https://doi.org/10.22034/aa.2026.2092485.1038

برآورد رتبه نهایی تیم‌ها در لیگ ناتمام فوتبال ایران با استفاده از مدل بردلی-تری-دیویدسون

ایمان سربیشه[۱] و آزاده ساکی[۲]* و مرتضی ادیبیان[۱]

[۱] کمیته تحقیقات دانشجویی، دانشگاه علوم پزشکی مشهد، مشهد، ایران

[۲] دانشیار آمار زیستی، گروه آمارزیستی، دانشکده بهداشت، دانشگاه علوم پزشکی مشهد، مشهد، ایران

تاریخ دریافت: ۱۴۰۵/۰۴/۰۳
تاریخ پذیرش: ۱۴۰۵/۰۸/۲۲

چکیده

توقف پیش از موعد لیگ‌های ورزشی به‌دلایلی از جمله بحران‌های بهداشتی، حوادث طبیعی و جنگ، فرایند رتبه‌بندی نهایی تیم‌ها را با چالش مواجه می‌کند. تکیه بر جدول کنونی مسابقات لزوماً منعکس‌کننده توان واقعی تیم‌ها نیست، زیرا تعداد و دشواری دیدارهای باقی‌مانده برای هر تیم متفاوت است. این مطالعه با هدف ارائه چارچوبی آماری مبتنی بر مدل بردلی-تری-دیویدسون برای برآورد توان نسبی تیم‌ها، پیش‌بینی نتایج مسابقات برگزارنشده و تعیین رتبه نهایی لیگ برتر فوتبال خلیج‌فارس ایران در فصل ۱۴۰۵-۱۴۰۴ انجام شد. داده‌های ۱۷۸ مسابقه برگزارشده لیگ برتر فوتبال ایران جمع‌آوری و مدل بردلی-تری-دیویدسون با درنظرگرفتن پارامتر گرایش به تساوی و مزیت میزبانی با روش درستنمایی بیشینه برازش شد. پس از برآورد توان نسبی تیم‌ها، احتمال برد میزبان، تساوی و برد مهمان برای ۶۲ مسابقه باقی‌مانده محاسبه شد. سپس امتیاز مورد انتظار هر تیم برآورد و رتبه‌بندی نهایی لیگ تعیین گردید. پارامتر گرایش به تساوی ۳۲۲.۱ و مزیت میزبانی ۵۳۸.۱ برآورد شد. تیم‌های تراکتور تبریز، استقلال تهران و چادرملو اردکان بالاترین توان نسبی را داشتند. نتایج نشان داد که استقلال تهران با ۸۱.۵۳ امتیاز در رتبه نخست قرار می‌گیرد و تراکتور تبریز و سپاهان اصفهان در رتبه‌های دوم و سوم قرار دارند.

واژه‌های کلیدی: مدل بردلی-تری-دیویدسون، لیگ فوتبال ایران، رتبه‌بندی تیم‌ها، پیش‌بینی مسابقات، مزیت میزبانی
رده‌بندی ریاضی: ۶۲F۰۷، ۶۲J۱۵.

۱ مقدمه

مسابقات ورزشی لیگ محور معمولاً بر اساس نتایج کامل یک فصل، تیم قهرمان، تیم‌های صعودکننده به رقابت‌های بین‌المللی و تیم‌های سقوط‌کننده به لیگ های پایین را تعیین می‌کنند. با این حال، در برخی شرایط نظیر بحران‌های بهداشتی، حوادث طبیعی، مسائل امنیتی و جنگ ممکن است فصل مسابقات پیش از برگزاری تمامی دیدارهای برنامه‌ریزی‌شده متوقف شود. در چنین شرایطی، تصمیم‌گیری درباره رتبه نهایی تیم‌ها و تخصیص سهمیه‌های بین‌المللی به یک چالش مهم ورزشی، مدیریتی و آماری تبدیل می‌شود در صورت ناتمام ماندن مسابقات ورزشی، رتبه‌بندی مبتنی بر جدول فعلی ممکن است تصویری ناقص از توان واقعی تیم‌ها ارائه دهد، زیرا برخی تیم‌ها ممکن است برنامه مسابقات آسان‌تر یا دشوارتری در دیدارهای باقی‌مانده داشته باشند. از این‌رو، استفاده از روش‌های آماری که بتوانند توان نسبی تیم‌ها را

*نویسنده مسئول مقاله: sakia@mums.ac.ir

از نتایج موجود استخراج کرده و نتایج احتمالی مسابقات برگزارنشده را برآورد کنند، اهمیت ویژه‌ای پیدا می‌کند. مدل بردلی–تری یکی از شناخته‌شده‌ترین مدل‌های آماری برای تحلیل داده‌های مقایسه‌ای دوتایی است [۱]. این مدل فقط دو نتیجه برد و باخت را در نظر می‌گیرد و امکان وقوع تساوی را لحاظ نمی‌کند از آنجا که در بسیاری از رقابت‌های ورزشی به ویژه فوتبال تساوی یکی از نتایج رایج مسابقات است مدل بردلی–تری–دیویدسون به عنوان توسعه‌ای از مدل بردلی–تری معرفی شد [۲]. سادگی تفسیر، امکان برآورد قدرت نهفته تیم‌ها و قابلیت استفاده در پیش‌بینی مسابقات آینده موجب شده است که مدل بردلی–تری–دیویدسون و توسعه‌های آن در رتبه‌بندی تیم‌های ورزشی، تحلیل رقابت‌ها و پیش‌بینی نتایج مسابقات مورد توجه قرار گیرند . همچنین مطالعات متعددی از نسخه‌های توسعه یافته این مدل برای بررسی اثر میزبانی تغییرات زمانی و شرایط ویژه مسابقات استفاده کرده‌اند [۳،۴،۵] کاربرد مدل‌های آماری و یادگیری ماشین در پیش‌بینی مسابقات بین‌المللی فوتبال، مانند جام جهانی فیفا یا لیگ قهرمانان اروپا، همواره مورد توجه بسیاری از تحلیلگران بوده است. از منظر آماری، نتیجه یک مسابقه فوتبال را می‌توان با استفاده از دو رویکرد متفاوت پیش‌بینی کرد. رویکرد اول مبتنی بر نتیجه است که به طور مستقیم نتیجه مسابقه را پیش‌بینی می‌کند و فرایند سه حالته (برد میزبان، تساوی یا برد مهمان) را با استفاده از رگرسیون لجستیک یا رگرسیون چندجمله‌ای مدل‌سازی می‌کند[۶]رویکرد دوم مبتنی بر گل است که گل‌های زده و خورده در هر مسابقه را با مدل‌سازی متغیرهای شمارشی، معمولاً با استفاده از رگرسیون پواسون مدل‌سازی می‌کند و سپس با تجمیع احتمالات برآوردشده، فرایند سه حالته مسابقه را استخراج می‌کند[۷]، برای مقایسه جامع‌تر از رویکردهای مبتنی بر گل و مبتنی بر نتیجه، می‌توان به مقاله اگیدی و تورلی (۲۰۲۱) مراجعه کرد[۸] لیگ برتر فوتبال ایران در فصل ۱۴۰۵-۱۴۰۴ به دلیل شرایط ناشی از جنگ قبل از برگزاری تمامی مسابقات متوقف شد. در زمان توقف مسابقات، همه تیم‌ها تعداد یکسانی مسابقه انجام نداده بودند و ۶۲ مسابقه برگزار نشده باقی مانده بود در چنین شرایطی، استفاده از رتبه‌بندی مبتنی بر امتیازات فعلی نمی‌تواند به طور کامل بیانگر توان نسبی تیم‌ها و جایگاه واقعی آن‌ها در پایان فصل باشد. در این پژوهش از مدل بردلی–تری–دیویدسون با در نظر گرفتن اثر میزبانی و گرایش به تساوی برای برآوری توان نسبی تیم‌های لیگ برتر فوتبال ایران استفاده شده است. در نهایت با استفاده از احتمالات پیش‌بینی شده برای مسابقات برگزار نشده، امتیاز مورد انتظار هر تیم محاسبه و رتبه نهایی تیم‌ها برآورد شده است. هدف این پژوهش ارائه یک چارچوب آماری شفاف برای تعیین رتبه‌های نهایی لیگ برتر فوتبال ایران در شرایط ناتمام ماندن مسابقات است نتایج این مطالعه می‌تواند در تصمیم گیری‌های سازمان لیگ، تعیین تیم قهرمان، تخصیص سهمیه‌های بین المللی و مشخص کردن تیم‌های سقوط کننده مورد استفاده قرار گیرد.

