابی یک بازیکن به طرح بازنشستگی کامل به کار گرفته شد. الگوریتم‌های زیر استفاده شدند:

الاستیک نت (Elastic Net)

الاستیک نت، که برای نخستین بار توسط Zou و Hastie در سال 2005 معرفی شد، یک مدل رگرسیونی پیشرفته با هدف منظم‌سازی و بهبود عملکرد پیش‌بینی مدل‌های خطی است. این روش ترکیبی از دو تکنیک مشهور منظم‌سازی، یعنی LASSO (که توسط Tibshirani در سال 1996 معرفی شد) و رگرسیون ریج (که توسط Hoerl و Kennard در سال 1970 ارائه شد) را در خود جای داده است. ترکیب این دو رویکرد به مدل امکان می‌دهد تا همزمان از ویژگی‌های انتخاب متغیر خودکار LASSO و ثبات پارامتری رگرسیون ریج بهره‌مند شود.

الاستیک نت به طور خاص در شرایطی مناسب است که میان متغیرهای پیش‌بینی‌کننده هم‌بستگی بالایی وجود دارد، مسئله‌ای که بسیاری از روش‌های رگرسیونی کلاسیک را با چالش مواجه می‌سازد. این مدل با افزودن یک اصطلاح جریمه ترکیبی به تابع زیان، از بیش‌برازش جلوگیری کرده و در عین حال امکان انتخاب متغیرهای مؤثر را فراهم می‌سازد. شدت این جریمه توسط دو پارامتر تنظیم‌پذیر کنترل می‌شود: یکی برای تعیین قدرت کلی منظم‌سازی، و دیگری برای تنظیم نسبت مشارکت هر یک از اجزای LASSO و ریج در جریمه ترکیبی.

			0.459	0.428	0.3	2.6	0	0.283	BLK
			0.432	0.623	0.5	2.2	0	0.488	STL
			0.753	1.184	1	4.2	0	0.508	TO
			0.82	1.212	1	6.667	0	0.129	ASTO
			4.105	9.313	9.1	44	0	0.169	ASPTTO
								0.1–97	DP

در این مطالعه، الاستیک نت به عنوان یک روش پایه برای مدل‌سازی احتمال وقوع "بازنشستگی کامل" بازیکنان مورد استفاده قرار گرفته و با بهره‌گیری از بسته نرم‌افزاری glmnet در محیط R پیاده‌سازی شده است.

رگرسیون تعقیب پروجکشن (Projection Pursuit Regression)

رگرسیون تعقیب پروجکشن، که در سال 1981 توسط Friedman و Stuetzle معرفی شد، روشی ناپارامتری برای مدل‌سازی رابطه بین متغیر پاسخ و مجموعه‌ای از متغیرهای پیش‌بینی‌کننده است. برخلاف مدل‌های خطی کلاسیک، این روش قادر است ساختارهای غیرخطی پیچیده و تعامل‌های پنهان بین متغیرها را شناسایی کند، بدون آن‌که نیاز به مشخص‌کردن شکل دقیق رابطه از پیش باشد.

در این روش، متغیر پاسخ به‌صورت ترکیبی از توابع هموارساز تعریف می‌شود که هر یک از آن‌ها بر روی ترکیبی خطی از متغیرهای پیش‌بینی‌کننده عمل می‌کنند. این رویکرد انعطاف‌پذیر باعث می‌شود تا مدل بتواند به‌صورت مؤثر ساختارهای غیرخطی و غیرخطی-ترکیبی را که ممکن است در داده‌ها نهفته باشند، استخراج نماید.

رگرسیون تعقیب پروجکشن به‌ویژه در مواردی که داده‌ها از ساختاری پیچیده یا چندوجهی برخوردارند مفید واقع می‌شود. در این تحقیق، این روش با استفاده از توابع پیش‌فرض موجود در زبان R به‌کار گرفته شده تا امکان مقایسه عملکرد آن با سایر الگوریتم‌ها فراهم شود.