۲ روش پژوهش

۱.۲ مدل بردلی–تری

مدل بردلی تری [1] یکی از شناخته‌شده‌ترین مدل‌های مقایسه زوجی است که تنها دو نتیجه برد و باخت را در نظر می‌گیرد و احتمال تساوی را مدل سازی نمی‌کند [۱]. اگر توان نسبی تیم‌های i و j به ترتیب با π_i و π_j نمایش داده شود، احتمال پیروزی تیم i بر تیم j برابر است با

$$P(i > j) = \frac{\pi_i}{\pi_i + \pi_j}. \quad (۱)$$

در واقع داریم

$$\log\left(\frac{P(i > j)}{P(j > i)}\right) = \beta_i - \beta_j. \quad (۲)$$

$$\pi_i = \exp(\beta_i), \qquad \pi_j = \exp(\beta_j). \quad (۳)$$

مزیت اصلی این مدل سادگی آن است، اما نادیده گرفتن نتیجه تساوی یکی از محدودیت‌های مهم آن در تحلیل مسابقات فوتبال محسوب می‌شود.

۲.۲ مدل بردلی–تری–دیویدسون

از آنجا که در مسابقات فوتبال نتیجه تساوی نیز رخ می دهد، مدل بردلی–تری–دیویدسون [2] به عنوان توسعه‌ای از مدل بردلی–تری معرفی شد [۲]. در این مدل یک پارامتر گرایش به تساوی با نماد νوارد مدل می‌شود. در مدل بردلی–تری–دیویدسون احتمال برد تیم میزبان، تساوی و برد تیم مهمان به صورت زیر تعریف می‌شود:

$$P(i) = \frac{\alpha\pi_i}{\alpha\pi_i + \pi_j + \nu\sqrt{\alpha\pi_i\pi_j}}. \quad (۴)$$

$$P(j) = \frac{\pi_j}{\alpha\pi_i + \pi_j + \nu\sqrt{\alpha\pi_i\pi_j}}. \quad (۵)$$

[1] Bradley-Terry
[2] Bradley–Terry–Davidson

$$P(D) = \frac{\nu\sqrt{\alpha\pi_i\pi_j}}{\alpha\pi_i + \pi_j + \nu\sqrt{\alpha\pi_i\pi_j}}. \quad (۶)$$

که در آن π_i توان نسبی تیم میزبان، π_j توان نسبی تیم مهمان، ν پارامتر گرایش به تساوی، αپارامتر مزیت میزبانی، $P(D)$ احتمال تساوی مسابقه است. در صورتی که مقدار α بزرگتر از یک باشد، نشان‌دهنده وجود مزیت میزبانی است و همچنین هر چه مقدار ν بزرگتر باشد، احتمال وقوع تساوی افزایش می‌یابد.

۳.۲ برآورد پارامترهای مدل بردلی-تری-دیویدسون

پارامترهای مدل شامل توان نسبی تیم‌ها، پارامتر مزیت میزبانی و پارامتر گرایش به تساوی با استفاده از روش درستنمایی بیشینه برآورد شدند. در این مدل، به دلیل آنکه پارامترهای توانایی تیم‌ها به‌صورت نسبی تعریف می‌شوند، مدل از قابلیت شناسایی‌پذیری کامل برخوردار نیست؛ به‌این معنا که اگر یک مقدار ثابت به تمامی پارامترهای توانایی نسبی اضافه شود، تغییری در لگاریتم شانس‌ها رخ نمی‌دهد. برای جلوگیری از این مشکل، پارامتر قدرت نسبی یکی از تیم‌ها به عنوان تیم مرجع برابر یک در نظر گرفته شد و توانایی سایر تیم‌ها نسبت به آن برآورد شدند. تابع درستنمایی مدل بر اساس نتایج مشاهده شده تمامی مسابقات برگزارشده تشکیل شد و مقادیر پارامترها به گونه‌ای تعیین شدند که مقدار تابع درستنمایی بیشینه شود.

۴.۲ محاسبه امتیازات مورد انتظار

پس از برآورد پارامترهای مدل بردلی–تری–دیویدسون، احتمال برد میزبان، احتمال تساوی و احتمال برد مهمان برای تمامی ۶۲ مسابقه باقی‌مانده محاسبه شد. و با توجه به سیستم امتیاز دهی در یک مسابقه فوتبال(برد برابر ۳ امتیاز، تساوی برابر ۱ امتیاز و باخت برابر ۰ امتیاز) امتیاز مورد انتظار برای هر تیم از یک مسابقه با رابطه زیر محاسبه شد:

$$E(\text{امتیاز}) = ۳ \times P(\text{برد}) + ۱ \times P(\text{مساوی}) + ۰ \times P(\text{باخت}). \quad (۷)$$

سپس مجموع امتیازات مورد انتظار حاصل از مسابقات باقی‌مانده به امتیازات واقعی کسب‌شده هر تیم اضافه شد و رتبه نهایی تیم‌ها محاسبه گردید.

۵.۲ اعتبارسنجی مدل

برای ارزیابی توان پیش‌بینی مدل ، داده‌های فصل کامل ۱۴۰۴-۱۴۰۳ لیگ برتر فوتبال ایران مورد استفاده قرار گرفت. داده‌ها به دو بخش آموزش (۱۷۸ بازی اول) و آزمون (۶۲ بازی باقی‌مانده) تقسیم شدند مدل بر روی داده‌های آموزش برازش داده شد و سپس نتایج مسابقات موجود در مجموعه آزمون، پیش‌بینی شدند. علاوه بر این عملکرد مدل با استفاده از ماتریس اغتشاش [۳]، دقت طبقه‌بندی، شاخص Loss Log [۹]. و شاخص Score Brier [۱۰] ارزیابی شد.

۶.۲ شاخص Loss Log

برای ارزیابی کیفیت احتمالات پیش‌بینی‌شده توسط مدل، از شاخص Log-Loss استفاده شد. این شاخص میزان فاصله بین احتمالات پیش‌بینی‌شده و نتایج واقعی را اندازه‌گیری می‌کند و به صورت زیر تعریف می‌شود:

$$\text{LogLoss} = -\frac{۱}{N}\sum_{i=۱}^{N}\sum_{k=۱}^{K} y_{ik}\log(p_{ik}), \quad (۸)$$

که در آن:

- N تعداد مسابقات،
- K تعداد حالات ممکن نتیجه مسابقه (در اینجا سه حالت برد میزبان،مساوی و برد مهمان)،
- y_{ik} متغیر شاخص وقوع نتیجه واقعی (اگر نتیجه واقعی با حالت ممکن kیکی باشد مقدار آن یک و در غیر این صورت صفر است)
- p_{ik} احتمال پیش‌بینی‌شده برای نتیجه k است

در این شاخص، مقادیر کوچکتر نشان‌دهنده عملکرد بهتر مدل هستند. این شاخص نسبت به پیش‌بینی‌های بسیار اشتباه حساسیت بالایی دارد و به همین دلیل یکی از متداول‌ترین معیارهای ارزیابی مدل‌های پیش‌بینی احتمالاتی محسوب می‌شود.

۷.۲ شاخص Score Brier

شاخصScore Brier میانگین توان دوم اختلاف بین احتمال پیش‌بینی‌شده و نتیجه واقعی را اندازه‌گیری می‌کند و به صورت زیر تعریف می‌شود:

[3] Confusion matrix

$$\text{Brier} = \frac{۱}{N}\sum_{i=۱}^{N}\sum_{k=۱}^{K}(p_{ik} - y_{ik})^۲. \quad (۹)$$

در این رابطه:

p_{ik} احتمال پیش‌بینی‌شده، y_{ik} نتیجه واقعی مشاهده‌شده است.

این معیار فاصله بین احتمالات پیش‌بینی‌شده و نتایج واقعی را اندازه‌گیری می‌کند و به‌طور هم‌زمان دو ویژگی مهم مدل یعنی دقت و کالیبره شدن را مورد ارزیابی قرار می‌دهد. برخلاف شاخص Log Loss، شاخص Score Brier نسبت به پیش‌بینی‌های بسیار اشتباه، حساسیت کمتری دارد و تصویری متعادل‌تر از عملکرد کلی مدل ارائه می‌دهد. مقدار صفر نشان‌دهنده پیش‌بینی کامل است و هرچه مقدار این شاخص افزایش یابد، کیفیت پیش‌بینی کاهش پیدا می‌کند. در این مطالعه عملکرد مدل‌های مختلف بر اساس مقادیر Loss Log و Brier Score مقایسه شدند. از آنجا که هر دو شاخص بر مبنای احتمالات پیش‌بینی‌شده محاسبه می‌شوند، مدلی که کمترین مقدار را در این دو معیار کسب کند، به‌عنوان مدل دارای بهترین توان پیش‌بینی در نظر گرفته می‌شود.

مقادیر کمتر برای هر دو شاخص نشان‌دهنده عملکرد بهتر مدل است و امکان مقایسه مستقیم بین مدل‌های با ساختارهای متفاوت را فراهم می‌سازد.

۸.۲ مقایسه عملکرد مدل پژوهش با مدل‌های مشابه

به‌منظور مقایسه عملکرد مدل‌های مختلف در مقایسه با مدل بردلی–تری–دیویدسون، از چهار مدل آماری شامل مدل ELO [۱۱] ، مدل بردلی–تری پایه، مدل پواسون [۱۲] و مدل دیکسون–کولز [۱۳][۴] استفاده شد. این مدل‌ها از نظر مبانی نظری، ساختار آماری و نوع اطلاعات مورد استفاده با یکدیگر تفاوت دارند و هر یک جنبه خاصی از عملکرد تیم‌ها را مدل‌سازی می‌کنند.

۱.۸.۲ مدل ELO

مدل ELO یک سیستم رتبه‌بندی پویا است که در ابتدا برای رتبه‌بندی بازیکنان شطرنج طراحی شد و بعدها در تحلیل مسابقات ورزشی به ویژه فوتبال کاربرد گسترده‌ای پیدا کرد [۱۱]. در این مدل برای هر تیم یک امتیاز قدرت (Rating) در نظر گرفته می‌شود که بیانگر توان نسبی آن تیم در مقایسه با سایر تیم‌ها است.

احتمال پیروزی تیم i در برابر تیم j از رابطه زیر محاسبه می‌شود:

$$P(i) = \frac{۱}{۱ + ۱۰^{(R_j - R_i)/۴۰۰}}, \quad (۱۰)$$

که در آن R_i و R_j به ترتیب رتبه ELO دو تیم هستند. پس از پایان هر مسابقه، رتبه تیم‌ها به صورت زیر به‌روزرسانی می‌شود:

$$R_i^{new} = R_i^{old} + K(S_i - E_i), \quad (۱۱)$$

که در آن:

- S_i نتیجه واقعی مسابقه،
- E_i نتیجه مورد انتظار،
- K ضریب به‌روزرسانی

است. در این مدل معمولا برد با کد ۱، مساوی با کد ۰.۵ و باخت با کد صفر در نظر گرفته می شود.

۲.۸.۲ مدل پواسون

مدل پواسون [۵] یکی از متداول‌ترین مدل‌های آماری برای تحلیل و پیش‌بینی نتایج مسابقات فوتبال است [۱۲].

در این مدل فرض می‌شود تعداد گل‌های به ثمر رسیده توسط هر تیم از توزیع پواسون پیروی می‌کند:

$$Y_{ij} \sim Poi(\lambda_{ij}), \quad (۱۲)$$

که در آن:

- Y_{ij} تعداد گل‌های تیم i در برابر تیم j،
- λ_{ij} میانگین مورد انتظار گل‌های تیم i

است.

برای مدل‌سازی میانگین تعداد گل‌ها از رابطه زیر استفاده می‌شود:

$$\log(\lambda_{ij}) = \mu + A_i - D_j + H \times I_{home}, \quad (۱۳)$$

که در آن:

- μ میانگین کلی گل‌زنی لیگ،

[4] Dixon-Coles
[5] Poisson

- A_i قدرت تهاجمی تیم،
- D_j قدرت دفاعی تیم،
- H اثر میزبانی،
- I_{home} متغیر نشان‌دهنده میزبانی

است.

پس از برآورد پارامترها، احتمال تمامی نتایج ممکن مسابقه محاسبه شده و از آن احتمال برد، تساوی و شکست هر تیم استخراج می‌شود.

۳.۸.۲ مدل دیکسون–کولز

این مدل نسخه اصلاح‌شده‌ای از مدل پواسون است که برای بهبود دقت پیش‌بینی نتایج کم‌گل در فوتبال توسعه داده شده است [۱۳]. یکی از محدودیت‌های مدل پواسون کلاسیک این است که همبستگی بین تعداد گل‌های دو تیم را در نظر نمی‌گیرد، در حالی که در عمل، رخدادهایی مانند صفر–صفر یا یک–یک با احتمال بیشتری نسبت به پیش‌بینی مستقل پواسونی رخ می‌دهند.

در این مدل، احتمال مشترک تعداد گل‌های دو تیم به صورت زیر اصلاح می‌شود:

$$P(X_i = x, Y_j = y) = \tau(x,y)\, Poi(\lambda_i, x)\, Poi(\lambda_j, y), \quad (۱۴)$$

که در آن $\tau(x,y)$ یک تابع اصلاحی برای نتایج کم‌گل است.

این تابع به صورت زیر تعریف می‌شود:

$$\tau(x,y) = \begin{cases} ۱ - \lambda_i \lambda_j \rho & \text{if} \quad x = ۰, y = ۰ \\ ۱ + \lambda_i \rho & \text{if} \quad x = ۰, y = ۱ \\ ۱ + \lambda_j \rho & \text{if} \quad x = ۱, y = ۰ \\ ۱ - \rho & \text{if} \quad x = ۱, y = ۱ \\ ۱ & \text{otherwise} \end{cases} \quad (۱۵)$$

که در آن λ_i میانگین مورد انتظار گل‌های تیم i در مسابقه و λ_j میانگین مورد انتظار گل‌های تیم j در همان مسابقه و ρ پارامتری است که میزان وابستگی بین گل‌های دو تیم را کنترل می‌کند. مقدار ρ معمولاً از داده‌ها برآورد می‌شود و می‌تواند اثر همبستگی منفی یا مثبت بین گل‌های دو تیم را مدل‌سازی کند.

۹.۲ چارچوب مقایسه مدل‌ها

برای مقایسه عملکرد مدل‌های مورد بررسی، همه مدل‌ها بر اساس یک چارچوب مشترک ارزیابی شدند. در این چارچوب، خروجی هر مدل برای سه حالت برد، مساوی و باخت استانداردسازی شد. سپس دو معیار Log-Loss و Score Brier برای تمامی مدل‌ها محاسبه شدند تا امکان مقایسه منصفانه بین روش‌هایی با ساختارهای آماری متفاوت فراهم شود.