ماشین بردار پشتیبان (Support Vector Machine)

ماشین بردار پشتیبان (SVM)، که توسط Cortes و

Vapnik در سال 1995 توسعه یافت، یکی از الگوریتم‌های قدرتمند یادگیری ماشین برای طبقه‌بندی داده‌ها است. این الگوریتم تلاش می‌کند مرزی بهینه را بین دو کلاس مختلف ایجاد کند به‌گونه‌ای که فاصله (حاشیه) بین نزدیک‌ترین نمونه‌های هر کلاس به این مرز بیشینه شود. این ویژگی، به SVM توانایی بالایی در تعمیم‌پذیری داده‌های جدید می‌دهد.

الگوریتم SVM با استفاده از نگاشت داده‌ها به فضای با بُعد بالاتر، قادر به تشخیص روابط غیرخطی پیچیده در میان داده‌هاست. در این فضای جدید، داده‌ها با استفاده از یک ابرصفحه بهینه از یکدیگر جدا می‌شوند. قدرت این مدل در استفاده از توابع هسته‌ای (kernel functions) نهفته است که امکان جداسازی داده‌های غیرخطی را فراهم می‌سازد.

در چارچوب این پژوهش، SVM به‌صورت خاص برای پیش‌بینی عضویت بازیکنان در یکی از دو گروه "بازنشسته کامل" یا "غیربازنشسته" به‌کار گرفته شده و از پیاده‌سازی آن در بسته e1071 در نرم‌افزار R استفاده شده است.

جنگل تصادفی (Random Forest)

جنگل تصادفی، که توسط Breiman در سال 2001 معرفی شد، یک الگوریتم یادگیری ماشین غیرخطی مبتنی بر تجمیع چندین درخت تصمیم است. ایده اصلی در این روش، ساخت مجموعه‌ای بزرگ از درخت‌های تصمیم است که هر یک بر اساس یک نمونه بوت‌استرپ از داده‌های آموزشی ساخته شده‌اند. سپس نتایج پیش‌بینی این درخت‌ها با یکدیگر ترکیب می‌شود تا یک پیش‌بینی نهایی و پایدار به‌دست آید.

یکی از مزیت‌های کلیدی جنگل تصادفی، توانایی آن در کنترل بیش‌برازش از طریق تنوع ایجادشده میان درخت‌هاست. این تنوع با انتخاب تصادفی زیرمجموعه‌ای از ویژگی‌ها برای هر درخت ایجاد می‌شود. نتیجه نهایی، مدلی است که از دقت بالا و مقاومت مناسب در برابر نویز برخوردار است.

در این تحقیق، یک مدل جنگل تصادفی متشکل از ۵۰۰ درخت تصمیم مورد استفاده قرار گرفته و با بهره‌گیری از بسته نرم‌افزاری ranger در محیط R پیاده‌سازی شده است. هدف از استفاده از این مدل، بررسی عملکرد آن در مقایسه با سایر الگوریتم‌ها در پیش‌بینی احتمال بازنشستگی کامل بازیکنان بوده است.

ماشین تقویت گرادیان (Gradient Boosting Machine)

ماشین تقویت گرادیان، که توسط Friedman در سال 2001 پیشنهاد شد، یکی از قدرتمندترین الگوریتم‌های یادگیری ترکیبی است که عملکرد بسیار خوبی در وظایف پیش‌بینی نشان داده است. برخلاف جنگل تصادفی که درخت‌ها را به‌صورت مستقل آموزش می‌دهد، GBMبه‌صورت مرحله‌ای و متوالی عمل می‌کند؛ بدین صورت که هر مدل جدید تلاش می‌کند خطاهای باقی‌مانده از مدل‌های قبلی را کاهش دهد.

در هر مرحله، تابع زیان مدل فعلی ارزیابی شده و گرادیان آن نسبت به مقادیر پیش‌بینی‌شده محاسبه می‌شود. این گرادیان‌ها به عنوان نوعی از باقیمانده در نظر گرفته شده و مدل بعدی با هدف کاهش این مقدار آموزش می‌بیند. این فرآیند تا زمانی ادامه می‌یابد که معیار همگرایی برآورده شود یا بهبود قابل‌توجهی در تابع زیان مشاهده نشود.