به طور کلی، مدل‌هایی که ساختار احتمالاتی دقیق‌تری دارند مانند پواسن و دیکسون–کولز انتظار می‌رود عملکرد بهتری در معیارLog Loss داشته باشند در حالی که مدل‌های ساده‌تر مانند ELO و بردلی–تری ممکن است در برخی شرایط از نظر پایداری عملکرد قابل قبول اما از نظر دقت احتمالاتی ضعیف‌تر باشند.

۱۰.۲ داده‌های مورد استفاده

اطلاعات مسابقات لیگ برتر فوتبال ایران درفصل ناتمام ۱۴۰۵–۱۴۰۴ و فصل کامل ۱۴۰۴–۱۴۰۳ از سایت رسمی ورزش سه استخراج شد [۱۴].از زمان شروع مسابقات تا تاریخ ۸ اسفند ۱۴۰۴ از مجموع ۲۴۰ مسابقه برنامه‌ریزی‌شده لیگ، تعداد ۱۷۸ مسابقه برگزار شده بود و ۶۲ مسابقه باقی مانده بود. در زمان توقف مسابقات، تیم‌های مختلف بین ۲۱ تا ۲۳ مسابقه انجام داده بودند. برای هر مسابقه، اطلاعات زیر جمع‌آوری شد:

تیم میزبان و تیم مهمان،

تعداد گل‌های زده شده توسط هر تیم،

مکان برگزاری مسابقه (خانه یا خارج از خانه)،

نتیجه نهایی مسابقه (برد،مساوی یا باخت)

۱.۱۰.۲ تقسیم داده‌ها:

برای ارزیابی عملکرد مدل‌ها بر روی داده‌های فصل کامل ۱۴۰۴–۱۴۰۳، داده‌ها به دو بخش آموزش و آزمون تقسیم شدند به این صورت که داده‌های مربوط به ۱۷۸ مسابقه ابتدای فصل به عنوان داده‌های آموزش و داده‌های مربوط به ۶۲ مسابقه باقیمانده به عنوان داده‌های آزمون در نظر گرفته شد تا شرایط پیش‌بینی واقعی‌تر شبیه‌سازی شود. این سناریو شبیه به شرایط واقعی در فصل ناقص ۱۴۰۵–۱۴۰۴ است.

۳ نتایج

۱.۳ نتایج مربوط به اعتبار سنجی مدل‌ها:

مدل‌ها با استفاده از داده‌های بخش آموزش برازش داده شدند و سپس نتایج مسابقات بخش آزمون پیش‌بینی گردید. ماتریس اغتشاش و دقت طبقه‌بندی (Accuracy) در جدول ۴ گزارش شدند این نتایج نشان می‌دهد که مدل بردلی–تری–دیویدسون در پیش‌بینی نتایج برد و باخت عملکرد قابل قبولی داشته است ولی همانند بسیاری از مدل‌های پیش‌بینی نتایج فوتبال، پیش‌بینی مسابقات مساوی همچنان چالش‌برانگیز باقی می‌ماند. شاخص Log-Loss و Score Brier برای مدل بردلی–تری–دیویدسون به ترتیب ۰/۶۷۱ و ۰/۲۳۸ محاسبه شد. همچنین این شاخص‌ها برای سایر مدل‌ها در جدول ۵ گزارش شده است. و منحنی راک برای هر مدل نیز در شکل یک رسم شده است.

۲.۳ نتایج مربوط به داده‌های فصل ناتمام ۱۴۰۵-۱۴۰۴:

از مجموع ۲۴۰ بازی کامل لیگ برتر فوتبال ایران (هر تیم ۳۰ بازی) تا ۸ اسفند ۱۴۰۴ تعداد ۱۷۸ بازی انجام شده است تیم‌های مختلف از حداقل ۲۱ بازی تا حداکثر ۲۳ بازی را انجام داده اند که از این تعداد ۶۵ بازی (۳۶/۵ درصد–۶۵ امتیاز) مساوی، ۴۶ بازی (۲۸/۸۴درصد–۱۳۸ امتیاز) برنده مهمان و ۶۷ بازی (۳۷/۶۴ درصد–۲۰۱ امتیاز) برنده میزبان بوده است. در داده‌های مطالعه ما پارامتر گرایش به مساوی در مدل بردلی–تری–دیویدسون برابر با $\nu = ۱/۳۲۲$ (95% CI: 1/287,1/356) و پارامتر مزیت میزبانی برابر $\alpha = ۱/۵۳۸$ (95% CI: 1/474,1/603) برآورد شد.

تیم گل‌گهر سیرجان به عنوان تیم مرجع در نظر گرفته شد یعنی توانایی نسبی سایر تیم‌ها نسبت به این تیم سنجیده شده است. برآورد پارامتر توانایی نسبی هر تیم در جدول ۱ گزارش شده است بر این اساس، تراکتور ، استقلال و چادر ملو اردکان بالاترین سطح توانایی را در میان تیم‌های لیگ را داشتند، در حالی که مس رفسنجان ، ذوب‌آهن و خیبر خرم آباد پایین‌ترین مقادیر توانایی نسبی را به خود اختصاص دادند، توان نسبی برآورد شده توسط مدل لزوماً معادل رتبه فعلی جدول لیگ نیست. مقدار احتمال برد میزبان، احتمال تساوی و احتمال برد مهمان برای ۶۲ بازی باقیمانده توسط مدل بردلی–تری–دیویدسون با استفاده از توان نسبی هر تیم و پارامتر گرایش به تساوی و پارامتر اثر میزبانی محاسبه شده است و با توجه به اینکه هر بازی حداکثر سه امتیاز دارد مقدار امتیاز مورد انتظار برای هر تیم در هر مسابقه نیز محاسبه شد (جدول۲). رتبه نهایی تیم‌ها با جمع امتیازات حاصل از بازی های برگزار شده و امتیازات مورد انتظار براورد شده محاسبه و در جدول ۳ گزارش شده است.

پذیرفته شده

جدول ۱: پارامترهای برآورد شده قدرت نسبی هر تیم با استفاده از مدل بردلی–تری–دیویدسون با در نظر گرفتن پارامتر گرایش به تساوی در لیگ فوتبال ایران و شرایط اثر میزبانی

رتبه	نام تیم	π_i	$\beta_i \pm SE$
۱	تراکتور تبریز	۱٫۴۵۵	۰٫۳۷۵ ± ۰٫۶۰۵
۲	استقلال تهران	۱٫۳۳۲	۰٫۲۸۶ ± ۰٫۶۳۹
۳	چادرملو اردکان	۱٫۰۷۸	۰٫۰۷۵ ± ۰٫۵۹۱
۴	گل گهر سیرجان	۱٫۰۰۰	۰ ± NA
۵	سپاهان اصفهان	۰٫۹۵۳	−۰٫۰۴۷ ± ۰٫۵۸۵
۶	ملوان بندر انزلی	۰٫۸۶۲	−۰٫۱۴۸ ± ۰٫۵۸۳
۷	فولاد خوزستان	۰٫۷۶۲	−۰٫۲۷۰ ± ۰٫۵۸۶
۸	استقلال خوزستان	۰٫۶۳۰	−۰٫۴۶۱ ± ۰٫۶۲۰
۹	شمس آذر قزوین	۰٫۵۴۳	−۰٫۶۱۰ ± ۰٫۶۰۰
۱۰	پرسپولیس تهران	۰٫۵۲۸	−۰٫۶۳۸ ± ۰٫۵۸۶
۱۱	پیکان تهران	۰٫۵۲۲	−۰٫۶۴۹ ± ۰٫۶۱۲
۱۲	فجر شهید سپاسی شیراز	۰٫۵۱۶	−۰٫۶۶۰ ± ۰٫۵۸۰
۱۳	آلومینیوم اراک	۰٫۳۵۴	−۱٫۰۳۸ ± ۰٫۶۰۷
۱۴	خیبر خرم‌آباد	۰٫۳۴۹	−۱٫۰۵۲ ± ۰٫۶۰۷
۱۵	ذوب‌آهن اصفهان	۰٫۱۶۵	−۱٫۷۹۷ ± ۰٫۷۰۲
۱۶	مس رفسنجان	۰٫۰۷۲	−۲٫۶۲۵ ± ۰٫۷۹۵

SE / NA: Not Available :انحراف معیار

جدول ۲: برآورد احتمال برد و مساوی تیم‌ها با استفاده از مدل بردلی–تری–دیویدسون و محاسبه امتیازات مورد انتظار برای تیم میزبان و مهمان در مسابقات باقی‌مانده لیگ برتر فوتبال ایران

امتیاز مورد انتظار تیم میزبان	امتیاز مورد انتظار تیم مهمان	احتمال برد تیم میزبان	احتمال تساوی	احتمال برد تیم مهمان	تیم مهمان	تیم میزبان
۸۲۰۰۰	۱٫۸۲۰	۰٫۱۵۳	۰٫۳۶۱	۰٫۴۸۶	آلومینیوم اراک	مس رفسنجان

پذیرفته شده

ادامه جدول

امتیاز مورد انتظار تیم میزبان	مورد امتیاز انتظار تیم مهمان	احتمال برد تیم میزبان	احتمال تساوی	احتمال برد تیم مهمان	تیم مهمان	تیم میزبان
۹۸۵.۰	۱٫۶۳۲	۰٫۲۰۱	۰٫۳۸۳	۰٫۴۱۶	پرسپولیس	ذوب آهن
۳۵۴.۱	۱٫۲۴۸	۰٫۳۱۹	۰٫۳۹۸	۰٫۲۸۳	ملوان	استقلال خوزستان
۹۵۵.۱	۰٫۷۰۶	۰٫۵۳۸	۰٫۳۳۹	۰٫۱۲۲	خیبر خرم آباد	گل گهر
۴۸۲.۲	۰٫۳۰۸	۰٫۷۵۸	۰٫۲۰۹	۰٫۰۳۳	مس رفسنجان	چادرملو اردکان
۰۴۶.۱	۱٫۵۶۶	۰٫۲۱۹	۰٫۳۸۸	۰٫۳۹۳	تراکتور	پرسپولیس
۴۹۲.۱	۱٫۱۱۵	۰٫۳۶۶	۰٫۳۹۳	۰٫۲۴۱	پیکان	فجر سپاسی
۷۴۹.۱	۰٫۸۸۱	۰٫۴۶۰	۰٫۳۷۰	۰٫۱۷۰	شمس آذر قزوین	سپاهان
۹۱۹.۰	۱٫۷۰۶	۰٫۱۸۱	۰٫۳۷۵	۰٫۴۴۴	استقلال	آلومینیوم اراک
۵۸۳.۱	۱٫۰۳۰	۰٫۳۹۹	۰٫۳۸۷	۰٫۲۱۴	استقلال خوزستان	فولاد خوزستان
۱۷۸.۲	۰٫۵۲۹	۰٫۶۲۸	۰٫۲۹۴	۰٫۰۷۸	ذوب آهن	ملوان
۷۸۵.۱	۰٫۸۴۹	۰٫۴۷۳	۰٫۳۶۵	۰٫۱۶۱	فولاد	تراکتور تبریز
۷۶۱.۰	۱٫۸۸۸	۰٫۱۳۷	۰٫۳۵۰	۰٫۵۱۳	سپاهان	ذوب آهن
۵۹۳.۱	۱٫۰۲۱	۰٫۴۰۲	۰٫۳۸۶	۰٫۲۱۲	چادرملو اردکان	استقلال تهران
۶۸۹.۰	۱٫۹۷۶	۰٫۱۱۸	۰٫۳۳۶	۰٫۵۴۷	فجر سپاسی	مس رفسنجان
۲۸۷.۱	۱٫۳۱۵	۰٫۲۹۶	۰٫۳۹۸	۰٫۳۰۶	ملوان	شمس آذر قزوین
۵۷۷.۱	۱٫۰۳۵	۰٫۳۹۷	۰٫۳۸۷	۰٫۲۱۶	پرسپولیس	استقلال خوزستان
۷۴۵.۰	۱٫۹۰۸	۰٫۱۳۳	۰٫۳۴۷	۰٫۵۲۰	گل گهر	ذوب آهن
۰۹۹.۲	۰٫۵۹۰	۰٫۵۹۶	۰٫۳۱۱	۰٫۰۹۳	خیبر خرم آباد	تراکتور تبریز
۰۰۴.۱	۱٫۶۱۱	۰٫۲۰۷	۰٫۳۸۴	۰٫۴۰۹	چادرملو اردکان	آلومینیوم اراک
۳۲۵.۱	۱٫۲۷۷	۰٫۳۰۹	۰٫۳۹۸	۰٫۲۹۳	فولاد	پیکان تهران
۶۴۸.۱	۰٫۹۷۱	۰٫۴۲۲	۰٫۳۸۱	۰٫۱۹۷	سپاهان	استقلال تهران
۷۲۲.۱	۰٫۹۰۴	۰٫۴۵۰	۰٫۳۷۳	۰٫۱۷۷	پیکان	ملوان
۴۸۴.۱	۱٫۱۲۲	۰٫۳۶۴	۰٫۳۹۳	۰٫۲۴۳	شمس آذر قزوین	پرسپولیس
۸۲۸.۱	۰٫۸۱۲	۰٫۴۹۰	۰٫۳۶۰	۰٫۱۵۱	ذوب آهن	خیبر خرم آباد
۷۰۵.۱	۰٫۹۲۰	۰٫۴۴۳	۰٫۳۷۵	۰٫۱۸۲	استقلال خوزستان	گل گهر
۶۶۸.۱	۰٫۹۵۳	۰٫۴۳۰	۰٫۳۷۹	۰٫۱۹۱	آلومینیوم اراک	فجر سپاسی
۳۶۰.۱	۱٫۲۴۲	۰٫۳۲۱	۰٫۳۹۸	۰٫۲۸۲	تراکتور تبریز	چادرملو اردکان
۲۴۴.۱	۱٫۳۵۸	۰٫۲۸۲	۰٫۳۹۸	۰٫۳۲۰	استقلال	فولاد
۴۵۲.۲	۰٫۳۲۹	۰٫۷۴۵	۰٫۲۱۹	۰٫۰۳۷	مس رفسنجان	سپاهان
۶۹۶.۱	۰٫۹۲۸	۰٫۴۴۰	۰٫۳۷۶	۰٫۱۸۴	خیبر خرم آباد	شمس آذر قزوین
۷۱۹.۰	۱٫۹۳۸	۰٫۱۲۶	۰٫۳۴۲	۰٫۵۳۲	چادرملو اردکان	ذوب آهن
۲۰۳.۱	۱٫۴۰۰	۰٫۲۶۹	۰٫۳۹۷	۰٫۳۳۵	گل گهر	پیکان