در این مطالعه، GBMبا استفاده از پیاده‌سازی آن در بسته gbm در محیط R مورد استفاده قرار گرفته و به عنوان یکی از مدل‌های رقابتی در پیش‌بینی متغیر بازنشستگی کامل بررسی شده است. توانایی این مدل در یادگیری ساختارهای پیچیده و تعامل‌های غیرخطی میان متغیرها، آن را به گزینه‌ای مطلوب برای مسائل پیش‌بینی با ساختار داده پیچیده تبدیل می‌کند.

روش‌شناسی

پروتکل اعتبارسنجی متقابل ۱۰-برابر (CV) (Hastie et al., 2009) برای ارزیابی عملکرد پیش‌بینی‌کننده الگوریتم‌ها به کار گرفته شد. در طول ۱۰ CV-برابر، داده‌ها به ۱۰ بخش مجزا به صورت طبقه‌بندی شده تقسیم می‌شوند و اطمینان حاصل می‌شود که توزیع رویدادها (طرح بازنشستگی کامل) در این بخش‌ها حفظ می‌شود. یک بخش را انتخاب کنید، آن را مجموعه آزمون نامیده و کنار بگذارید. یک الگوریتم را در بخش‌های باقیمانده (مجموعه آموزش) آموزش دهید و سپس از پیش‌بینی‌کننده‌ها در مجموعه آزمون برای پیش‌بینی مقادیر پاسخ‌ها (در مجموعه آزمون) استفاده کنید. دقت پیش‌بینی‌ها را با استفاده از یک معیار عملکردی تخمین بزنید و این فرآیند را ۱۰ بار تکرار کنید تا زمانی که تمام بخش‌ها نقش مجموعه آزمون را ایفا کرده باشند. مساحت زیر منحنی (AUC) برای اندازه‌گیری عملکرد پیش‌بینی‌کننده الگوریتم‌ها در طول پروتکل CV استفاده شد. ۱۰ CV-برابر ۲۰ بار تکرار شد تا منابع احتمالی تغییرات در بین تقسیم‌بندی‌ها را در نظر بگیرد.

علاوه بر این، انتخاب متغیر (VS) به عنوان یک گام اضافی از تحلیل با استفاده از الگوریتم VS غیرخطی Boruta (Kursa et al., 2010) که در بسته Boruta (Kursa and Rudnicki, 2010) موجود است، انجام شد. الگوریتم Boruta به طور تکراری از الگوریتم RF برای انجام هدف خود (VS) استفاده می‌کند و این امکان را برای محاسبه اهمیت متغیر در هر مرحله فراهم می‌کند. این روش VS و عملکرد پیش‌بینی‌کننده (AUC) هر الگوریتم، دوباره با استفاده از پروتکل ۱۰ CV-برابر که ۲۰ بار تکرار شد، اعتبارسنجی متقابل شدند.

کاربرد در سال‌های اخیر

عملکرد مدل با استفاده از داده‌های اخیرتر و به طور خاص ۷۷ بازیکن که در سال‌های ۲۰۱۳ و ۲۰۱۴ درفت شده بودند، ارزیابی شد. آمارهای عملکردی آن‌ها در سال دوم دوباره استخراج شد و تنها ۱۲ متغیر مورد علاقه در نظر گرفته شد. یک تفاوت قابل توجه بین این مجموعه داده و مجموعه داده مورد استفاده در مقاله، نسبت بازیکنانی است که حداقل ۱۰ سال در لیگ دوام می‌آورند. از ۳۲۲ بازیکن درفت شده بین سال‌های ۱۹۹۹ و ۲۰۰۶، ۱۵۶ نفر (۴۸.۴۵٪) موفق به دستیابی به طرح بازنشستگی کامل شدند، در حالی که در سال‌های اخیرتر، از ۷۷ بازیکن درفت شده در سال‌های ۲۰۱۳ و ۲۰۱۴، تنها ۲۶ نفر (۳۳.۷۷٪) ۱۰ سال یا بیشتر در لیگ مانده‌اند.