پذیرفته شده

ادامه جدول

امتیاز مورد انتظار تیم میزبان	مورد امتیاز تیم	احتمال برد تیم میزبان	احتمال تساوی	احتمال برد تیم مهمان	تیم مهمان	تیم میزبان
۱۲۱.۱	۱/۴۸۵	۰/۲۴۳	۰/۳۹۳	۰/۳۶۴	تراکتور	استقلال خوزستان
۹۰۳.۱	۰/۷۴۹	۰/۵۱۸	۰/۳۴۸	۰/۱۳۴	پرسپولیس	استقلال تهران
۵۳۸.۰	۲/۱۶۶	۰/۰۸۰	۰/۲۹۶	۰/۶۲۳	ملوان	مس رفسنجان
۲۱۹.۱	۱/۳۸۴	۰/۲۷۴	۰/۳۹۷	۰/۳۲۹	سپاهان	فجر سپاسی
۱۵۱.۱	۱/۴۵۴	۰/۲۵۲	۰/۳۹۵	۰/۳۵۳	فولاد	آلومینیوم اراک
۳۴۳.۲	۰/۴۰۶	۰/۶۹۷	۰/۲۵۱	۰/۰۵۲	ذوب آهن	تراکتور تبریز
۷۷۰.۱	۰/۸۶۳	۰/۴۶۸	۰/۳۶۷	۰/۱۶۵	شمس آذر قزوین	گل گهر
۵۰۲.۱	۱/۱۰۶	۰/۳۷۰	۰/۳۹۲	۰/۲۳۸	پیکان	پرسپولیس
۲۲۹.۱	۱/۳۷۴	۰/۲۷۷	۰/۳۹۷	۰/۳۲۵	استقلال خوزستان	پیکان
۹۳۰.۱	۰/۷۲۶	۰/۵۲۹	۰/۳۴۴	۰/۱۲۸	آلومینیوم اراک	سپاهان
۸۲۴.۱	۰/۸۱۶	۰/۴۸۸	۰/۳۶۰	۰/۱۵۲	فجر سپاسی	چادرملو اردکان
۲۹۹.۱	۱/۳۰۳	۰/۳۰۰	۰/۳۹۸	۰/۳۰۲	استقلال	ملوان
۳۹۴.۲	۰/۳۶۹	۰/۷۱۹	۰/۲۳۶	۰/۰۴۴	مس رفسنجان	فولاد
۰۴۱.۱	۱/۵۷۱	۰/۲۱۸	۰/۳۸۸	۰/۳۹۴	تراکتور	پیکان
۳۱۵.۱	۱/۲۸۷	۰/۳۰۶	۰/۳۹۸	۰/۲۹۶	پرسپولیس	آلومینیوم اراک
۲۶۴.۱	۱/۳۳۸	۰/۲۸۹	۰/۳۹۸	۰/۳۱۳	ملوان	فجر سپاسی
۰۶۶.۲	۰/۶۱۶	۰/۵۸۳	۰/۳۱۸	۰/۰۹۹	خیبر خرم آباد	استقلال تهران
۴۹۹.۰	۲/۲۱۶	۰/۰۷۲	۰/۲۸۴	۰/۶۴۴	گل گهر	مس رفسنجان
۲۵۳.۱	۱/۳۴۹	۰/۲۸۵	۰/۳۹۸	۰/۳۱۷	چادرملو اردکان	استقلال خوزستان
۵۹۸.۱	۱/۰۱۶	۰/۴۰۴	۰/۳۸۶	۰/۲۱۰	فولاد	سپاهان
۰۰۸.۲	۰/۶۶۲	۰/۵۶۰	۰/۳۳۰	۰/۱۱۱	ذوب آهن	شمس آذر قزوین
۹۱۵.۰	۱/۷۱۱	۰/۱۸۰	۰/۳۷۵	۰/۴۴۵	استقلال خوزستان	ذوب آهن
۸۸۹.۱	۰/۷۶۱	۰/۵۱۳	۰/۳۵۰	۰/۱۳۷	آلومینیوم اراک	ملوان
۹۲۸.۱	۰/۷۲۸	۰/۵۲۸	۰/۳۴۴	۰/۱۲۸	شمس آذر قزوین	تراکتور
۲۸۸.۲	۰/۴۴۶	۰/۶۷۴	۰/۲۶۶	۰/۰۶۰	مس رفسنجان	پرسپولیس
۶۷۳.۱	۰/۹۴۹	۰/۴۳۱	۰/۳۷۹	۰/۱۹۰	فجر سپاسی	فولاد
۳۱۴.۱	۱/۲۸۸	۰/۳۰۵	۰/۳۹۸	۰/۲۹۷	پیکان	خیبر خرم آباد
۵۵۳.۱	۱/۰۵۸	۰/۳۸۸	۰/۳۸۹	۰/۲۲۳	سپاهان	چادرملو اردکان
۳۶۶.۱	۱/۲۳۶	۰/۳۲۳	۰/۳۹۷	۰/۲۸۰	استقلال	گل گهر

پذیرفته شده

جدول ۳: رتبه نهایی تیم‌ها بر اساس امتیازات مورد انتظار و امتیازات کسب شده

رتبه	نام تیم	میزبان	مهمان	مجموع	کل امتیازات
۱	استقلال تهران	۷٫۲۱۰	۵٫۶۰۳	۱۲٫۸۱۴	۸۱۴.۵۳
۲	تراکتور تبریز	۸٫۱۵۵	۵٫۸۶۵	۱۴٫۰۲۰	۰۲۰.۵۳
۳	سپاهان اصفهان	۷٫۷۳۰	۵٫۳۰۱	۱۳٫۰۳۱	۰۳۱.۵۲
۴	گل گهر سیرجان	۶٫۷۹۱	۵٫۵۲۴	۱۲٫۳۱۵	۳۱۵.۴۸
۵	چادرملو اردکان	۷٫۲۹۱	۵٫۹۲۰	۱۳٫۱۳۹	۱۳۹.۴۸
۶	فولاد خوزستان	۶٫۸۹۵	۴٫۵۹۷	۱۱٫۴۹۲	۴۹۲.۴۲
۷	ملوان بندر انزلی	۷٫۰۸۸	۶٫۱۵۹	۱۳٫۲۴۷	۲۴۷.۴۲
۸	فجر شهید سپاسی شیراز	۵٫۶۴۴	۳٫۷۴۰	۹٫۳۸۴	۳۸۴.۳۹
۹	خیبر خرم آباد	۴٫۳۷۱	۲٫۸۳۴	۷٫۲۰۵	۲۰۵.۳۶
۱۰	استقلال خوزستان	۵٫۳۰۲	۵٫۰۳۵	۱۰٫۳۳۷	۳۳۷.۳۵
۱۱	پرسپولیس تهران	۶٫۳۲۰	۴٫۷۰۱	۱۱٫۰۲۲	۰۲۲.۳۵
۱۲	پیکان تهران	۳٫۵۶۹	۴٫۴۱۴	۷٫۹۸۳	۹۸۳.۳۱
۱۳	آلومینیوم اراک	۴٫۳۸۹	۴٫۲۵۰	۸٫۶۴۰	۶۴۰.۳۰
۱۴	شمس آذر قزوین	۴٫۹۹۱	۳٫۵۹۴	۸٫۵۸۴	۵۸۴.۳۰
۱۵	ذوب آهن اصفهان	۴٫۱۲۰	۲٫۴۰۹	۶٫۵۳۰	۵۳۰.۲۵
۱۶	مس رفسنجان	۲٫۵۴۵	۱٫۴۵۳	۳٫۹۹۸	۹۹۸.۱۸