این کاهش در درصد را می‌توان به عوامل زیادی نسبت داد. به عنوان مثال، بازار به طور قابل توجهی در سراسر جهان گسترش یافته است و تیم‌های NBA طیف وسیع‌تری از فرصت‌ها را، خارج از بازیکنان کالج، در اختیار دارند. برعکس، بازیکنان بیشتری تصمیم می‌گیرند از NBA خارج شوند زیرا حقوق در لیگ‌های اروپایی به طور قابل توجهی افزایش یافته و قراردادهای رقابتی ارائه می‌دهند. در ابتدا، ویژگی‌های داده‌های فعلی ممکن است تا حدی شبیه به ویژگی‌های مشاهده شده در داده‌های گذشته باشند اما یکسان نیستند و با توجه به اندازه نمونه نسبتاً کوچک، اطلاعات موجود کافی نبود. این مربوط به این واقعیت است که NBA یک سیستم در حال تکامل است که دائماً در حال تغییر است و نسبت به گذشته ورزشی‌تر شده است. بازیکنانی که برای ساخت RR استفاده شدند بین سال‌های ۱۹۹۹ و ۲۰۰۶ درفت شدند، در حالی که گروه اعتبارسنجی شامل بازیکنانی بود که در سال‌های ۲۰۱۳ و ۲۰۱۴ درفت شدند و حداکثر ۱۵ سال تفاوت بین برخی بازیکنان را نشان می‌داد.

نتیجه‌گیری

این مقاله به بررسی این موضوع پرداخت که کدام متغیرها و با چه دقتی می‌توانند احتمال بقای یک بازیکن NBA در لیگ برای حداقل ۱۰ سال را پیش‌بینی کنند و در نتیجه حق دریافت طرح بازنشستگی کامل را تعیین کنند. تحلیل نشان داد که از میان ۱۹ معیار عملکردی به کار

گرفته شده، ۱۲ مورد برای این منظور از نظر آماری مهم تلقی شدند. به استثنای سن که به وضوح تأثیر منفی بر مدت زمان حرفه دارد، سایر متغیرها به طور مثبتی بر احتمال بقا برای حداقل ۱۰ سال در لیگ تأثیر گذاشتند.

با وجود استفاده از الگوریتم‌ها و تکنیک‌های پیشرفته یادگیری ماشین، ما به رگرسیون لجستیک ریج به عنوان مدل بهینه، از نظر عملکرد پیش‌بینی‌کننده، دست یافتیم. مدل نهایی به مقدار پیش‌بینی‌کننده AUC برابر با ۰.۷۹۱ رسید، در حالی که AUC تخمین زده شده هنگام آزمایش در مجموعه آموزش برابر با ۰.۷۹۴ بود، بنابراین شواهدی وجود دارد که نشان می‌دهد ما از پدیده بیش‌برازش در طول پروتکل ۱۰ CV-برابر تکراری اجتناب کردیم. استفاده از آمارهای بازیکنان درفت شده اخیرتر نه تنها ارزیابی عملکرد اضافی مدل RR را تسهیل کرد، بلکه نقاط ضعف این تحلیل را نیز نشان داد. NBA سالانه تغییر می‌کند و به همین دلیل مدل باید به طور مداوم به‌روزرسانی شود. یک سؤال جالب این است که پنجره زمانی باید چقدر باشد. آیا مدل باید از یک پنجره متحرک استفاده کند و فقط بازیکنانی را که در ۱۵ سال اخیر درفت شده‌اند، نگه دارد، یا سال‌های بیشتری لازم است؟ علاوه بر این، تحلیل تنها شامل بازیکنانی بود که در حین تحصیل در کالج درفت شده بودند. چگونه گنجاندن بازیکنان حرفه‌ای (مانند اروپایی‌ها) بر نتایج تأثیر می‌گذارد؟