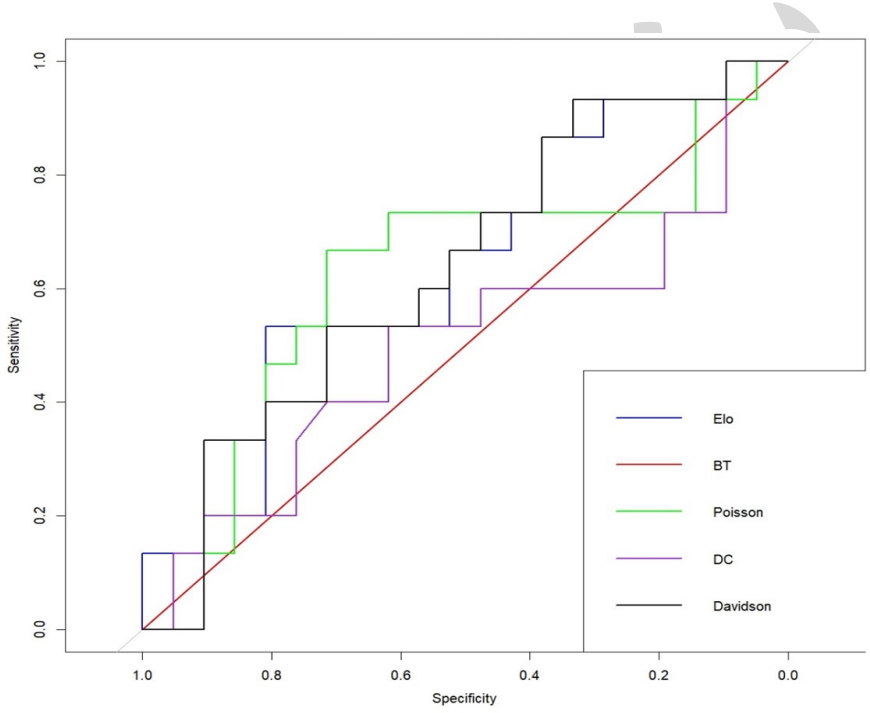
جدول ۴: نتایج ارزیابی دقت مدل BTD

نتیجه واقعی	تعداد پیش‌بینی صحیح	تعداد واقعی	دقت (%)
برد مهمان	۱۱	۱۷	۶۴/۷
مساوی	۳	۲۴	۱۲/۵
برد میزبان	۱۸	۲۱	۸۵/۷
کل	۳۲	۶۲	۵۱/۶

پذیرفته شده

جدول ۵: شاخص Log loss و شاخص Brier Score روی داده های فصل کامل ۱۴۰۴-۱۴۰۳ در مدل های مرسوم برای بیش بینی نتایج مسابقات ورزشی

مدل	Log-Loss	Brier
بردلی-تری-دیویدسون	۰٫۶۷۱	۰٫۲۳۸
Elo	۰٫۶۷۲	۰٫۲۳۹
بردلی-تری	۰٫۶۹۳	۰٫۲۵۰
پواسن	۰٫۷۰۶	۰٫۲۴۹
دیکسون-کولز	۰٫۷۶۷	۰٫۲۷۴



شکل ۱: مقایسه منحنی‌های ROC مدل‌های ELO ، بردلی-تری(BT)، پواسن(Poisson)، دیکسون-کولز (DC) بردلی-تری-دیویدسون(Davidson) بر روی داده‌های مسابقات فصل کامل ۱۴۰۴-۱۴۰۳

۴ بحث:

۱.۴ عملکرد تیم‌ها و پارامترهای مدل بردلی-تری-دیویدسون:

جدول ۱ برآورد پارامترهای مدل بردلی-تری-دیویدسون با در نظر گرفتن اثر میزبانی و پارامتر گرایش به تساوی را نشان می‌دهد. بر اساس توان نسبی برآورد شده ، تیم تراکتور تبریز با مقدار ۱/۴۵۵ بالاترین قدرت نسبی را در میان تیم‌های لیگ در فصل ناتمام ۱۴۰۵-۱۴۰۴ کسب کرد و پس از آن استقلال تهران (۱/۳۳۲)، چادرملو اردکان (۱/۰۷۸) و گل‌گهر سیرجان (۱) قرار گرفتند. در مقابل، تیم‌های مس رفسنجان (۰٫۰۷۲) ذوب‌آهن اصفهان (۰٫۱۶۵)خیبر خرم‌آباد (۰٫۳۴۹) و آلومینیوم اراک (۰٫۳۵۴) کمترین مقادیر قدرت نسبی را به خود اختصاص دادند. در داده‌های مطالعه حاضر، پارامتر گرایش به تساوی برابر با $\nu = ۱٫۳۲۲$ و پارامتر مزیت میزبانی برابر با $\alpha = ۱٫۵۳۸$ برآورد شد. از آنجا که این

پارامتر مقداری بزرگ‌تر از یک براورد شده است می‌توان نتیجه گرفت که حضور در زمین خودی در لیگ فوتبال ایران شانس کسب نتیجه مطلوب را افزایش می‌دهد. از این‌رو، انتظار می‌رود تیم‌هایی که تعداد بیشتری از مسابقات باقی‌مانده خود را در زمین خودی برگزار می‌کنند، یا در خانه با حریفانی دارای قدرت نسبی بیشتر و یا در خارج از خانه با حریفانی دارای قدرت نسبی پایین‌تر روبه‌رو می‌شوند، نسبت به تیم‌هایی که برنامه مسابقات دشوارتری دارند، امتیازات بیشتری کسب کنند.

در همین راستا یافته‌های این مطالعه نشان داد که، مجموع امتیازات مورد انتظار کسب‌شده توسط تیم‌های میزبان برابر با ۴۱.۹۱ و برای تیم‌های مهمان برابر با ۳۹.۷۱ بود. این یافته‌ها با نتایج مطالعه Silva و همکاران (۲۰۲۵) مبنی بر اینکه مزیت میزبانی،بیشتر نشان‌دهنده کاهش احتمال شکست تیم میزبان است تا افزایش احتمال پیروزی آن، همسو است [۳].

همچنین یافته‌های این مطالعه نشان داد که تیم پرسپولیس، با وجود قرار گرفتن در رتبه پنجم جدول تا قبل از ۸ اسفند ۱۴۰۴ بر اساس مسابقات برگزارشده، پس از برآورد امتیازات مورد انتظار مسابقات باقی‌مانده تا رتبه یازدهم سقوط می‌کند. بررسی برنامه مسابقات باقی‌مانده این تیم نشان داد که پرسپولیس چهار مسابقه خارج از خانه در پیش دارد که سه مورد از آن‌ها برابر تیم‌هایی با توانایی برآوردشده بالاتر از این تیم است. در مقابل، چهار مسابقه خانگی این تیم عمدتاً برابر حریفانی با توانایی برآوردشده پایین‌تر برگزار می‌شود. لذا احتمال کسب امتیاز در مسابقات خارج از خانه در برابر تیم‌های قوی‌تر کمتر از مسابقات خانگی در برابر تیم‌های ضعیف‌تر است. بنابراین، ترکیب نامتعادل برنامه مسابقات باقی‌مانده، شامل رویارویی با حریفان قدرتمند در زمین آن‌ها، موجب کاهش امتیاز مورد انتظار پرسپولیس در مقایسه با برخی رقبای مستقیم شده است. این موضوع می‌تواند یکی از دلایل اصلی افت جایگاه پیش‌بینی‌شده این تیم در جدول نهایی باشد.

علاوه بر این یافته های این مطالعه در مورد تیم ملوان بندر انزلی که از رتبه دهم در مسابقات برگزار شده تا هفتم تغییر میکند به این دلیل است که بازی های باقیمانده این تیم در خارج از خانه عمدتا با تیم هایی است که قدرت نسبی کمتری نسبت به این تیم دارند و بازی‌های باقیماند در خانه با تیم‌هایی است که قدرت نسبی آن‌ها بیشتر از این تیم است (جدول ۲).این یافته با نتایج مطالعه هان و همکاران (۲۰۰۸) در خصوص مزیت میزبانی همخوانی دارد. آن‌ها با تحلیل مسابقات نشان دادند که بازی خارج از خانه احتمال موفقیت تیم‌ها را کاهش می‌دهد [۱۵].