از جنبه اقتصادی، پیامدهای اقتصادی مدل عمدتاً برای بازیکنان NBA از اهمیت بالایی برخوردار است. بر اساس آمارهای سال دوم خود، آن‌ها می‌توانند احتمال ماندن در لیگ برای حداقل ۱۰ سال را محاسبه کنند و بر روی بهبود کدام آمارها تمرکز کنند تا شانس خود را برای تضمین طرح بازنشستگی کامل افزایش دهند. سرمایه‌گذاری تیم‌ها در توسعه مهارت یک استراتژی اقتصادی حیاتی است، با درک پتانسیل بازیکنان برای داشتن حرفه‌های طولانی‌مدت در لیگ. این رویکرد فعال نه تنها به نفع تیم‌ها است، بلکه بازیکنان را نیز ترغیب می‌کند تا قابلیت‌های خود را بر اساس بینش‌های به دست آمده از مدل‌های آماری به طور استراتژیک افزایش دهند. بازیکنان می‌توانند از این بینش‌ها برای اولویت‌بندی حوزه‌های مهارتی مرتبط با موفقیت پایدار استفاده کنند، چه اصلاح دقت شوت‌زنی، تقویت توانایی دفاعی، یا بهینه‌سازی آمادگی جسمانی متناسب با نقاط قوت و ضعف منحصر به فرد آن‌ها، متغیرهایی که همانطور که قبلاً نشان داده شد، می‌توانند تأثیر معنی‌داری بر طول عمر حرفه‌ای داشته باشند.

علاوه بر این، بازیکنانی که احتمال آماری بالاتری برای حرفه‌های طولانی‌مدت (۱۰ سال یا بیشتر) دارند، اهرم قابل توجهی در مذاکرات قرارداد به دست می‌آورند. تیم‌ها تمایل دارند قراردادهای قابل توجه‌تر و بلندمدت‌تری را به این بازیکنان پیشنهاد دهند و بازده سرمایه‌گذاری خود را به حداکثر برسانند. علاوه بر این، همانطور که توسط بسیاری از ورزشکاران موفق اثبات شده است، افزایش چشم‌انداز موفقیت بلندمدت در لیگ اغلب معاملات اسپانسری و تأییدیه سودآوری را جذب می‌کند و برندها را با ورزشکاران تثبیت شده همسو می‌کند. برعکس، بازیکنانی که احتمال طول عمر کمتری دارند، ممکن است در تضمین معاملات تأییدیه با چالش‌هایی مواجه شوند که بر درآمد بالقوه آن‌ها خارج از زمین تأثیر می‌گذارد.

استفاده استراتژیک از مدل‌های آماری برای تقویت مهارت‌ها نه تنها ارزش بازار بازیکنان را در طول بازار آزاد افزایش می‌دهد، بلکه فرهنگ بهبود مستمر را در لیگ نیز تقویت می‌کند. تیم‌ها تمایل بیشتری به سرمایه‌گذاری در بازیکنانی دارند که چنین تعهد و پتانسیلی را برای موفقیت پایدار نشان می‌دهند، در نتیجه تحرک بازیکنان را افزایش داده و به طور بالقوه حقوق را در سراسر لیگ بالا می‌برند. علاوه بر این، با گسترش ردپای جهانی NBA، مدل‌های آماری در شناسایی استعدادها از پیشینه‌ها و مناطق مختلف نقش مهمی ایفا می‌کنند. با رمزگشایی ماهرانه شاخص‌های آماری موفقیت، استعدادیاب‌ها می‌توانند به طور مؤثر استعدادها را از بازارهای بین‌المللی

کشف و پرورش دهند و لیگ را با یک مجموعه بازیکن متنوع‌تر و رقابتی‌تر غنی سازند.

قدرداني

نویسنده بر خود لازم می‌داند از پلتفرم کوئرا (Quera) قدردانی نماید که یکی از پرسش‌های مطرح‌شده در مسابقه‌ی هوش مصنوعی در ورزش کشتی، الهام‌بخش ایده‌ی اصلی این پژوهش شد و مسیر شکل‌گیری آن را هموار ساخت. همچنین از آقای شهریار حکیمی و کانال تخصصی ایشان Sport Whisper که همواره یکی از منابع اصلی نگارنده برای پیگیری اخبار، تحلیل‌ها و روندهای روز دنیای ورزش بوده‌اند، صمیمانه سپاس‌گزاری می‌شود.

نقش محوری این منبع در تقویت علاقه به ورزش و به‌ویژه آشنایی و یادگیری عمیق‌تر درباره لیگ‌های حرفه‌ای ایالات متحده آمریکا، در شکل‌گیری زمینه ذهنی این پژوهش قابل توجه بوده است.

مراجع

- *Staw, B. M., & Hoang, H. (1995). Sunk Costs in the NBA: Why Draft Order Affects Playing Time and Survival in Professional. Administrative Science Quarterly, 40(3), 474-494.*
- *Groothuis, P. A., & Hill, J. R. (2004). Exit discrimination in the NBA: A duration analysis of career length. Economic Inquiry, 42(2), 341-349.*
- *Groothuis, P. A., & Hill, J. R. (2018). Career Duration in the NBA: Do Foreign Players Exit Early?. Journal of Sports Economics, 19(6), 873-883.*
- *Petersen, A. M., Jung, W.-S., Yang, J.-S., & Stanley, H. E. (2011). Quantitative and empirical demonstration of the Matthew effect in a study of career longevity. Proceedings of the National Academy of Sciences, 108(1), 18-23.*
- *Coates, D., & Oguntimein, B. (2010). The Length and Success of NBA Careers: Does College Production Predict Professional Outcomes?. International Journal of Sport Finance, 5(1).*
- *Barnes, J. C. (2008). Relationship of selected pre-NBA career variables to NBA players' career longevity. The Sport Journal.*
- *Miguel, C. G., M´ılan, F. J., Soares, A. L., Quinauad, R. T., K´os, L. D., Palheta, C. E., Mendes, F. G., & Carvalho, H. M. (2019). Modelling the relationship between NBA draft and the career longevity of players using generalized additive models. Revista de Psicolog´ıa del Deporte, 28(3), 0065-70.*
- *Fynn, K. D., & Sonnenschein, M. (2012). An Analysis of the Career Length of Professional Basketball Players. The Macalester Review, 2(2), Article 3.*
- *Kester, B. S., Behery, O. A., Minhas, S. V., & Hsu, W. K. (2017). Athletic performance and career longevity following anterior cruciate ligament reconstruction in the National Basketball Association. Knee Surgery, Sports Traumatology, Arthroscopy, 25, 3031-3037.*
- *Khalil, L. S., Jildeh, T. R., Tramer, J. S.,*

Abbas, M. J., Hessburg, L., Mehran, N., & Okoroha, K. R. (2020). Effect of Achilles Tendon Rupture on Player Performance and Longevity in National Basketball Association Players. Orthopaedic Journal of Sports Medicine, 8(11).

- Johns, W., Walley, K. C., Seedat, R., Thordarson, D. B., Jackson, B., & Gonzalez, T. (2021). Career Outlook and Performance of Professional Athletes After Achilles Tendon Rupture: A Systematic Review. Foot and Ankle International, 42(4), 495–509.
- Martin, C. L., Arundale, A. J., Kluzek, S., Ferguson, T., Collins, G. S., & Bullock, G. S. (2021). Characterization of Rookie Season Injury and Illness and Career Longevity among National Basketball Association Players. JAMA Network Open, 4(10).
- Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society Series B: Statistical Methodology, 67(2), 301–320.
- Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society Series B: Statistical Methodology, 58(1), 267–288.
- Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: applications to nonorthogonal problems. Technometrics, 12(1), 69–82.
- Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. Journal of Statistical Software, 33(1), 1.
- Friedman, J. H., & Stuetzle, W. (1981). Projection pursuit regression. Journal of the American Statistical Association, 76(376), 817–823.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20, 273–297.
- Breiman, L. (2001). Random Forests. Machine Learning, 45, 5–32.
- Wright, M. N., & Ziegler, A. (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. Journal of Statistical Software, 77(1), 1–17.
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. Annals of Statistics, 29(5), 1189–1232.
- Greg, R., & Developers, G. (2024). gbm: Generalized Boosted Regression Models (Version 2.1.9).
- Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). The elements of statistical learning: Data mining, inference, and prediction. Springer.
- Kursa, M. B., Jankowski, A., & Rudnicki, W. R. (2010). Boruta–a system for feature selection. Fundamenta Informaticae, 101(4), 271–285.
- Kursa, M. B., & Rudnicki, W. R. (2010).