۲.۴ عملکرد و مقایسه مدل‌های پیش‌بینی

عملکرد پنج مدل مرسوم در پیش‌بینی نتایج مسابقات فوتبال شامل مدل‌های بردلی تری دیویدسون ،ELO بردلی–تری پایه، پواسون و دیکسون–کولز با استفاده از شاخص‌های Loss Log و Score Brier مورد ارزیابی قرار گرفت. نتایج ارائه شده در جدول ۵ نشان می‌دهد که مدل بردلی–تری–دیویدسون با مقادیر ۰/۶۷۱ برای Loss Log و ۰/۲۳۸ برای Score Brier عملکرد بهتری نسبت به سایر مدل‌ها داشته است. برای مقایسه یک مدل مرجع مبتنی بر احتمال‌های مساوی (مدل تصادفی) برای سه نتیجه ممکن ، مقادیر متناظر Score Brier و Log-loss به ترتیب ۰/۶۶۷ و ۱/۰۹۸ است. بنابراین مدل بردلی–تری–دیویدسون نسبت به مدل تصادفی به ترتیب ۶۴/۳کاهش در Brier Score و ۳۸/۹ کاهش در Log-loss نشان داد که بیانگر توانایی مناسب مدل در تخصیص احتمال‌های واقع‌بینانه به نتایج مسابقات و قدرت پیش‌بینی قابل قبول آن است. همچنین منحنی راک مدل بردلی–تری–دیویدسون نشان دهنده توانایی مناسب این مدل نسب به سایر مدل‌ها در تمایز بین نتایج مختلف مسابقات است. این یافته نشان می‌دهد که در نظر گرفتن هم‌زمان قدرت نسبی تیم‌ها، اثر میزبانی و گرایش به تساوی می‌تواند به بهبود عملکرد پیش‌بینی کمک کند (شکل ۱).

مدل بردلی–تری–دیویدسون نسبت به مدل بردلی–تری پایه که فاقد پارامتر تساوی است، بهبود قابل توجهی نشان می‌دهد. کاهش Loss Log از ۶۹۳.۰ به ۰/۶۷۱ (بهبود حدود. درصدی) و کاهش Brier Score از. به. (بهبود حدود. درصدی) بیانگر اهمیت در نظر گرفتن احتمال تساوی در پیش‌بینی نتایج مسابقات فوتبال است. این یافته با نتایج دیویدسون (۱۹۷۰) همخوانی دارد که نشان داد افزودن پارامتر تساوی به مدل بردلی–تری، قدرت پیش‌بینی مدل را به طور معناداری افزایش می‌دهد [۲].

۵ نتیجه گیری

یافته‌های این مطالعه می تواند در اختصاص دادن عادلانه سهمیه‌های بین‌المللی و سهمیه‌های سقوط به لیگ‌های پایین‌تر و همچنین معرفی قهرمان لیگ فوتبال در فصل ۱۴۰۴–۱۴۰۵مورد استفاده قرار گیرد.

۶ محدودیت‌های مطالعه

با وجود مزایای مدل بردلی–تری–دیویدسون مطالعه حاضر دارای محدودیت‌هایی نیز است. در این مدل تنها نتایج نهایی مسابقات مورد استفاده قرار گرفت و اطلاعاتی نظیر کیفیت بازیکنان، مصدومیت‌ها، تغییرات کادر فنی و شرایط زمانی مسابقات در نظر گرفته نشد. همچنین فرض شده است که توانایی تیم‌ها در طول فصل ثابت باقی می‌ماند، در حالی که در عمل قدرت تیم‌ها می‌تواند در طول زمان تغییر کند. پیشنهاد می‌شود در مطالعات آینده از مدل‌های پویاتر استفاده شود تا تغییرات توانایی تیم‌ها در طول فصل نیز لحاظ گردد. همچنین ترکیب مدل بردلی–تری–دیویدسون با اطلاعات آماری دقیق‌تر نظیر تفاضل گل، شاخص‌های عملکرد بازیکنان و متغیرهای مرتبط با شرایط مسابقه مانند دمای هوا میزان الودگی هوا، ساعت برگزاری و تعداد تماشاگران می‌تواند دقت پیش‌بینی را افزایش دهد.

مراجع

[1] Bradley, R. A. and Terry, M. E. (1952). Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, 39(3/4), 324–345.

[2] Davidson, R. R. (1970). On extending the Bradley-Terry model to accommodate ties in paired comparison experiments. *Journal of the American Statistical Association*, 65(329), 317–328.

[3] Silva, H., Oliveira, A., and Marcelino, R. (2025). Home advantage in soccer: being invincible during home matches may be the reason. *TRENDS in Sport Sciences*, 32(4).

[4] Liu, H. H., Battaglia, J., Wu, T. T. (2025). Impact of COVID-19 on English Football Premier League: Analyzing rankings and home advantage using extended Bradley-Terry models. *PLOS ONE*, 20(10), e0332627.

[5] Islam, M. M., Khan, J. R., and Raheem, E. (2017). Bradley-Terry model for assessing the performance of ten ODI cricket teams adjusting for home ground effect. *Journal of Data Science*, 16, 657–668.

[6] Carpita, M., Ciavolino, E., and Pasca, P. (2019). Exploring and modelling team performances of the Kaggle European Soccer database. *Statistical Modelling*, 19(1), 74–101.

[7] Goddard, J. (2005). Regression models for forecasting goals and match results in association football. *International Journal of Forecasting*, 21(2), 331–340.

[8] Egidi, L. and Torelli, N. (2021). Comparing goal-based and result-based approaches in modelling football outcomes. *Social Indicators Research*, 156(2/3), 801–813.

[9] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

[10] Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477), 359–378.

[11] Hvattum, L. M. and Arntzen, H. (2010). Using Elo ratings for match result prediction in association football. *International Journal of Forecasting*, 26(3), 460–470.

[12] Maher, M. J. (1982). Modelling association football scores. *Statistica Neerlandica*, 36(3), 109–118.

[13] Dixon, M. J. and Coles, S. G. (1997). Modelling association football scores and inefficiencies in the football betting market. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 46(2), 265–280.

[14] Varzesh3 website. Available at: https://www.varzesh3.com

[15] Han, B., Yang, L., Pan, P., García-de-Alcaraz, A., Yang, C., and Liu, T. (2022). The influence of removing home advantage on the Chinese Football Super League. *BMC Sports Science, Medicine and Rehabilitation*, 14(1), 208.

پذیرفته شده

Andisheye Amari, 2025, Vol. 30, No. 1
https://doi.org/10.22034/aa.2026.2092485.1038

Estimating the Final League Standings of an Incomplete Iranian Football League Using the Bradley–Terry–Davidson Model

Iman Sarbisheh[1] , **Azadeh Saki**[2] *, **Morteza Adibian**[1]

[1] *Student Research Committee, Mashhad University of Medical Sciences, Mashhad, Iran*

[2] *Associate Professor of Biostatistics, Department of Biostatistics, School of Public Health, Mashhad University of Medical Sciences, Mashhad, Iran*

Received: 2026/06/24

Accepted: 2026/11/13

Abstract

Premature termination of sports leagues due to factors such as public health crises, natural disasters, or armed conflicts creates substantial challenges for determining final team rankings. Relying solely on the current standings may not accurately reflect the true strength of teams, as the number and difficulty of the remaining matches differ across teams. This study aims to provide a statistical framework based on the Bradley–Terry–Davidson (BTD) model to estimate the relative strengths of teams, predict the outcomes of unplayed matches, and determine the final standings of the Persian Gulf Pro League in the 2024–2025 season. Data from 178 completed matches of the Iranian Persian Gulf Pro League were collected. The BTD model, incorporating both a tie propensity parameter and a home advantage effect, was fitted using the maximum likelihood estimation method. Based on the estimated team strengths, the probabilities of home win, draw, and away win were calculated for the remaining 62 matches. The expected points for each team were then computed to obtain the projected final league standings. The estimated tie propensity parameter was 1.322, while the home advantage was 1.538. Tractor Tabriz, Esteghlal Tehran, and Chadormalu Ardakan exhibited the highest relative strengths. The projections indicated that Esteghlal Tehran would finish first with 53.81 expected points and Tractor Tabriz and Sepahan Isfahan in second and third place, respectively.

Keywords: Bradley–Terry–Davidson, Iranian Football League, Team Ranking, Match Prediction, Home Advantage

Mathematics Subject Classification: 62F07, 62J15.

*Corresponding author, *sakia@mums.ac.ir*