Feature Selection with the Boruta Package. .Journal of Statistical Software, 36(11)

- Goldstein, A., Kapelner, A., Bleich, J., & Pitkin, E. (2015). Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. Journal of Computational and Graphical Statistics, .24(1), 44–65
- Székely, G. J., & Rizzo, M. L. (2004). Testing for equal distributions in high dimension. .InterStat, 5(16.10), 1249–1272
- James, G. (1954). Tests of linear hypotheses in univariate and multivariate analysis when the ratios of the population variances are unknown. Biometrika, 41(1/2), .19–43
- R Core Team. (2023). R: A language and environment for statistical computing. R .Foundation for Statistical Computing
- Psathas, A., Rallatou, D., & Tsagris, M. (2023). Skin tone of nba players and performance statistics. is there a relationship?. Communications in Statistics: Case Studies, Data Analysis and .Applications, 9(3), 234–251
- Number Analytics. (2025). Top 4 survival .analysis tactics in sports analytics
- White Rose eTheses Online. (2025). Player performance, salary and survival analysis in .the NBA
- FUPRESS. (2023). A survival analysis study to discover which skills determine a higher .scoring in basketball
- MDPI. (2023). Injury patterns and impact on performance in the NBA league using .sports analytics
- MDPI. (2024). Unsupervised learning in NBA injury recovery: Advanced data mining to decode recovery durations and economic .impacts
- MDPI. (2023). Machine learning at the ....service of survival analysis: Predictions
- MDPI. (2019). Mortality of NBA players: Risk factors and comparison with the .general US population
- MDPI. (n.d.). Applications of machine learning in sports medicine, physical activity, .posture, and rehabilitation
- PubMed Central. (2025). The application of artificial intelligence techniques in predicting game outcomes of professional .basketball league: A systematic review
- PubMed Central. (2024). Diagnostic applications of AI in sports: A comprehensive .review of injury risk prediction methods
- PubMed Central. (2024). Performance and healthcare analysis in elite sports teams using artificial intelligence: A scoping .review
- PubMed Central. (n.d.). Artificial

- *intelligence and machine learning approaches in sports: Concepts, applications, challenges, and future perspectives.*
- *PubMed Central. (2021). Basketball sports injury prediction model based on the grey theory neural network.*
- *PubMed Central. (2022). Return to performance following severe ankle, knee, and hip injuries in National Basketball Association players.*
- *PubMed Central. (2025). Enhancing game outcome prediction in the Chinese basketball league through a machine learning framework based on performance data.*
- *PubMed Central. (n.d.). Artificial intelligence in health and sport sciences: Promise, progress, and prudence.*
- *PubMed Central. (2018). It's a hard-knock life: Game load, fatigue, and injury risk in the ....*
- *Oyelabs. (2025). AI in sports – Revolutionizing training and performance.*
- *ResearchGate. (2014). Application of survival models in sports injury prevention research: A systematic review.*
- *ResearchGate. (2022). Machine learning for predicting lower extremity muscle strain in National Basketball Association athletes.*
- *ResearchGate. (2023). The time course of injury-risk after return-to-play in professional football.*
- *ResearchGate. (2023). Predicting full retirement attainment of NBA players.*
- *ResearchGate. (2025). Leveraging machine learning for accurate prediction of NBA player salaries.*
- *Beyond Cox models: Assessing the performance of machine-learning methods in non-proportional hazards and non-linear survival analysis. (2025). arXiv:2504.17568.*
- *MOTOR: A time-to-event foundation model for structured medical records. (2023). arXiv:2301.03150.*
- *Expected points above average: A novel NBA player metric based on Bayesian hierarchical modeling. (2025). arXiv:2405.10453v2.*
- *NHSJS. (2025). Computational analysis of NBA players with machine and deep learning.*
- *SciTePress. (2023). A comparison of machine learning algorithms for National Basketball Association (NBA) Most Valuable Player (MVP) vote share prediction.*

پاورقی‌ها:

پانویس‌ها: